



Kent Academic Repository

John Moore, Simon (2011) *Elucidation of the anaerobic biosynthesis of vitamin B₁₂ in Bacillus megaterium*. Doctor of Philosophy (PhD) thesis, University of Kent.

Downloaded from

<https://kar.kent.ac.uk/94539/> The University of Kent's Academic Repository KAR

The version of record is available from

<https://doi.org/10.22024/UniKent/01.02.94539>

This document version

UNSPECIFIED

DOI for this version

Licence for this version

CC BY-NC-ND (Attribution-NonCommercial-NoDerivatives)

Additional information

This thesis has been digitised by EThOS, the British Library digitisation service, for purposes of preservation and dissemination. It was uploaded to KAR on 25 April 2022 in order to hold its content and record within University of Kent systems. It is available Open Access using a Creative Commons Attribution, Non-commercial, No Derivatives (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) licence so that the thesis and its author, can benefit from opportunities for increased readership and citation. This was done in line with University of Kent policies (<https://www.kent.ac.uk/is/strategy/docs/Kent%20Open%20Access%20policy.pdf>). If you ...

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in **Title of Journal**, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

**Elucidation of the
Anaerobic Biosynthesis of
Vitamin B₁₂ in
*Bacillus megaterium***

A thesis submitted to the University of Kent for the
degree of PhD in the Faculty of Science, Technology
and Medical Studies

2011

Simon John Moore

Declaration

Name: Simon Moore

Degree: PhD-Biochemistry

Title: Elucidation of the Anaerobic Biosynthesis of Vitamin B₁₂ in *Bacillus megaterium*

No part of this thesis has been submitted in support of an application for any degree or other qualification of the University of Kent, or any other University or Institution of learning

Simon John Moore

Abstract

Vitamin B₁₂ (cobalamin) is one of Nature's most complex small molecules. It is a cobalt-containing modified tetrapyrrole that plays a number of key metabolic roles in many prokaryotic systems and is also essential to the biochemistry of higher animals. However, it is made by only certain bacteria, requiring around thirty enzymatic steps for its complete *de novo* construction. Surprisingly, the nutrient is synthesised by one of two related, though genetically distinct, pathways that represent aerobic or anaerobic routes. The anaerobic pathway has remained poorly characterised due to the instability of the pathway intermediates to oxygen and the low activity of enzymes, and is the focus of the research reported in this thesis.

The Gram-positive aerobe *Bacillus megaterium* has previously been used for the commercial production of cobalamin and has a complete anaerobic pathway. Several genes, (termed *cbi* for cobinamide biosynthesis) from the cobalamin biosynthetic pathway have been cloned and overexpressed individually within the host *B. megaterium* DSM319. One of the major bottlenecks in the anaerobic pathway is the ring contraction step, where only limited yields (< 5%) of product had previously been obtained using crude-lysate based incubations. By studying the purified *B. megaterium* CbiH₆₀ enzyme, an efficient method was developed to allow the generation of the ring-contracted product cobalt-factor IV in high yields. This breakthrough then permitted the characterisation of many of the down-stream steps in the pathway to be characterised by using other purified cobalamin biosynthetic enzymes from *B. megaterium*, by incubating combinations of purified enzymes with cobalt-factor IV and the appropriate cofactors (SAM, NADH). This has resulted in the successful step-by-step elucidation of the anaerobic cobalamin pathway, including the isolation of the long sought-after and elusive intermediates cobalt-precorrin-5, -6A, -6B, -7 and -8. An enzyme-cocktail approach has also resulted in a full *in vitro* synthesis of cobyrinic acid (the first known stable intermediate in the pathway) from 5-aminolevulinic acid has been demonstrated, involving a full 15 enzymatic steps. The exposition of the anaerobic pathway has been interpreted with respect to the chemical logic of the metabolic process and the evolution of multifaceted biochemical pathways.

Table of Contents

Abstract.....	i
Table of contents.....	ii
Figures.....	iii
Tables.....	iv
Abbreviations.....	v
Acknowledgements.....	vi

Section	Chapter 1	Page
1	Introduction to Vitamin B₁₂ and its Biosynthesis in <i>B. megaterium</i>	
1.1	Vitamin B ₁₂ and modified tetrapyrroles	2
1.1.1	Introduction to modified tetrapyrroles	2
1.1.2	Vitamin B ₁₂	4
1.2	Coenzyme B ₁₂	5
1.2.1	Adenosylcobalamin-dependent isomerases	5
1.2.2	Methylcobalamin-dependent methyltransferases	6
1.3	The biosynthesis of vitamin B ₁₂	8
1.3.1	An overview of tetrapyrrole biosynthesis	8
1.3.2	The biosynthesis of 5-aminolevulinic acid	8
1.3.3	The biosynthesis of uroporphyrinogen III	9
1.3.4	An overview of the vitamin B ₁₂ pathway	10
1.3.5	Comparisons of the aerobic and anaerobic pathways	11
1.3.6	The chelatase branch point: sirohaem and vitamin B ₁₂	15
1.3.7	The oxygen-dependent (aerobic) pathway	16
1.3.8	The oxygen-independent (anaerobic) pathway	18
1.4	<i>Bacillus megaterium</i>	24
1.4.1	Background	24
1.4.2	Recombinant protein production	25
1.4.3	<i>xyl</i> operon	25
1.4.4	Plasmids for intracellular protein production	27
1.4.5	Location of vitamin B ₁₂ genes in <i>B. megaterium</i> DSM319	30
1.5	Aims of this thesis	33
	Chapter 2	
2	Materials and Methods	
2.1	Materials	35
2.1.1	Chemicals and kits	35
2.2	Bacterial strains and plasmids	36

2.2.1	Bacterial strains	36
2.2.2	Plasmids	37
2.3	Microbiological techniques	42
2.3.1	Sterilisation	42
2.3.2	Media for growth of bacterial strains	42
2.3.3	Liquid cultures of <i>E. coli</i>	43
2.3.4	Liquid cultures of <i>B. megaterium</i>	43
2.3.5	Antibiotics and additives	43
2.3.6	Spectrometric determination of cell density	44
2.3.7	Storage of bacteria	44
2.3.8	Bioassay of vitamin B ₁₂ production	45
2.4	Molecular biology techniques	46
2.4.1	Isolation of genomic DNA from <i>B. megaterium</i>	46
2.4.2	Isolation of total RNA from <i>B. megaterium</i>	46
2.4.3	Determination of DNA concentration	46
2.4.4	Determination of RNA concentration	46
2.4.5	Plasmid purification from <i>E. coli</i>	47
2.4.6	Agarose gel electrophoresis	47
2.4.7	Amplification of DNA by polymerase chain reaction	47
2.4.8	Oligonucleotide primers for PCR	47
2.4.9	Digestion of plasmid DNA with restriction endonucleases	49
2.4.10	Purification of PCR products and linear plasmid fragments	50
2.4.11	Ligation of DNA – direct cloning	50
2.4.12	Cloning into pGEM [®] -T Easy – indirect cloning	50
2.4.13	Preparation of competent cells	50
2.4.14	Transformation of <i>E. coli</i>	51
2.4.15	DNA sequencing	51
2.4.16	Oligonucleotide primers used for sequencing	51
2.4.17	Semi-quantitative reverse transcription-PCR	52
2.4.18	Oligonucleotide primers used for sqRT-PCR	52
2.4.19	Protoplast transformation	53
2.4.20	Preparation of protoplasts	53
2.4.21	Transformation	54
2.4.22	Media	54
2.5	Protein biochemistry	56
2.5.1	Recombinant protein production in <i>E. coli</i>	56
2.5.2	Recombinant protein production in <i>B. megaterium</i>	56
2.5.3	Preparation of cell-lysates of <i>B. megaterium</i>	56
2.5.4	His ₆ -bind column purification	56
2.5.5	Anaerobic purification of proteins	57
2.5.6	Buffer exchange (desalting)	57
2.5.7	A ₂₈₀ protein concentration estimation	57
2.5.8	Gel filtration	57

2.5.9	SDS-PAGE	58
2.5.10	2D-PAGE	60
2.5.11	Bradford assay for 2D-PAGE	60
2.5.12	MALDI-TOF	60
2.5.13	Western blotting	61
2.6	Biochemical techniques	62
2.6.1	Anaerobic techniques	62
2.6.2	Preparation of cobalt(II)-factor III	62
2.6.3	Ultraviolet-visible (UV/Vis) spectrophotometry	62
2.6.4	Electron paramagnetic resonance (EPR)	62
2.6.5	Anion-exchange chromatography	63
2.6.6	RP-18 chromatography	63
2.6.7	Reverse phase HPLC	64
2.6.8	Electrospray ionisation mass spectrometry (ESI-MS)	64
2.6.9	Inductively coupled plasma mass spectrometry (ICP-MS)	64
2.6.10	Iron determination by colorimetric assay	64
Chapter 3		
3	The Regulation of Vitamin B₁₂ Biosynthesis in <i>B. megaterium</i>	
3.1	Introduction	67
3.2	Results	70
3.2.1	Cobalt stimulates growth and vitamin B ₁₂ production	70
3.2.2	CbiX ^L is regulated by vitamin B ₁₂ feedback	71
3.2.3	The <i>cobI</i> operon is regulated by a cobalamin riboswitch	73
3.2.4	Development of plasmids to overproduce the <i>cobI</i> operon	75
3.2.5	Vitamin B ₁₂ production in recombinant strains	78
3.2.6	Isolation and characterisation of a novel zinc tetrapyrrole	80
3.2.7	Identification of BMD_0328 as a putative cobalt/nickel transporter	82
3.2.8	Investigation into the role of BMD_0328 as a cobalt transporter	84
3.3	Discussion	87
Chapter 4		
4	Enzyme-capturing of Anaerobic Pathway Intermediates	
4.1	Introduction	92
4.2	Results	94
4.2.1	Cloning of the anaerobic pathway	94
4.2.2	Purification of the B ₁₂ anaerobic enzymes	95
4.2.3	Enzyme bound pathway intermediates	99
4.2.4	Ring contraction is a bottleneck in the anaerobic pathway	99
4.2.5	CbiD purifies bound to cobyrinic acid	100
4.2.6	Cobalamin reduces levels of protein-bound B ₁₂ intermediates	104
4.3	Discussion	106

	Chapter 5	
5	The Characterisation of the Anaerobic Ring Contraction Enzyme CbiH₆₀	
5.1	Introduction	109
5.2	Results	112
5.2.1	CbiH ₆₀ and CbiH ₃₀	112
5.2.2	Purification of CbiH ₆₀	115
5.2.3	CbiH ₆₀ contains a 4Fe-4S cluster	117
5.2.4	Purification of CbiH ₃₀	117
5.2.5	Preparation of the substrate, cobalt-factor III	121
5.2.6	UV-Vis of CbiH ₆₀ and CbiH ₃₀ titrations with cobalt(II)-factor III	122
5.2.7	CbiH ₆₀ catalyses the C-17 methylation and ring contraction reaction	123
5.2.8	EPR of CbiH ₆₀ and cobalt-factor III incubations	127
5.3	Discussion	135
	Chapter 6	
6	The Biosynthesis of the Corrin Macrocycle in <i>B. megaterium</i>	
6.1	Introduction	141
6.2	Results	145
6.2.1	Isolation of cobalt-precorrin-5	145
6.2.3	Isolation of cobalt-precorrin-6A	152
6.2.4	The characterisation of the corrin ring reductase, CbiJ	155
6.2.5	The biosynthesis of cobyrinic acid from cobalt-precorrin-6A	157
6.2.6	Isolation of cobalt-precorrin-7	157
6.2.7	Isolation of cobalt-precorrin-8X and cobyrinic acid	160
6.3	Discussion	171
	Chapter 7	
7	General Discussion	179
	References	189
	Appendices	202

Figures

Chapter 1 Introduction to Vitamin B₁₂ and its Biosynthesis in *B. megaterium*

- Figure 1.1 The family of modified tetrapyrroles (p 3)
- Figure 1.2 The two major classes of coenzyme B₁₂ dependent reactions (p 7)
- Figure 1.3 A diagram of the porphyrinogen and corrin rings (p 10)
- Figure 1.4 The anaerobic biosynthesis of adenosylcobalamin from uroporphyrinogen III (p 22-23)
- Figure 1.5 The xylose inducible promoter system and shuttle vectors for intracellular protein production in *B. megaterium* (p 29)
- Figure 1.6 Organisation of genes involved in the biosynthesis of adenosylcobalamin in *B. megaterium* DSM319 (p 32)

Chapter 2 Materials and Methods

- Figure 2.1 Method A - LC timetable for routine separation of tetrapyrroles using 0.1% TFA (solvent A) and acetonitrile (solvent B) (p 63)

Chapter 3 The Regulation of Vitamin B₁₂ Biosynthesis in *B. megaterium*

- Figure 3.1 Nickel affinity purification and Western blot of *B. megaterium* DSM319 wild-type cells grown under minimal conditions (p 72)
- Figure 3.2 Western blots of the CbiX^L protein in response to cobalt, vitamin B₁₂ and transition metals (p 73)
- Figure 3.3 2D-PAGE of recombinant strains of *B. megaterium* DSM319 (p 76-77)
- Figure 3.4 Western blotting, measuring the levels of CbiX^L from *B. megaterium* pSJM129 (p 79)
- Figure 3.5 Characterisation of the zinc-tetrapyrrole accumulated in *B. megaterium* DSM319 overexpressing the *cobI* operon (p 81)
- Figure 3.6 Hydrophobicity plot (Kyte and Doolittle) of BMD_0328 (p 83)
- Figure 3.7 Pairwise alignment of BMD_0328, NhlF and HoxN (p 83)
- Figure 3.8 Effect of BMD_0328 on cobalt minimal toxicity concentration (p 85)

Chapter 4 Enzyme-capturing of Anaerobic Pathway Intermediates

- Figure 4.1 HBA bound to CobH (PDB_111H) from *P. denitrificans* (p 93)

- Figure 4.2 SDS-PAGE of purification process of *B. megaterium* DSM509 cobalamin enzymes (p 98)
- Figure 4.3 Intermediates bound to enzymes from the anaerobic pathway (p 99)
- Figure 4.4 UV-Vis spectrum of intermediate isolated from CbiD (p 101)
- Figure 4.5 Bioassay for the vitamin B₁₂ intermediates cobyrinic acid and cobinamide (p 102)
- Figure 4.6 The effect of cyanocobalamin on protein-bound B₁₂ intermediates (p 105)
- Chapter 5 The Characterisation of the Ring Contraction Enzyme CbiH₆₀**
- Figure 5.1 Cloning of CbiH₆₀ and truncated CbiH₃₀ from *B. megaterium* (p 112)
- Figure 5.2 Alignment of CbiH₆₀, CbiH and CobJ sequences (p 113-114)
- Figure 5.3 UV-Vis spectra of purified CbiH₆₀^{C-His} (p 116)
- Figure 5.4 CbiH₆₀ contains a [4Fe-4S] centre (p 119)
- Figure 5.5 Purification of *B. megaterium* CbiH (p 120)
- Figure 5.6 UV-Vis spectra of reduced and oxidised cobalt-factor III (p 121)
- Figure 5.7 UV-Vis of 5 µM cobalt(II)-factor III incubated with increasing CbiH concentration (p 122)
- Figure 5.8 LC-MS of cobalt-factor IV (p 125)
- Figure 5.9 Conversion of cobalt(II)-factor III to cobalt-factor IV by UV-Vis spectroscopy (p 126)
- Figure 5.10 UV-Vis spectra of purified cobalt-factor IV (p 126)
- Figure 5.11 UV-Vis spectra and inset picture of EPR sample containing cobalt(II)-factor III, CbiH₆₀^{C-His}, SAM and dithiothreitol (p 128)
- Figure 5.12 EPR trace of cobalt(II)-factor III (100 µM) incubated on its own (A), SAM (B) or dithiothreitol (C) (p 130)
- Figure 5.13 EPR time trace of cobalt(II)-factor III (100 µM) incubated with CbiH₆₀^{C-His} (175 µM) in the presence of SAM (5 mM) (p 132)
- Figure 5.14 EPR time trace of cobalt(II)-factor III (100 µM) incubated with CbiH₆₀^{C-His} (175 µM) in the presence of SAM (5 mM) and dithiothreitol (10 mM) (p 134)
- Figure 5.15 Alignment of a short sequence domain conserved in class III methyltransferase activity (p 138)

Figure 5.16 Regions of conservation between CbiH and CobJ using the structure of the *Thermus thermophilus* Hb8 CbiH as a model (p 139)

Chapter 6 The Biosynthesis of the Corrin Macrocycle in *B. megaterium*

Figure 6.1 The missing steps of the anaerobic pathway (p 144)

Figure 6.2 The visual appearance of reactions containing different combinations of CbiH₆₀, CbiF, CbiG with 30 μ M cobalt-factor III, 1 mM SAM and 10 mM DTT (p 146)

Figure 6.3 UV-Vis spectrum and visual appearance of new intermediate produced from cobalt-factor III, CbiH₆₀, CbiF, SAM and DTT (p 147)

Figure 6.4 UV-Vis scanning spectra of 5 μ M cobalt(II)-factor III incubated with CbiH₆₀, CbiF, CbiG, 1 mM SAM and 10 mM DTT (p 148)

Figure 6.5 LC-MS UV-Vis chromatogram of CbiH₆₀, CbiF, CbiG incubations with cobalt-factor III, SAM and DTT (p 150)

Figure 6.6 UV-Vis and MS of oxidised cobalt-precorrin-5 eluting at 19.7 min (p 150)

Figure 6.7 UV-Vis spectra of cobalt-precorrin-5 incubated with CbiD and SAM (p 153)

Figure 6.8 UV-Vis and MS of cobalt-precorrin-6A eluting at 17.4 min (p 154)

Figure 6.9 UV-Vis of cobalt-precorrin-6A incubated with CbiJ and NADH (p 156)

Figure 6.10 UV-Vis chromatogram of cobalt-precorrin-5A incubated with CbiG, -D, -J, -E, -T and cofactors (1 mM SAM, 50 μ M NADH), then separated and analysed by LC-MS (p 158)

Figure 6.11 UV-Vis and MS of putative cobalt-precorrin-7 eluting at 16.5 min (p 159)

Figure 6.12 The visual appearance and UV-Vis spectra of reactions containing cobalt-precorrin-6A, CbiJ, CbiET and CbiC with SAM (p 162)

Figure 6.13 UV-Vis spectra showing the oxidation of cob(II)yrinic acid to cob(III)yrinic acid (p 163)

Figure 6.14 UV-Vis chromatogram of cobalt-precorrin-6A incubated with CbiJ, CbiET, CbiC and cofactors (5 mM SAM, 250 μ M NADH), then separated and analysed by LC-MS (p 165)

Figure 6.15 UV-Vis and MS of (2) cobalt-precorrin-8X eluting at 15 min MS (p 166)

- Figure 6.16 UV-Vis and MS of (2*) oxidised cobalt-precorrin-8X eluting at 13 min MS (p 167)
- Figure 6.17 UV-Vis and MS of (3) cob(II)yrinic acid eluting at 24 min MS (p 168)
- Figure 6.18 UV-Vis and MS of (4) unknown species produced by CbiJ, CbiET, CbiC and cobalt-precorrin-6A, eluting at 26 min (p 169)
- Figure 6.19 UV-Vis and MS of (5) unknown species produced by CbiET, CbiC and cobalt-precorrin-6A, eluting at 11.9 min (p 170)
- Figure 6.20 Protein-ligand docking model of cobalt-precorrin-5B and *A. fulgidus* CbiD (PDB_1SR8) (p 174)

Chapter 7: General Discussion

- Figure 7.1. Theoretical model for tetrapyrrole pathway evolution from an ancestral semi-biosynthetic sirohaem pathway (p 187)

Tables

Chapter 1 Introduction to Vitamin B₁₂ and its Biosynthesis in *B. megaterium*

- Table 1.1 The two sub-classes of adenosylcobalamin-dependent isomerases (p 6)
- Table 1.2 Table 1.2. Comparison of the aerobic and anaerobic pathways, adapted from (Warren *et al.*, 2002) (p 14)

Chapter 2 Materials and Methods

- Table 2.1 Bacterial Strains (p 36)
- Table 2.2 Plasmids (p 37-41)
- Table 2.3 Antibiotics and additives (p 44)
- Table 2.4 Oligonucleotides for PCR (p 48)
- Table 2.5 PCR reaction mix (p 49)
- Table 2.6 PCR protocol for standard amplifications (Up to 2500 bp) (p 49)
- Table 2.7 Sequencing oligonucleotides (p 51-52)
- Table 2.8 Oligonucleotide primers used for sqRT-PCR (p 53)
- Table 2.9 Denaturing molecular mass marker Dalton VII (p 58)
- Table 2.10 Recipe for SDS gels (p 58)
- Table 2.11 IEF method (p 59)

Chapter 3 The Regulation of Vitamin B₁₂ Biosynthesis in *B. megaterium*

- Table 3.1 Cobalamin riboswitches in *B. megaterium* DSM319 (p 69)
- Table 3.2 Effect of cobalt on growth and vitamin B₁₂ production (p 71)
- Table 3.3 Transcript levels in *B. megaterium* DSM319 (p 74)
- Table 3.4 Vitamin B₁₂ produced in plasmid strains of *B. megaterium* DSM319 (p 79)
- Table 3.5 Analysis of cobalt and nickel levels in *B. megaterium* DSM319 plasmid strains (p 86)

Chapter 4 Enzyme-capturing of Anaerobic Pathway Intermediates

Table 4.1 Cloning of anaerobic pathway genes into plasmids (p 94)

Table 4.2 Purification summary of the *B. megaterium* DSM509 cobalamin enzymes produced and purified from *B. megaterium* DSM319 (p 97)

Table 4.3 Summary of enzyme-intermediate trapping experiments (p 103)

Chapter 5 The Characterisation of the Ring Contraction Enzyme CbiH₆₀

Table 5.1 Reactions for EPR analysis (p 127)

Chapter 6 The Biosynthesis of the Corrin Macrocycle in *B. megaterium*

Table 6.1 Reaction list to study the activity of CbiJ, CbiET and CbiC (p 160)

Table 6.2 Summary of the anaerobic pathway intermediates from factor II to cobyrinic acid (p 176-177)

Abbreviations

A₂₈₀	Absorbance at 280 nm
Abs	Absorbance
ALA	5-aminolevulinic acid
ALAS	ALA synthase
APS	Ammonium persulphate
AU	Absorbance units
B₁₂	Vitamin B ₁₂ , Cobalamin, Cyanocobalamin
BSA	Bovine serum albumin
CcpA	Catabolite control protein
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
cre	Catabolite response element
Da	Dalton
DAD	Diode array detector
DMB	Dimethylbenzimidazole
DMF	Dimethyl-formamide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
dsDNA	Double stranded deoxyribonucleic acid
ε	Extinction coefficient
EPR	Electroparamagnetic resonance
EDTA	Ethylenediaminetetraacetic acid
ESI-MS	Electrospray ionisation mass spectrometry
FAD	Flavin adenine dinucleotide
FPLC	Fast performance liquid chromatography
gap	Glyceraldehyde-3-phosphate dehydrogenase
G+C (%)	Genome percentage of guanine and cytosine residues
GluRS	Glutamyl-tRNA synthase
GFP	Green fluorescence protein
GSA	Glutamate-1-semialdehyde
HBA	Hydrogenobyric acid
His₆-tag	Hexahistidine tag
HMB	Hydroxymethylbilane
HRP	Horseradish peroxidase
HPLC	High performance liquid chromatography
HPr	Histidine containing phosphocarrier protein
ICP-MS	Inductively coupled plasma mass spectrometry
IEF	Isoelectric focusing
IPTG	Isopropyl β-D-1-thiogalactopyranoside
K	Kelvin
LB	Luria-Bertani
MALDI-TOF	Matrix assisted laser desorption/ionization time-of-flight
MCS	Multiple cloning site
MD-01	Molecular dimensions - structure screen 1 (50 conditions)
MPD	2-methyl-2,4-pentanediol
mRNA	Messenger RNA
MS/MS	Tandem mass spectrometry
MSR	Methionine synthase reductase

Abbreviations

MTC	Minimum toxicity concentration
NAD	β -Nicotinamide adenine dinucleotide
NADH	β -Nicotinamide adenine dinucleotide hydrate (reduced)
NADPH	β -Nicotinamide adenine dinucleotide 3'-phosphate hydrate (reduced)
NiR/SiR	Nitrite/sulphite reductase
NMR	Nuclear magnetic resonance
NLS	Nucleotide loop assembly
OD	Optical density
O_L/O_R	P _{xyIA} operator sequences, left and right
PAGE	Polyacrylamide gel electrophoresis
PBG	Porphobilinogen
PBGS	Porphobilinogen synthase
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PEP	Phosphoenolpyruvate
PVDF	Polyvinylidene fluoride
PTS	Phosphoenolpyruvate: glucose phosphotransferase system
P_{xyIA}	Xylose-inducible promoter
RBS	Ribosome binding site
RNA	Ribonucleic acid
SAM	S-adenosyl-L-methionine
SAH	S-adenosyl-L-homocysteine
SDS	Sodium dodecyl sulphate
sqRT-PCR	Semi-quantitative reverse transcription polymerase chain reaction
ssRNA	Single stranded ribonucleic acid
Strep	Streptavidin
SUMT	S-adenosyl-L-methionine uroporphyrinogen III methyltransferase
SYBR	SYBR green I fluorescent DNA binding dye
TAE	Tris-acetate-EDTA
TBST	Tris buffered saline-Tween [®] -20
TEMED	N, N, N', N'-tetramethylethylenediamine
TEV	Tobacco etch virus
TFA	Trifluoroacetic acid
TLC	Thin layer chromatography
TPP	Thiamine pyrophosphate
tRNA	Transfer ribonucleic acid
2D-PAGE	Two-dimensional SDS-PAGE
UV	Ultra violet
v/v	Volume/volume
w/w	Weight/weight
w/v	Weight/volume
X-gal	Bromo-chloro-indolyl-galactopyranoside
XLSA	X-linked sideroblastic anaemia

For the completion of this PhD thesis I would like to thank the following. I pay sincere gratitude to Professor Martin Warren. Martin is one of the most influential and inspirational people that I have met. He has given me confidence and has taught me that success is achieved through hard work and perseverance. I have learnt many skills from his experience, especially his desire and ideas that drive every PhD. Moreover he is very supportive and I rate him as an exceptional supervisor.

Next I would like to thank Dr Rebekka Biedendieck. She is a kind friend that has been highly influential on my work and continued development. Not only did she show me how to work with *B. megaterium*, but most importantly taught me to “schnell, schnell”, dithering only delays shall I say. She is also very meticulous, something I can only try to be. I wish her the very best in her future career, as she undoubtedly can become a Professor one day. I will always stay in touch, thank you so much for your patience and help.

My thanks goes out to all the members of the Warren lab. Dr Steffi Frank was my initial co-supervisor along with Rebekka. She is very kind and always has good advice. She provided me with the initial training to begin my research into the anaerobic pathway. Also sincere thanks to Dr Evelyne Deery the “master of cloning”, with all her experience and interest in all things B₁₂ she is great asset to the lab and has provided me with much needed help. Thanks to Dr Andrew Lawrence who offers a different perception. He has a mind to solve challenging questions, whilst also providing valuable technical support. Dr Amanda Brindley also performed a lot of the earlier research into the anaerobic pathway along with Evelyne. I miss working with her particularly as some of her ideas helped me tackle the CbiH problem. Also thanks to Dr Susi Schroeder for her constant assistance and help in the lab.

Thanks to Dr Steve Rigby who can only be described as legendary. Steve talks the hind legs off a donkey, but he is very friendly and gave essential advice to a “non-physicist/numpty biochemist” in the dark art of EPR. Thanks so much for some fascinating experiments. Also thanks to Dr Martin Gamer for protoplast transformations. In addition thanks to Professor Andy Munro, Dr Kirsty McClean

(redox potentiometry) and Professor Nigel Robinson, Dr Kevin Waldron (ICP-MS) for their valuable assistance in this research.

I turn my attention to individuals that have driven my enthusiasm for science. I particularly thank my neighbours Robin and Sue Darvell. Robin is a retired eye surgeon and has provided invaluable help throughout my life. Together they have also helped me to purchase many books that have aided in my development. In addition school teachers are sometimes overlooked. In particular in year 10, I met a new maths teacher Mark Towlson. He is completely crazy, funny and taught me to play chess. His favourite phrase was “Winning isn’t everything; it’s the only thing”. He is one of the finest teachers I have met and gives students confidence and inspiration to achieve, a priceless skill for a teacher.

Finally I would like to thank my family and friends. Particularly my long-term partner Emma, She has always supported me and hopefully one day we will be millionaires! To my father Roger I thank you for all your support. Finally I would like to dedicate this PhD thesis in the memory of my mum, Angela Moore.



Chapter 1

An Introduction to Vitamin B₁₂ and its Biosynthesis in *B. megaterium*

1.1 Vitamin B₁₂ and Modified Tetrapyrroles

1.1.1 Introduction to Modified Tetrapyrroles

The modified tetrapyrrole family includes amongst its members a diverse range of prosthetic groups that are indispensable to the biochemistry of life. A distinguishing feature of these molecules is the unique colour they exhibit, which together with their function gives rise to their title as “The pigments of life” (Battersby, 2000). Each tetrapyrrole is derived from the primogenitor uroporphyrinogen III that serves as the main macrocycle from which each tetrapyrrole is derived and is distinguished by the extent and type of modification that occurs. Moreover, at some point during each biosynthetic process, a metal ion is inserted into the core of the ring macrocycle, which in most cases is of central importance to the chemistry and function of the molecule. Haem and sirohaem both require a Fe²⁺ ion to be inserted, but differ significantly in their degree of modifications and the biochemical reactions for which they are required. Haem is a red pigment that is multifunctional and ubiquitous in all forms of life. In summary it is required for oxygen transport, respiration and electron transport to name a few. Sirohaem is a green isobacteriochlorin that is required as a cofactor in nitrite and sulphite reductases (NiR/SiR), which catalyse a 6-electron reduction of nitrite or sulphite to ammonia or sulphide, respectively (Crane & Getzoff, 1996; Einsle *et al.*, 2002). Chlorophylls and bacteriochlorophylls are green pigments that are involved in the harvesting of light and its conversion into chemical energy. Chlorophyll is the most abundant tetrapyrrole on Earth and can be observed from outer space. The yellow coenzyme F₄₃₀ is uniquely found in methanogenic archaea and is the coenzyme for methyl coenzyme M reductase, which catalyses the reduction of methyl coenzyme M to methane and coenzyme M (Pfaltz *et al.*, 1987). Finally the most complex of all tetrapyrroles is vitamin B₁₂. This is a pink pigment that uniquely has a smaller metal binding cavity (ring contracted) into which a cobalt ion is bound. The biologically active forms of B₁₂ include adenosylcobalamin and methylcobalamin, which are required for rearrangement and methyl transfer reactions, respectively. Structures of these tetrapyrroles are shown in Figure 1.1.

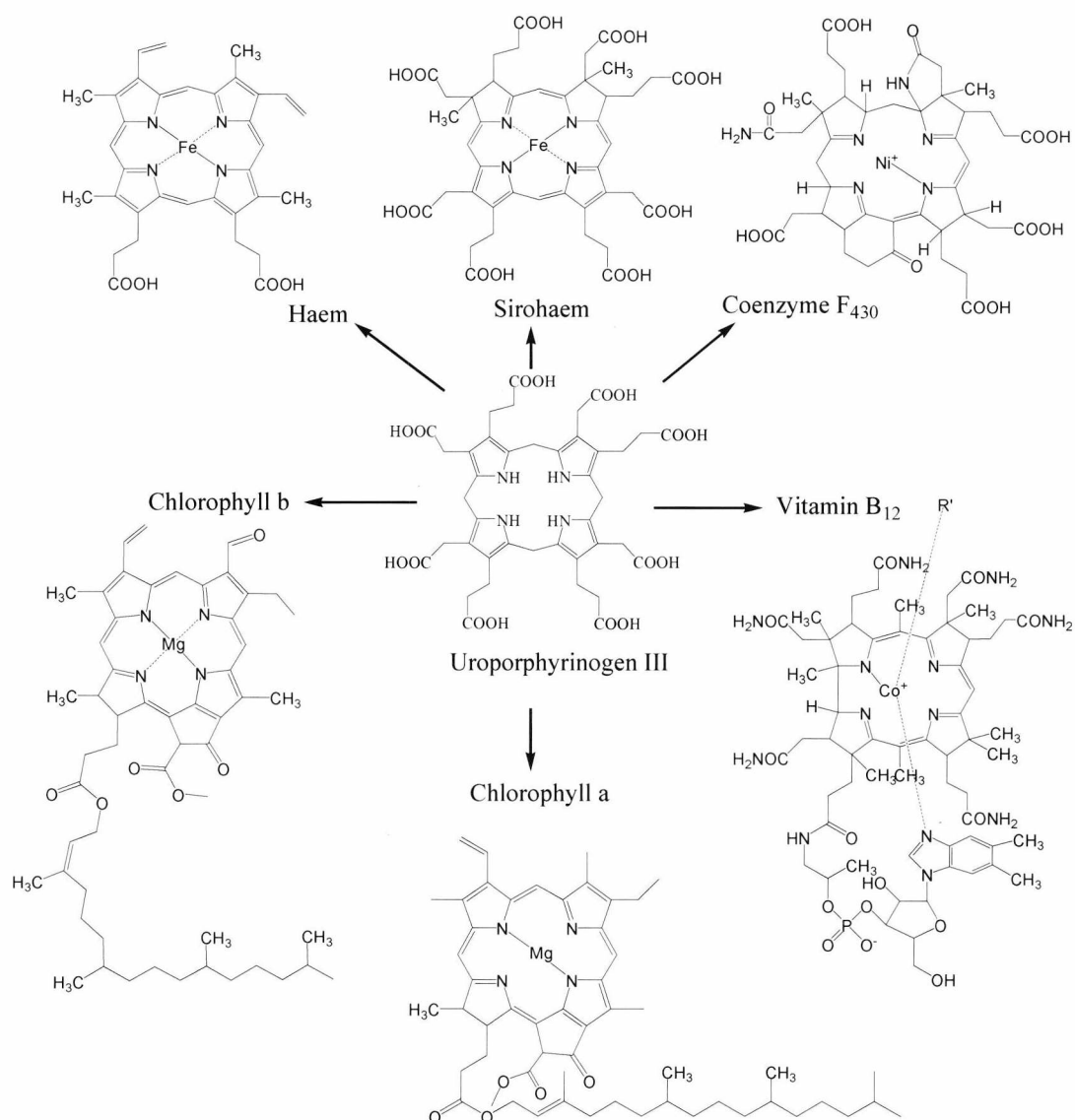


Figure 1.1. The family of modified tetrapyrroles. Uroporphyrinogen III is the primogenitor, from which all modified tetrapyrroles are derived. Each tetrapyrrole has a metal ion inserted within the ring macrocycle, with haem and sirohaem requiring $-Fe^{2+}$, coenzyme F₄₃₀ - Ni^{+} , chlorophyll - Mg^{2+} and vitamin B₁₂ - Co^{2+} . The R' upper ligand in cobalamin is variable with either cyano, hydroxy, methyl or adenosyl groups attached.

1.1.2 Vitamin B₁₂

During the 1920's Minot and Murphy first described the treatment of pernicious anaemia with crude liver extracts (Minot & Murphy, 1926). Following its isolation and crystallisation in 1948, the molecule was named vitamin B₁₂ (cyanocobalamin) (Rickes *et al.*, 1948; Smith, 1948). Shortly afterwards its structure was deduced by Dorothy Hodgkin using X-ray crystallography, revealing Nature's most complex of small molecules (Hodgkin *et al.*, 1955). Vitamin B₁₂ is a modified tetrapyrrole with a unique cobalt atom coordinated within four equatorial nitrogen ligands, which are donated by the four surrounding pyrrole groups A-D. It has both upper (β) and lower (α) axial ligands. During extraction the upper ligand is displaced by a cyano group and thus cyanocobalamin is referred to as vitamin B₁₂. However, the biological forms of B₁₂ have a different form of the upper ligand, usually either a methyl or adenosyl group, which are adenosylcobalamin or methylcobalamin, respectively. The lower ligand is 5,6-dimethylbenzimidazole (DMB), which is attached within a nucleotide loop. In some organisms pseudocobalamin is produced and contains adenine as the lower ligand.

Central to the function of cobalamin is the unique nature of the Co-C bond, which is vital for reactions involving methyl transfer, rearrangement and degradation of inorganic chemicals such as propanediol and ethanolamine. Mammals are auxotrophic for vitamin B₁₂ and have to acquire it from their diet, with deficiencies of coenzyme B₁₂ resulting in diseases such as megaloblastic anaemia and congenital defects of the newborn (Wadia, 2000). Plants and fungi are able to survive without vitamin B₁₂, with all other species having to synthesise it, or acquire from their environment. The sole source of vitamin B₁₂ is from bacteria, which have also provided the solution for elucidating its remarkable biosynthesis. As a highly complex molecule, its biosynthesis demands approximately thirty enzymatic steps and can comprise up to 1% of an entire microorganisms genome (Roth *et al.*, 1993). In comparison, the complete chemical synthesis of vitamin B₁₂ was elaborately deduced in the laboratories of Woodward and Eschenmoser, requiring approximately sixty steps (Eschenmoser & Wintner, 1977; Woodward, 1973). Although ingenious, unfortunately this process ultimately remains impractical due to the high cost and low yield (Martens, 2002). In humans vitamin B₁₂ is mainly

acquired through dietary consumption. However for therapeutic applications and as food supplements approximately 10 tonnes per annum is produced and purified from a range of industrially cultured bacterial species (Marten et al, 2002).

1.2 Coenzyme B₁₂

Coenzyme B₁₂ is a highly versatile cofactor in biochemical reactions, which relies mainly on the unique nature of the organometallic Co-C bond. There are currently three main classes of cobalamin dependent reactions. The adenosylcobalamin-dependent isomerases are the largest family with a variety of sub-classes. The majority of these enzymes are involved in fermentation pathways in bacteria, with the important exception being the methylmalonyl-CoA mutase, which is essential in animals (Banerjee & Ragsdale, 2003). Furthermore there is a coenzyme B₁₂ dependent ribonucleotide reductase class II, which is found in some bacteria and is required for DNA synthesis and repair. The methylcobalamin-dependent methyltransferases are the second largest family and are required in amino acid and one-carbon metabolism, whilst some anaerobic bacteria can also utilise this coenzyme for CO₂ fixation (Banerjee & Ragsdale, 2003). Finally coenzyme B₁₂ is also utilised by some anaerobic bacteria for the breakdown of chlorinated aromatics and aliphatic organics.

1.2.1 Adenosylcobalamin-dependent isomerases

The adenosylcobalamin-dependent isomerases are a family of enzymes that catalyse the rearrangement of a variable substituent (R' group) and a hydrogen atom, between two adjacent carbons. They can be classified into two groups that depend on the mode of cofactor binding with the lower axial ligand DMB. Normally under physiological pH the cobalt ion is bonded to DMB. In this state the bonding is considered base-on (DMB-on) and belongs to the Class II enzymes. Class I enzymes replace the DMB with a histidine residue to attach to the cobalt ion. This is named the His-on conformation (or base-off), with the histidine being present in a highly conserved DXHXXG motif (Banerjee & Ragsdale, 2003). Enzymes belonging to these classes are summarised in Table 1.1. The unique Co-C bond in adenosylcobalamin is highly labile and is essential to the rearrangement reactions it catalyses (Banerjee & Ragsdale, 2003). The reaction process is summarised in

Figure 1.2, where the methylmalonyl-CoA mutase is shown as an example. It begins with substrate binding to the holoenzyme, before the Co-C bond of adenosylcobalamin is cleaved homolytically. The hydrogen from the substrate is transferred to the adenosyl radical to form 5'-deoxyadenosine. The substrate radical is then rearranged, with the R' group moving to the carbon previously attached to the hydrogen. This yields a product radical. The hydrogen from 5'-deoxyadenosine is then transferred to the free bond on the product radical, producing the product and regenerating the adenosylcobalamin cofactor.

Class I (His-on family)	Class II (DMB-on family)
Methylmalonyl-CoA mutase	Propanediol dehydratase
Glutamate mutase	Glycerol dehydratase
Methyleneglutarate mutase	Ethanolamine ammonia lyase
Isobutyryl-CoA mutase	Ribonucleotide reductase
Lysine 5,6 aminomutase	(Class II)

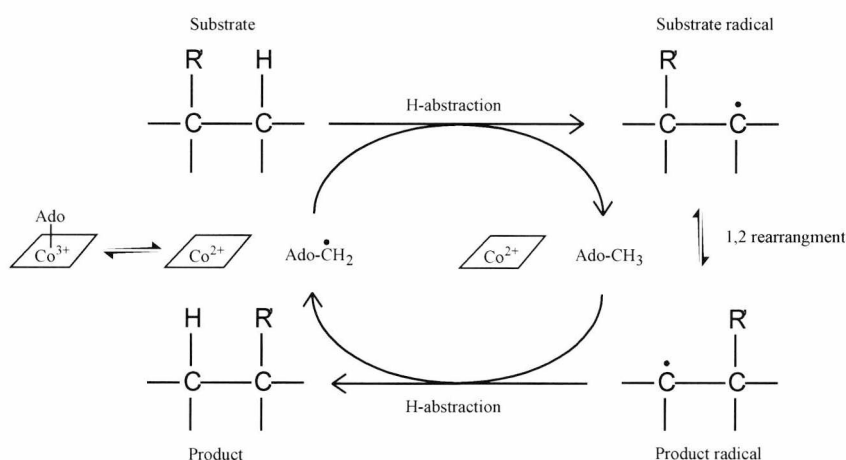
Table 1.1. The two sub-classes of adenosylcobalamin-dependent isomerases. The class I family contain the DXHXXG motif, with the histidine replacing the bond between DMB and the cobalt ion. The class II family have no specific motif, but instead share a common cofactor binding mode, with DMB bound as the lower ligand.

1.2.2 Methylcobalamin-dependent methyltransferases

The methylcobalamin-dependent methyltransferases represent the second largest class and are involved in amino acid and one-carbon metabolism and in some anaerobes, CO₂ fixation. The most widespread methylcobalamin-dependent enzyme is the methylcobalamin-dependent methionine synthase MetH. This is a large enzyme (~140 kDa) that catalyses the transfer of a methyl group from methyltetrahydrofolate to homocysteine, forming methionine and tetrahydrofolate. MetH is composed of four modular domains to bind homocysteine, methyltetrahydrofolate, methylcobalamin and *S*-adenosyl-L-methionine (SAM). This reaction, summarised in Figure 2, requires heterolytic bond fission to produce a Co⁺ and a CH₃⁺ species, with the methyl group transferred to homocysteine to produce methionine. The methylcobalamin cofactor is then regenerated by

acquiring a new methyl group from methyltetrahydrofolate. This cycle can occur approximately 2000 times, before the Co^+ species is oxidised to Co^{2+} (Drummond *et al.*, 1993). To reactivate the cofactor, SAM provides the methyl group with the reduction provided by a reduced flavodoxin in *E. coli* (Fujii & Huennekens, 1974) or methionine synthase reductase (MSR) in humans (Olteanu & Banerjee, 2001).

A. Adenosylcobalamin-dependent rearrangement mechanism (Methylmalonyl-CoA mutase example)



B. Methylcobalamin-dependent methyltransferase mechanism (Methionine synthase example)

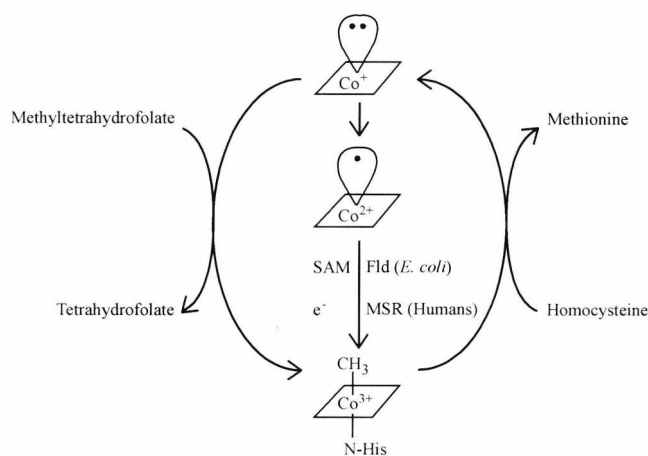


Figure 1.2. The two major classes of coenzyme B₁₂ dependent reactions. A. The adenosylcobalamin dependent isomerase reaction mechanism is provided using the methylmalonyl-CoA mutase as an example. **B.** The methionine synthase reaction is provided as an example of a methylcobalamin-dependent methyltransferase reaction. Illustrations modified from (Banerjee & Ragsdale, 2003).

1.3 The Biosynthesis of Vitamin B₁₂

1.3.1 An overview of tetrapyrrole biosynthesis

Tetrapyrrole biosynthesis is a large branched pathway that begins with the synthesis of 5-aminolevulinic acid (ALA), which is itself made via two pathways. The next steps leading to the formation of uroporphyrinogen III are conserved within all organisms, with this being the last intermediate common to all modified tetrapyrroles. From this branch point haem, haem *d1*, sirohaem, chlorophyll, coenzyme F₄₃₀, vitamin B₁₂ can all be synthesised. For haem and chlorophyll biosynthesis the last common intermediate is protoporphyrin IX. In contrast, for haem *d1*, sirohaem, coenzyme F₄₃₀ and vitamin B₁₂, precorrin-2 is the last shared intermediate. Although no organism can synthesise all of these tetrapyrroles, some organisms can make four different derivatives. Little is known about the control and regulation that dictates which tetrapyrrole is synthesised at a given time point. However, it seems likely that the availability of the metals that are required may determine which tetrapyrrole is synthesised (Raux *et al.*, 2000). The following sections shall now concentrate on the biosynthesis of the intermediates ALA, uroporphyrinogen III and adenosylcobyrinic acid.

1.3.2 The biosynthesis of 5-aminolevulinic acid

ALA is the precursor for all modified tetrapyrroles. This small molecule provides all the nitrogen and carbon atoms required for the synthesis of the macrocycle. ALA itself is synthesised by one of two pathways. The first to be discovered was the “Shemin” pathway, which involves the condensation of succinyl-CoA and glycine, with the subsequent loss of CO₂ (Neuberger & Scott, 1953; Shemin *et al.*, 1955). This reaction is catalysed by ALA synthase (ALAS), which is found in animals, fungi and the α -proteobacteria. The first crystal structure of ALAS (from *Rhodobacter capsulatus*) was published recently in complex with its substrates (glycine and succinyl-CoA) and cofactor pyridoxal phosphate (Astner *et al.*, 2005). The sequence identity between human erythroid ALAS and the published structure is 49% and has provided insights into X-linked sideroblastic anaemias (XLSA's), which are caused by mutations in the human erythroid ALAS. For the remaining bacterial species, archaea and plants, a different route for ALA synthesis is required

and is named the C5-pathway. This route begins with the substrate glutamyl-tRNA, which is required for both protein and tetrapyrrole biosynthesis. Glutamyl-tRNA is synthesised in an ATP dependent process catalysed by the enzyme glutamyl-tRNA synthase (GluRS). The first unique step in the C5-pathway is an NADPH-dependent reduction of glutamyl-tRNA by HemA, to produce the unstable intermediate glutamate-1-semialdehyde (GSA). At this point GSA is released by the enzyme via a “backdoor” exit and transferred to HemL, which is docked onto the HemA enzyme (Luer *et al.*, 2007; Moser *et al.*, 2001). This is essential as GSA is highly labile and the close association of HemA and HemL represents an excellent example of metabolic channelling.

1.3.3 The biosynthesis of uroporphyrinogen III

This stage of tetrapyrrole biosynthesis is shared between all forms of life and involves the construction of the tetrapyrrole nucleus uroporphyrinogen III, which is in the form of an architectural blueprint for all modified tetrapyrroles. Uroporphyrinogen III is produced from ALA by three steps, using the enzymes HemB, HemC and HemD [for a detailed review see (Heinemann *et al.*, 2008)]. The first step requires the asymmetrical condensation of two molecules of ALA to produce porphobilinogen (PBG) by porphobilinogen synthase (PBGS) HemB. Following this four molecules of PBG are linked together by porphobilinogen deaminase (PBGD) HemC, which releases four molecules of ammonia. The product of this reaction is the linear tetrapyrrole hydroxymethylbilane (HMB). To produce uroporphyrinogen III the enzyme uroporphyrinogen III synthase (HemD) catalyses the inversion of ring D and forms a methylene bridge between rings A and D.

The type III isomer of uroporphyrinogen is the only asymmetric form and results from the inversion of ring D. The cyclisation of the macrocycle forms what is known as a porphyrinogen ring. Later in the biosynthesis of vitamin B₁₂, this is converted into a corrin ring whereby the methylene bridge between rings A and D is uniquely removed during a process called ring contraction (Figure 1.3).

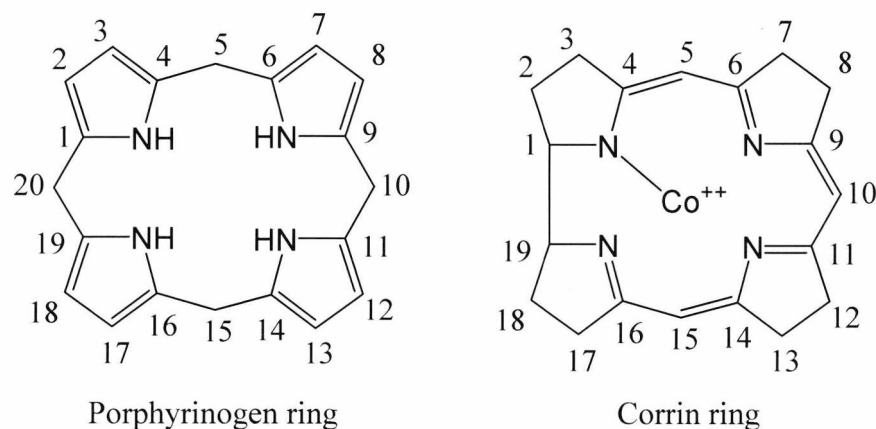


Figure 1.3. A diagram of the porphyrinogen and corrin rings. The carbon atoms are labelled 1-20, starting at ring A. All the following methylations or reductions refer to these carbon positions. The C-20 methylene bridge between ring A and D is removed during the ring contraction to produce the corrin ring.

1.3.4 An overview of the vitamin B₁₂ pathway

Adenosylcobalamin is synthesised along a biochemical pathway that is comprised of 30 enzymes starting from ALA and which is found in only certain eubacteria and archaea. Essentially the elaborate construction of cobalamin can be simplified to three stages. This firstly involves the synthesis of the ring macrocycle uroporphyrinogen III from ALA. The second stages see the (aerobic and anaerobic) biosynthesis of adenosylcobyrinic acid from uroporphyrinogen III, whilst the final stage completes the biosynthesis of adenosylcobalamin via the attachment of the lower (α) axial ligand DMB. This final stage also involves the nucleotide loop assembly (NLA) pathway which catalyses the covalent linkage of adenosylcobyrinic acid, 1-aminopropanol, α -ribose and DMB. This forms the nucleotide loop and attachment of the lower ligand base DMB, forming the final product adenosylcobalamin (Maggio-Hall & Escalante-Semerena, 1999).

The formation of adenosylcobyrinic acid from uroporphyrinogen III is referred to as the biosynthesis of the corrin macrocycle. Here two pathways (aerobic and anaerobic) have evolved and share a requirement of eight methylations, ring contraction, extrusion of the *meso* C-20 position, corrin ring reduction, decarboxylation, methyl group rearrangement, six side-chain amidations, cobalt insertion, cobalt reduction ($\text{Co}^{2+} \text{e}^- \rightarrow \text{Co}^+$) and attachment of the adenosyl upper (β) ligand group to the cobalt ion.

Early experiments with pulse-labelling NMR, established that the order of methylations in both pathways occurred as follows: C-2, C-7, C-20, C-17, C-11, C-1, C-5 and C-15 (Scott *et al.*, 1989; Uzar *et al.*, 1987). However, it should be noted that an element of uncertainty remains over the exact order of the final two methylations. This initial work allowed the current nomenclature for the pathway intermediates to be devised. The series of intermediates were named precorrin-*n*, with *n* referring to the number of methylations that had occurred (Uzar *et al.*, 1987). For example, the first step in the pathway results in the addition of two methyl groups at C-2 and C-7 to produce precorrin-2 from uroporphyrinogen III. Later it was established that an alternative anaerobic pathway existed and that this required an early cobalt insertion. The names of these intermediates were given the prefix cobalt-precorrin-*n*. This is shown as a reaction sequence in Table 1.2.

1.3.5 Comparisons of the aerobic and anaerobic pathways

Two major pathways for the biosynthesis of cobalamin have evolved. These are cryptically referred as the aerobic (oxygen-dependent) and anaerobic (oxygen-independent) pathways, with the latter most likely evolving the earliest when the Earth's atmosphere was devoid of molecular oxygen (Raux, 1999). However, it was the aerobic pathway that was first elucidated. Early interest in the biosynthesis of vitamin B₁₂ in the 1980-1990's, began with an industrial overproducing strain of *Pseudomonas denitrificans*, which houses a complete aerobic pathway (Cameron, 1989). By using a combination of microbiology and recombinant DNA technology, the majority of the genes for cobalamin biosynthesis were isolated (Cameron *et al.*, 1991; Crouzet *et al.*, 1990a; Crouzet *et al.*, 1990b; Crouzet *et al.*, 1991). To begin with these genes were given the prefix *cob*, with the letter following the prefix indicating the order of genes in the operon from which they were isolated, rather than representing the order of the biosynthetic pathway (Warren *et al.*, 2002). Based on the initial industrial strain, genetic engineering was used to develop new strains with increased levels of enzymes from the cobalamin biosynthetic pathway, which lead to increased metabolic flux (Debussche *et al.*, 1993). This technique allowed the isolation and characterisation of individual enzymes and intermediates (Blanche *et al.*, 1991a; Blanche *et al.*, 1991b; Blanche *et al.*, 1992a; Blanche *et al.*, 1992b; Blanche *et al.*, 1992c; Debussche *et al.*, 1991; Debussche *et al.*, 1992; Debussche *et al.*, 1993; Thibaut *et al.*, 1990b; Thibaut *et al.*, 1990c; Thibaut *et al.*, 1992). This

pioneering work established the order of reactions in the pathway (see Table 1.2) and also provided a fascinating insight in to how metabolic engineering can aid in the development of strains to produce commercially important metabolites (Martens *et al.*, 2002).

Following this, cobalamin genes from *Salmonella enterica* serovar Typhimurium and *Bacillus megaterium* were isolated, [see (Raux *et al.*, 1998a; Raux *et al.*, 1998b; Roth *et al.*, 1993)]. Analysis of these genes in *Escherichia coli* revealed that vitamin B₁₂ biosynthesis could occur, but only under anaerobic conditions. In addition to the isolation of early cobalt intermediates it established that an alternative anaerobic pathway must exist. These genes were given the prefix *cbi* (cobinamide synthesis) to distinguish the pathway from the aerobic pathway (Warren *et al.*, 2002). However, it was soon realised that the anaerobic pathway would provide a far greater challenge to elucidate in comparison to the ease and speed that accompanied the unravelling of the aerobic pathway. It was initially believed that this was due to the unstable nature of the intermediates of the anaerobic pathway as they required the early insertion of cobalt into oxygen sensitive intermediates (Balachandran *et al.*, 1994). In comparison to the aerobic pathway cobalt insertion occurs at a much later stage, and is a key difference between the two pathways. In the aerobic pathway cobalt is inserted into a stable intermediate called hydrogenobyric acid *a,c* diamide. Cobalt chelation into this ring contracted intermediate is catalysed by an ATP-dependent cobaltochelatease, which is a large multi-enzyme complex of CobN, CobS and CobT (Heldt *et al.*, 2005). In contrast the anaerobic pathway utilises a small single enzyme (CbiX or CbiK) to insert cobalt at an early stage in the pathway into a pyrrolic ring system.

Later it was established that another key difference between the two pathways occurred with the ring contraction process. This reaction involves methylation at C-17 and the removal of the methylene bridge between the rings A and D. This event forms a new macrocycle, whereby the pyrrolic ring system is converted into the corrin ring (see Figure 3). This is unique in tetrapyrrole biosynthesis and is believed to allow tighter coordination of the cobalt ion within the macrocycle (Warren *et al.*, 2002). Significantly the timing of the ring contraction process occurs at the same relative position in the two pathways. However, the fundamental difference between

the pathways is that the aerobic pathway requires molecular oxygen and two enzymes CobG and CobJ (Schroeder *et al.*, 2009). The anaerobic pathway is much simpler, requiring only a single enzyme CbiH in a process that is oxygen-independent.

The exact mechanism of ring contraction (for both pathways) remains unclear, but it is likely methylation occurs first, closely followed by ring contraction. For the aerobic pathway a proposal for the mechanism of the ring contraction process has been outlined (Schroeder *et al.*, 2009). The enzyme CobG contains a $[4\text{Fe-4S}]^{2+/1+}$ cluster and a non-haem iron. The C-terminus containing these redox groups shares similarity to the NiR/SiR. It has been shown to utilise molecular oxygen by binding to the non-haem iron. An electron is provided by the reduced $[4\text{Fe-4S}]^{1+}$ cluster to generate firstly a peroxy intermediate, before a $\text{Fe}^{3+}\text{-O}^+$ species is formed. This oxygen species is used to hydroxylate the C-20 position on the substrate precorrin-3A, with the extra proton presumably gained from solvent (H_2O) exchange. In addition, the acetate side chain at C-2 forms a γ -lactone at C-1, thus producing precorrin-3B. Subsequently CobJ methylates at C-17 and the following ring contraction step results in precorrin-4. The next sections describe the early stages of the anaerobic pathway and discuss how this branch point is shared with sirohaem biosynthesis.

Table 1.2. Comparison of the aerobic and anaerobic pathways, adapted from (Warren *et al.*, 2002).

Aerobic pathway		Anaerobic pathway	
	Uro'gen III		Uro'gen III
CobA	↓	← C2, C7 methylation →	↓
	Precorrin-2		Precorrin-2
	↓	Cobalt insertion →	↓
CobI	↓	← C20 methylation →	Cobalt-precorrin-2
	Precorrin-3A		↓
CobG	↓	← C-20 hydroxylation ← γ -lactone ring formation	Cobalt-precorrin-3
	Precorrin-3B		↓
CobJ	↓	← Ring contraction → δ -lactone ring formation →	↓
	Precorrin-4		Cobalt-precorrin-4
CobM	↓	← C11 methylation →	↓
	Precorrin-5		Cobalt-precorrin-5A
	↓	Acetaldehyde extrusion → ← Acetic acid extrusion	↓
CobF	↓	← C1 methylation →	Cobalt-precorrin-5B
	Precorrin-6A		↓
CobK	↓	← C18-19 reduction →	Cobalt-precorrin-6A
	Precorrin-6B		↓
CobL	↓	← C5, C15 methylation and decarboxylation →	Cobalt-precorrin-6B
	Precorrin-8X		↓
CobH	↓	← C11-C12 methyl rearrangement →	Cobalt-precorrin-8X
	HBA		↓
CobB	↓	← <i>a, c</i> -amidation →	Cobyric acid
	HBA <i>a, c</i> -diamide		↓
CobNST	↓	← Cobalt insertion	↓
	Cob(II)yrinic acid <i>a, c</i> -diamide		Cobyric acid <i>a, c</i> - diamide
CobR	↓	← Cobalt reduction →	↓
	Cob(I)yrinic acid <i>a, c</i> -diamide		Cob(I)yrinic acid <i>a, c</i> - diamide
CobA^{Ado}	↓	← Adenosylation →	↓
	Ado-cob(II)yrinic acid <i>a, c</i> -diamide		Ado-cob(II)yrinic acid <i>a, c</i> -diamide
CobQ	↓	← <i>b, d, e, g</i> -amidation →	↓
	Ado-cobyric acid		Ado-cobyric acid
CobCD	↓	← Aminopropanol attachment →	↓
	Ado-cobinamide		Ado-cobinamide
CobP	↓	← Phosphorylation and GMP addition →	↓
	Ado-GDP- cobinamide		Ado-GDP- cobinamide
CobUV	↓	← Nucleotide loop assembly →	↓
	Ado-cobalamin		Ado-cobalamin

1.3.6 The chelatase branch point: Sirohaem and anaerobic pathways

For the anaerobic biosynthesis of cobalamin and sirohaem the first couple of steps are shared. The first step is rate-limiting and is catalysed by the *S*-adenosyl-L-methionine uroporphyrinogen III methyltransferase (SUMT) enzyme. Using *S*-adenosyl-L-methionine (SAM) as a methyl donor, SUMT methylates at C-2 and C-7 of uroporphyrinogen III to produce precorrin-2. This step is also shared by the aerobic pathway and was the first to be characterised when the SUMT enzyme CobA from *P. denitrificans* SC510 was isolated and purified (Blanche *et al.*, 1989). It was shown that CobA is sensitive to inhibition by its substrate uroporphyrinogen III and that it is competitively inhibited by *S*-adenosyl-L-homocysteine (SAH), which is a by-product of the reaction (Blanche *et al.*, 1989; Robin *et al.*, 1991). Later the same research group isolated and characterised another SUMT (SirA, subsequently shown to belong to a sirohaem operon) enzyme from *B. megaterium* DSM509 and this showed similar enzymatic properties to CobA (Robin *et al.*, 1991). This enzyme is shared by both sirohaem and cobalamin pathways and, in some organisms such as *B. megaterium*, several copies (CysG^A, SirA, NasF) can be found located in the genome. CysG^A is located within a large cobalamin biosynthetic operon, which will be discussed later (Raux *et al.*, 1998b).

NasF is situated within a nitrite reductase assimilation operon (*nasDEFG*). *nasD* and *nasE* encode the large and small components of a nitrite reductase, respectively, with *nasG* predicted to encode a nitrite transporter. Although it is clear that this enzyme system requires sirohaem as a cofactor, a question remains as to why only the SUMT enzyme is associated with the operon. To make sirohaem two more steps are required and it is interesting that a complete sirohaem biosynthetic operon (*sirABC*) is also present in *B. megaterium*. The *sirA* gene, encodes for the enzyme SUMT (SirA) whilst the next enzyme SirC catalyses an NAD⁺ dependent macrocycle ring oxidation at C-14 to C-15 to produce sirohydrochlorin (Leech, 2002). Following this, the ferrochelatase SirB inserts ferrous iron into sirohydrochlorin to produce sirohaem (Leech, 2002; Raux *et al.*, 2003). In other organisms such as *S. typhimurium* and *Lactobacillus reuteri*, instead of three separate genes encoding for sirohaem biosynthesis, a single multifunctional gene named CysG is found. This gene was first characterised in *E. coli* (Warren *et al.*, 1990a; Warren *et al.*, 1994). The C-terminus of CysG encodes the SUMT enzyme

and is termed CysG^A, whilst the N-terminus encodes both dehydrogenase and ferrochelatase activities and is termed CysG^B. For the anaerobic biosynthesis of vitamin B₁₂ cobalt is inserted into sirohydrochlorin (otherwise known as factor II) by the enzyme CbiX^L (*B. megaterium*) to give cobalt-factor II. This is also catalysed in archaea by a smaller version of the enzyme called CbiX^S (Brindley *et al.*, 2003). In other bacteria such as *S. typhimurium* a different enzyme CbiK is required. It should be briefly explained that factor intermediates can occur in both aerobic and anaerobic pathways, and refers to the level of oxidation (saturation) in the ring macrocycle. In brief factor intermediates contain an extra double bond and are thus the oxidised form of the precorrin. At some point in both pathways, factor intermediates are converted back to the level of hexahydroporphyrin (precorrin), although no enzyme has been attributed to catalyse this reaction.

The regulation and control between the sirohaem and the anaerobic pathway is unclear and may depend on the bioavailability of the metals iron and cobalt inside the cell (Raux *et al.*, 2000). For the next steps in the anaerobic pathway limited experimental evidence is available. However, based on the similarity shared between some of the enzymes of the two pathways the various intermediates have been predicted to occur based on the characterised aerobic pathway intermediates, [see (Debussche *et al.*, 1993)].

1.3.7 The oxygen-dependent (aerobic) pathway

The aerobic biosynthesis of adenosylcobyrinic acid from precorrin-2 was elucidated over 20 years ago. For a comprehensive review, see (Warren *et al.*, 2002). This was achieved through a major effort to understand the regulation and biochemistry of the pathway, with the ultimate aim of producing an enhanced commercial B₁₂ producing strain of *P. denitrificans* (Martens, 2002).

The aerobic pathway proceeds through a series of cobalt-free intermediates. Beginning with precorrin-2, the first reaction involves the methylation at position C-20. This is catalysed by CobI, which forms the intermediate precorrin-3A (Debussche *et al.*, 1993). Following this the ring contraction process takes place. In the aerobic pathway this requires a two-stage process and is dependent on oxygen. Initially, the iron-sulphur enzyme CobG (or the flavoprotein CobZ in some other

organisms) hydroxylates the *meso* C-20 position to form precorrin-3B, which also results in the formation of a γ -lactone ring (Schroeder *et al.*, 2009). Then, CobJ methylates the C-17 position with ring contraction believed to occur afterwards, forming the product precorrin-4 (Debussche *et al.*, 1993). After precorrin-4 the next enzyme CobM methylates at the C-11 position to produce the highly unstable precorrin-5 (Debussche *et al.*, 1993). Subsequently the enzyme CobF is responsible for the C-1 methylation and removal of the extruded C-20 position as acetic acid (Battersby *et al.*, 1981; Mombelli *et al.*, 1981), producing precorrin-6A (Roessner *et al.*, 1994; Thibaut *et al.*, 1990a; Thibaut *et al.*, 1990c). It is at this point that a two-electron reduction of the corrin ring is required. This is catalysed by CobK, which uses NADPH as an electron donor. Unusually the sequence of CobK does not reveal any typical motifs for NAD(P)H binding, but the reaction was shown to be highly specific for NADPH instead of NADH producing precorrin-6B from precorrin-6A (Blanche *et al.*, 1992c).

The proceeding reactions require methylations at C-5 and C-15, followed by a decarboxylation of the *e* acetate side chain to produce precorrin-8. However, the order of the C-5 and C-15 methylations remains unclear. Nonetheless, for most aerobic pathway organisms this reaction is catalysed by a single enzyme CobL (Blanche *et al.*, 1992a). This protein is composed of a fusion of two proteins, with the N-terminus and C-terminus encoding two domains most similar to CbiE and CbiT, respectively from the anaerobic pathway. At first it was thought the N-terminal CbiE domain catalyses both methylations, as its sequence was most similar to other class III methyltransferases (Roth *et al.*, 1993). This view was supported initially with the determination of the structure of CbiE (PDB_2BB3) from *A. fulgidus*. However, the structure for CbiT (PDB_1F38) from *Methanothermobacter thermoautotrophicus*, also revealed this to be an unusual class I methyltransferase. This is unique amongst all the tetrapyrrole biosynthetic methyltransferases and established that the CbiT domain of CobL is likely to catalyse one of the methylations and the decarboxylation event.

After this a highly unusual 1,5 sigmatropic rearrangement takes place as the C-11 methyl group is migrated to the C-12 position. This is catalysed by CobH, which converts precorrin-8 to hydrogenobyric acid (HBA) (Thibaut *et al.*, 1992). HBA is

a highly stable intermediate that is amidated at side chains *a* and *c*, using glutamine or ammonia as a nitrogen source. This reaction is catalysed by an ATPase, CobB, converting HBA into HBA *a,c*-diamide. This intermediate serves as the substrate for cobalt insertion, whereby a multienzyme complex of CobN, CobS and CobT inserts cobalt in an ATP-dependent fashion, producing cobyrinic acid *a, c* diamide. This remarkable complex shares a high degree of similarity to the magnesium chelatase, which is an essential step in chlorophyll biosynthesis (Heldt *et al.*, 2005).

Around this point the aerobic and anaerobic pathways are thought to rejoin, although the process and timing of the adenosylation process in the anaerobic pathway still remains unclear. In the aerobic pathway it is known that the flavoprotein CobR reduces the cob(II)alt ion to the cob(I)alt species in cobyrinic acid *a,c*-diamide, which is quickly followed by an ATP dependent adenosylation by the CobA^{ado} adenosyltransferase (Lawrence *et al.*, 2008). This enzyme attaches an adenosyl group to the cobalt ion as an upper axial ligand. This enzyme is not to be confused with the CobA SUMT enzyme. Finally the ATPase CobQ, amidates the *b*, *d*, *e* and *f* carboxylate side chains producing adenosylcobyrinic acid.

1.3.8 The oxygen-independent (anaerobic) pathway

After cobalt insertion into factor II by the cobaltochelatease, a SAM dependent methylation at C-20 is catalysed by CbiL to produce cobalt-factor III. However despite reasonable similarity to CobI, CbiL does not methylate precorrin-2 (Spencer, 1994). CbiL has also been studied from the thermophilic archaea *M. thermoautotrophicus*, and has been shown to prefer the oxidised substrate cobalt-factor II to cobalt-precorrin-2, suggesting the factor intermediates are preferred during the early stages of the anaerobic pathway (Frank *et al.*, 2007). It should also be considered that the cobalt ion can exist in three different oxidation states, Co⁺, Co²⁺ and Co³⁺. The Co⁺ species is rare and only Co²⁺ and Co³⁺ species have been observed as cobalt(II)-factor II/III and cobalt(III)-factor II/III intermediates, and confirmed by EPR (Frank *et al.*, 2007). CbiL shows no preference for either of the two cobalt oxidation states (Frank *et al.*, 2007).

The next step of the pathway is the ring contraction event, catalysed by CbiH. This reaction utilises either cobalt-precorrin-3 or cobalt-factor III and involves a SAM-

dependent C-17 methylation, ring contraction and formation of a δ -lactone ring to produce cobalt-precorrin-4 or cobalt-factor IV, respectively. Three publications have described the isolation of the products and brief reaction conditions. Firstly cobalt-factor IV was isolated from a crude cell extract of *Propionibacterium shermanii* and characterised by NMR (Scott *et al.*, 1996). The enzyme that produced this intermediate was identified by first synthesising precorrin-3 and then chemically inserting cobalt. This produced cobalt-precorrin-3, which was incubated with SAM and an *E. coli* crude cell-lysate overexpressing CbiH from *S. enterica*. This yielded a green pigment that was shown to contain cobalt-factor IV (Santander *et al.*, 1997). This new intermediate could then be shown to be converted in low yields (< 1%) to cobyrinic acid, using a crude-cell lysate of *P. shermanii*, an organism that contains a complete anaerobic pathway.

However, recent work has shown an alternative route which likely reflects the true pathway. Instead of first synthesising precorrin-3 and chemically inserting the cobalt, cobalt-precorrin-3 or cobalt-factor III was prepared using a multi-enzyme approach including the cobaltochelatase CbiX. By using an *E. coli* crude-cell lysate overexpressing either the CbiH from *M. thermoautotrophicus* or *S. enterica*, it was shown that cobalt-precorrin-3 or cobalt-factor III could be converted into cobalt-precorrin-4 and cobalt-factor IV, respectively. Although the yield was still low (< 5%), small quantities could be separated by HPLC and characterised by UV-Vis and LC-MS. Unfortunately, little is known about the mechanism of the reaction and the oxidation state of the cobalt ion. As molecular oxygen is absent in the reaction, redox chemistry with the cobalt ion may be involved in this alternative mechanism (Warren *et al.*, 2002).

The next confirmed reaction of the anaerobic pathway is the C-11 methylation of cobalt-precorrin-4 to produce cobalt-precorrin-5A. This was investigated by using an *E. coli* lysate with the *S. enterica* CbiF overproduced (Kajiwara *et al.*, 2006). Unfortunately the intermediate was extracted as an octamethylester, whereby the surrounding carboxylic acid side chains were esterified in methanol and concentrated sulphuric acid (95:5 v/v). The advantage of using esters is they are more stable. The problem, however, is this method does not provide any knowledge regarding the oxidation state of the cobalt ion or ring macrocycle. In addition, it is

not known how and when the pathway re-enters the precorrin state. Although UV-Vis or mass spectra have been recorded, these have not been fully published, which is presumably due to the very low yields. In the same publication it was also suggested that the *S. enterica* CbiG could catalyse the removal of the δ -lactone ring to produce cobalt-precorrin-5B (Kajiwara *et al.*, 2006). Previous research had identified the δ -lactone ring is released as acetaldehyde (Wang *et al.*, 1996). CbiG is unique to the anaerobic pathway, with only weak similarity (in a short C-terminal sequence) to the hypothetical CobE protein from the aerobic pathway (Vevodova *et al.*, 2005). Although CobE is essential to cobyrinic acid biosynthesis, no role has yet been assigned to the protein. The intermediate cobalt-precorrin-5B was also isolated as an octamethylester form with only NMR data published, although the supplementary information suggests the intermediate is spectrally different from cobalt-precorrin-5A.

Along with CbiG, another unique protein CbiD exists in the anaerobic pathway. This is believed, although not proven, to catalyse the C-1 methylation of cobalt-precorrin-5B to produce cobalt-precorrin-6A. A crystal structure of the *Archaeoglobus fulgidus* CbiD was recently released, although not published (PDB_1SR8). The sequence and structure of CbiD bears no similarity to CobF or indeed any other methyltransferase. Moreover the enzyme shows no similarity to any other known enzyme, in sequence or structure. Normally methyltransferases typically possess a SAM binding domain which is a conserved GXGXG motif found at the N-terminus. As mentioned CbiD shares no similarity with any of the methyltransferases, but it does have a GXGXG motif in the centre of its sequence. Recent evidence suggests CbiD is the C-1 methyltransferase. This was predicted indirectly by using strains of *E. coli* expressing all the anaerobic enzymes from *S. enterica*, in an attempt to produce adenosylcobyrinic acid. This work showed that the strains could accumulate a mixture of cobyrinic acid *a,c*-diamide (25%) and 1-desmethylcobyrinic acid *a,c*-diamide (75%), but were strangely missing the final amidations that were expected to be catalysed by the cobyrinic acid synthase (CbiP). A deletion of *cbiD* resulted in production of only 1-desmethylcobester derivatives, showing a distinct lack of C-1 methylation (Roessner *et al.*, 2005). Furthermore only 1-desmethylcobester derivatives could be isolated if either the genes *cbiA* or *cbiP* were absent. Although the evidence is not direct, it raises other interesting

questions about the pathway. This is based on the assumption that the amidases CbiA and CbiP are involved at a much later step in the pathway.

For the rest of the pathway the intermediates and reactions have yet to be isolated. However, the enzymes CbiJ, CbiE, CbiT and CbiC share reasonable homology to enzymes from the aerobic pathway that have been characterised and shown to convert precorrin-6 to HBA. Of particular note in *B. megaterium* DSM509 is that it has a very similar protein to CobL which is named CbiET, whereas in many other organisms such as *S. enterica*, *A. fulgidus*, *P. shermanii* and *M. thermoautotrophicus* these enzymes are found as two separate enzymes (CbiE and CbiT). It is expected that these steps should occur in a similar order to the aerobic pathway. A summary of the anaerobic biosynthesis of vitamin B₁₂ is provided in Figure 1.4.

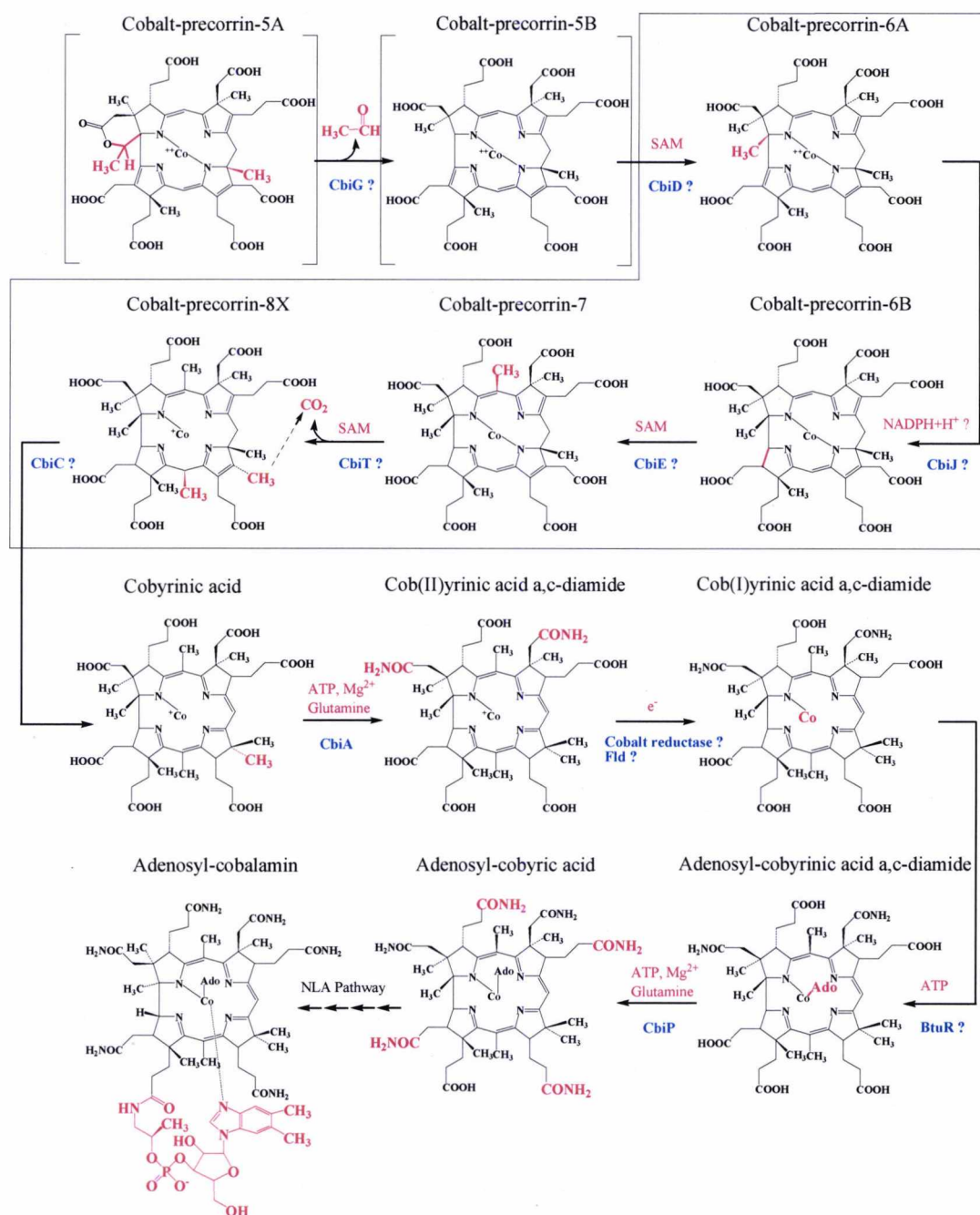


Figure 1.4. The anaerobic biosynthesis of adenosylcobalamin from uroporphyrinogen III. An outline of the anaerobic pathway is illustrated. Intermediates and reactions in open brackets have been isolated, but require further experimental validation. Reactions and intermediates in the closed box have yet to be characterised. These steps are predicted based on an assumed similarity with the aerobic pathway. The enzyme or exact timing that the ring macrocycle is reduced back to the precorrin level is not understood. The cobalt reductase is yet to be identified, but is required for the adenosyl group (Ado) to be attached as the upper ligand. The later steps in the NLA pathway leading to the formation of adenosylcobalamin are simplified as a single step.

1.4 *Bacillus megaterium*

1.4.1 Background

The “Big beast” (Vary, 1994) *B. megaterium* belongs to the *Firmicutes* phylum which is mainly composed of Gram-positive bacteria, and includes species such as *Bacillus*, *Clostridium* and *Staphylococcus*. Other members of the family include groups such as *Megasphaera*, *Pectinatus*, *Selenomonas*, and *Zymophilus*, which appear Gram-negative due to the presence of a pseudo porous outer membrane. These organisms all share a low G+C (%) content.

B. megaterium is a large mesophilic Gram-positive rod and has previously been used for the commercial production of vitamin B₁₂ (Raux *et al.*, 1998b). It utilises the anaerobic pathway, but remarkably produces vitamin B₁₂ aerobically or anaerobically. A similar trait is observed with *Rhodobacter capsulatas*, which despite possessing an aerobic pathway, can produce vitamin B₁₂ both under aerobic or anaerobic conditions (Raux, 1999). *B. megaterium* is highly versatile as its metabolism allows it to utilise inexpensive carbon sources such as corn syrup, meat and petrochemical waste. It is capable of surviving in a diverse range of environments, including nutrient poor niches such as the desert (Vary *et al.*, 2007). It has a relatively low G+C (%) content of 38.2% (Eppinger *et al.*, 2011), encodes polycistronic mRNA and has a profound ability to sporulate, which is stimulated by the presence of manganese (Kunkel, 2004). Furthermore, many strains stably maintain multiple plasmids, such as *B. megaterium* QM B1551, which harbours seven plasmids of sizes 5.4, 9.1, 26.3, 53.9, 71, 108, and 165 kb [designated pBM100-700 respectively (Rosso & Vary, 2005)]. These plasmids comprise approximately 11% of the total cellular DNA. Most importantly, under standard laboratory conditions, all seven of these plasmids are stably maintained with no known positive selection (Rosso & Vary, 2005). This is one of the key advantages of *B. megaterium* and helps to distinguish it over its better characterised *Bacillus subtilis* cousin. The two smallest plasmids (pBM100 and pBM200) replicate by a rolling-circle mechanism, whilst the other larger plasmids replicate using cross-hybridisation replicons (theta replication). The presence of these plasmids provides a unique asset to the cells capability to survive in diverse environments. Sequencing

of these plasmids revealed many genes encoding a range of transporters (including heavy metal exporters), dehydrogenases, amylases, a cytochrome P-450, monooxygenases, carboxylases, transposases and integrases. Furthermore, there are genes encoding proteins for cell division, germination, rifampin resistance and rather surprisingly, an entire rRNA operon (located on pBM400) (Kunnimalaiyaan & Vary, 2005). Another strain, *B. megaterium* DSM319, is plasmidless and has been used as the model strain for studying recombinant protein production. This has led to further advances in recombinant gene expression (T7 and *xylA* promoters), which has widened the use of *B. megaterium* in modern biotechnology (Biedendieck *et al.*, 2007b; Gamer *et al.*, 2009; Vary *et al.*, 2007).

1.4.2 Recombinant protein production

A wide variety of organisms have been developed for recombinant protein production, each with their own advantages and disadvantages. For a comprehensive review please see (Demain & Vaishnav, 2009). *B. megaterium* offers the advantages of high productivity from inexpensive materials, is non-pathogenic, has low protease activity and is able to maintain plasmids without positive selection (Vary *et al.*, 2007). Its main disadvantage, like other prokaryotic systems, is the inability to make post-translational modifications, which is an asset unique to eukaryotic systems. *E. coli* is regarded as a well-established and researched host for recombinant protein production in prokaryotes. However, previous research has shown that a number of *B. megaterium* cobalamin biosynthetic enzymes are produced as inclusion bodies in *E. coli* (Evelyne Raux, unpublished data). In order to characterise the anaerobic pathway, large quantities of soluble enzymes are required. One of the initial aims of this thesis was to see if this solubility could be improved by homologous production of these enzymes in the host *B. megaterium*. This next section describes recent developments in recombinant protein production in *B. megaterium* DSM319.

1.4.3 The *xyl* operon

B. megaterium has been developed as an excellent host for recombinant production of homologous and heterologous proteins. This area of research began with the isolation and characterisation of a xylose utilisation operon in *B. subtilis* (Gartner *et al.*, 1988), which was later characterised in a number of related *Bacillus* strains

including *B. megaterium* (Rygus *et al.*, 1991). Many of the genes in prokaryotes that are involved together in sections of metabolism are clustered together in operons. This is very much true with xylose utilisation in *B. megaterium*. Xylose utilisation requires the XylA (xylose isomerase), XylB (xylulokinase) and a permease (XylT). The genes encoding these proteins are organised into an operon (*xylABT*) that is transcribed from the xylose-inducible promoter (P_{xylA}). The genes are regulated at the level of transcription by the gene product of *xylR*, which is divergently transcribed and shares the same intergenic region as *xylABT*. In the absence of xylose, XylR can bind to two operator sequences (O_L/O_R) that overlap by 4 bp and is released upon the binding of xylose to XylR. This allows for transcription to start and acts as a sensing mechanism for the cell to switch its metabolism and utilise xylose as a carbon source when it is available. In addition, Gram-positive low G+C (%) bacteria control central metabolism by down-regulating alternative carbon utilisation genes/operons in response to glucose. Transcription of the *xylABT* operon is also controlled by a *cis*-acting catabolite response element (*cre*), located between bp 23 to 200 of the *xylA* open reading frame and a *trans*-active catabolite controlled protein (CcpA), which act as a dual mode of regulation (Rygus & Hillen, 1992). CcpA activity is regulated by the phosphoenolpyruvate: glucose phosphotransferase system (PTS). In the presence of glucose, the histidine containing phosphocarrier protein (HPr) becomes phosphorylated at His¹⁵ and Ser⁴⁶ by phosphoenolpyruvate (PEP) and ATP, respectively (Monedero *et al.*, 2001). This phosphorylation event enhances the binding of CcpA to the *cre* sequence.

The initial discovery of the xylose inducible promoter system was subsequently exploited for use in recombinant gene expression. A shuttle vector named pWH1520 was constructed containing *xylR*, the P_{xylA} promoter, the first 195 bp of *xylA*, followed by a multiple cloning site (MCS) (Rygus & Hillen, 1991). This plasmid was derived from pBC16 (Bernhard *et al.*, 1978), which allows for cloning in *E. coli* due to the presence of the origin of replication *colEI* and the β -lactamase gene for positive selection. Furthermore for replication and selection in *B. megaterium*, the rolling-circle origin of replication *oriU*, *repU* and the gene for tetracycline resistance are present. Using this promoter for recombinant gene expression, expression can be induced by 350 fold. Initially the expression vector

was successfully used for production of a number of recombinant proteins from prokaryotes and eukaryotes, including *E. coli* β -galactosidase, *B. megaterium* glucose dehydrogenase and human urokinase-like plasminogen activator (Rygus & Hillen, 1991). Interestingly, the *Clostridium difficile* toxin A with a molecular mass of 308 kDa, was also successfully produced at higher levels when compared to recombinant production in *E. coli*.

1.4.4 Plasmids for intracellular protein production

To improve this expression plasmid further, the *cre* sequence in *xyIA* was removed and the original MCS was replaced with an improved version to create the plasmid pMM1520 (PhD Thesis: Malten, 2005). To allow for easy purification of recombinant proteins, affinity tags such as the hexahistidine-tag (His₆) and the streptavidin(II)-tag (StrepII) were positioned in the MCS, allowing the gene to be cloned in frame with the tag. The overproduced fusion protein can then be easily purified in a one-step procedure (Biedendieck *et al.*, 2007b). Furthermore, several varieties of this new improved vector were designed so the tags could be added to either the N-terminus (pN-His-TEV1622 and pN-Strep-TEV1622) or C-terminus (pC-His1622 and pC-Strep1622) of the protein, with the added option of a tobacco etch virus protease cleavage site (TEV) for later removal of the tag [see Figure 1.5, adapted from (Biedendieck *et al.*, 2007b)].

As a model for studying this new system, the production of recombinant green fluorescent protein (GFP) from *Aequorea victoria* was studied in *B. megaterium* DSM319. This strain does not contain any indigenous plasmids and is a good producer of recombinant proteins using the xylose inducible promoter system. All vectors were tested and showed that the strains could produce GFP up to 17.9 mg L⁻¹, with up to 9 mg L⁻¹ isolated in a one-column affinity purification (Biedendieck *et al.*, 2007b). By modifying the 5' untranslated region and optimising the ribosome binding site (RBS), this plasmid system can produce GFP up to 124 mg L⁻¹ and 1.25 g L⁻¹ in shaking flask culture and high cell density cultivation, respectively (Stammen *et al.*, 2010). Alternative vectors have also been designed with the xylose inducible promoter system and incorporated with different origins of replication for compatibility. In this way pMGBm19 was created, a plasmid that contains the rolling circle origin of replication *repM100* derived from the pBM100 indigenous

plasmid of *B. megaterium* QM B1551. This also has the chloramphenicol resistance gene *cat* (chloramphenicol acetyltransferase) so is compatible as a second plasmid when used as a dual system (Gamer *et al.*, 2009).

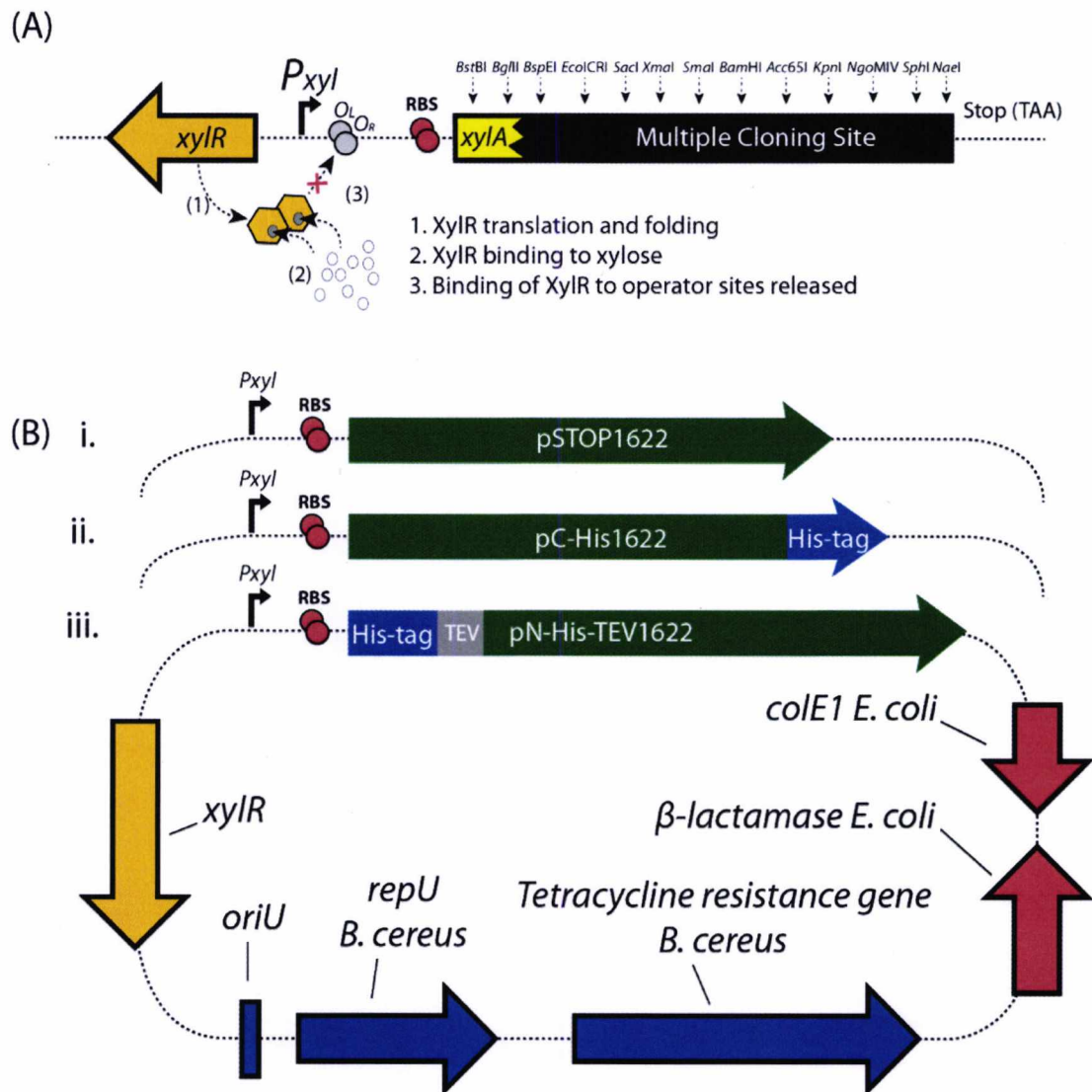


Figure 1.5. The xylose inducible promoter system and shuttle vectors for intracellular protein production in *B. megaterium*. **A.** The xylose inducible promoter system and multiple cloning site (MCS) are highlighted. The *xylR* gene (orange arrow) encodes the xylose repressor protein that is transcribed and translated (1) before binding to xylose (2). This causes a conformation change and releases the protein from binding to the operator sequence, which is located downstream of the xylose inducible promoter (3). This activates gene expression. The first 195 bp of *xylA* (broken yellow arrow) is fused to the MCS. Restriction sites are shown with arrows, with a stop codon (TAA) present downstream. **B.** The xylose inducible promoter system was adapted for a variety of expression plasmids, with extracellular protein production vectors shown. i. pSTOP1622, for cloning with no tags; ii. pC-His1622, for cloning with a C-terminal His₆-tag; iii. pN-His-TEV1622, for cloning with a N-terminal His₆-tag and TEV protease cleavage site. The origins for replication and antibiotic resistance genes are shown for *E. coli* (red arrows) and *B. megaterium* (blue arrows).

1.4.5 Location of vitamin B₁₂ genes in *B. megaterium* DSM319

With the recent release of the *B. megaterium* DSM319 genome (Eppinger *et al.*, 2011), this has allowed for detailed analysis of vitamin B₁₂ biosynthesis. The complete *de novo* biosynthesis of vitamin B₁₂ requires around thirty genes that are mainly organised into large operons. Furthermore, other genes encoding coenzyme B₁₂ dependent enzymes have been located within the genome, including the vitamin B₁₂ dependent methionine synthase MetH. Others include the ribonucleotide reductase NrdJ, ethanolamine ammonia lyase EutBC and the methylmalonyl CoA mutase MutAB. For the biosynthesis of vitamin B₁₂, the genes are localised within four operons with a few single genes located separately. This arrangement is integral to the regulation of the pathway, which shall be later discussed in detail (see Chapter 3).

In Gram-positive bacteria such as the *Bacillus*, *Clostridium*, *Mycobacterium* and *Staphylococcus* species, the genes encoding for the synthesis of uroporphyrinogen III from glutamyl-tRNA are organised into a single operon consisting of *hemAXCDBL*, with a few subtle differences in particular species (Johansson & Hederstedt, 1999). This gene arrangement is less conserved in other species of eubacteria and in particular in the archaea where genes are more generally scattered throughout the whole genome. The majority of the anaerobic pathway is encoded in the *cobI* operon containing the genes *cbiW-cbiH₆₀-cbiX^L-cbiJ-cbiC-cbiD-cbiET-cbiL-cbiF-cbiG-cbiA-cysG^A-cbiY-btuR*. This operon was first isolated from the industrial vitamin B₁₂ producer strain *B. megaterium* DSM509 and was shown to promote vitamin B₁₂ biosynthesis in *E. coli*, but only under anaerobic conditions (Raux *et al.*, 1998a). In this operon, the genes *cbiW* and *cbiY* remain uncharacterised. The former, *cbiW*, is predicted to encode a 2Fe-2S thioredoxin with no known functional requirement in cobalamin biosynthesis (Raux *et al.*, 1998a). However, it is positioned next to *cbiH₆₀*, which is unusual for a ring contraction enzyme, as it encodes an extra C-terminal extension that shares similarity with a range of different NiR/SiR and may contain an iron-sulphur cofactor (Raux *et al.*, 1998a). More detail on these enzymes shall be discussed in Chapter 4 and 5. CbiY represents a small flavoprotein that has been previously been overproduced and purified from *E. coli* (Andrew Lawrence, personal communication). No function has been described for this protein, but it appears not

to be essential for adenosylcobyrinic acid biosynthesis (Raux *et al.*, 1998a). BtuR shows similarity to the large family of ATP:cob(I)yrinic acid *a,c*-diamide adenosyltransferases, although this has not been experimentally confirmed. This operon is highly conserved in many of the vitamin B₁₂ producing *Bacillus* species, with some *Bacilli* containing extra genes such as cobalt transporters (*cbiM*, *cbiN*, *cbiQ* and *cbiO*), the precorrin-2 dehydrogenase (*sirC*) and the cobyrinic acid synthase (*cbiP*). However in *B. megaterium* DSM319, these are located separately within the genome.

The enzymes for the later stages of vitamin B₁₂ biosynthesis are encoded within a smaller *cobII* operon. This operon encodes the genes *cbiZ-cbiB-cobD-cobU-cobS-cobC-btuF-cobA^{Ado}*. The majority of these genes are involved in the NLA pathway and have been characterised in detail (Escalante-Semerena *et al.*, 1990; Escalante-Semerena *et al.*, 1992; Escalante-Semerena, 2007; Maggio-Hall & Escalante-Semerena, 1999; Woodson *et al.*, 2003), although *btuF* is a vitamin B₁₂ transporter and *cobA^{Ado}* is another ATP:cob(I)yrinic acid *a,c*-diamide adenosyltransferase. Two other genes that are closely associated with the *cobII* operon are *tnrA* (a putative transcriptional regulator) and a gene encoding a putative metal binding protein, the latter displaying some similarity to CobW from the aerobic pathway. Although the function of CobW is unclear, it is a GTPase that may be involved in the cobalt chelation process (Heldt *et al.*, 2005) and is essential to cobalamin biosynthesis in *P. denitrificans* (Crouzet *et al.*, 1991). An additional copy of this gene is also found on the *B. megaterium* DSM319 genome, located on its own and containing a histidine rich-motif. Other genes involved in the biosynthesis of the lower axial ligand DMB are also required, but these have yet to be identified and annotated in the *B. megaterium* DSM319 genome. All the genes that are required or associated with adenosylcobalamin biosynthesis are summarised in Figure 1.6.

Vitamin B₁₂ biosynthetic genes in *B. megaterium* DSM319

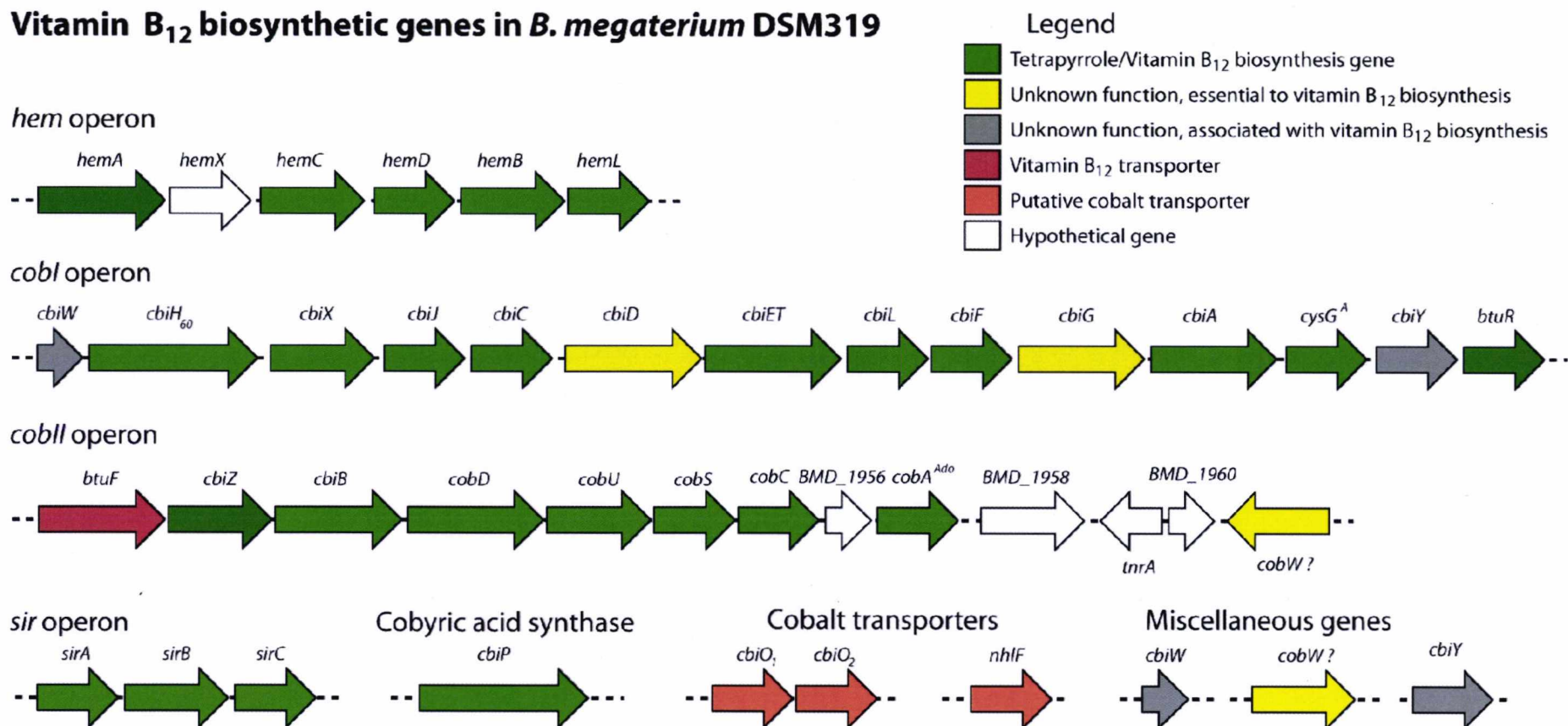


Figure 1.6. Organisation of genes involved in the biosynthesis of adenosylcobalamin in *B. megaterium* DSM319. All genes are represented as coloured arrows as indicated in the figure legend. All genes clustered in operons or situated on their own are annotated. Hypothetical genes are annotated as open reading frames (BMD_000) as shown on the *B. megaterium* DSM319 genome viewer (<http://megabac.tu-bs.de>).

1.5 Aims of this thesis

The aerobic biosynthesis of the corrin macrocycle has been solved by a range of techniques that focused on an industrial B₁₂ producing strain of *P. denitrificans*, SC510 (Debussche *et al.*, 1993). However, the characterisation of the anaerobic pathway has been harder to achieve. Although some information has been gained with *P. shermanii*, recent investigations have studied enzymes from *S. enterica* and *M. thermoautotrophicus* expressed in *E. coli*. In this thesis a more challenging and dynamic approach will be attempted. *B. megaterium* is ideally suited in modern biotechnology, with the recent advances in recombinant protein production. This shares similar advantages that made the elucidation of the aerobic biosynthesis of the corrin ring in *P. denitrificans* possible. Already the cobaltochelatease CbiX^L from *B. megaterium* has been characterised and shown to be more active than any other anaerobic cobaltochelatease (PhD Thesis: Leech, 2004). It is hoped that similar observations may be seen with further enzymes. This may also provide further insights into why *B. megaterium* is capable of vitamin B₁₂ production commercially, whilst *S. enterica* and recombinant strains of *E. coli* can only produce comparatively small quantities. This makes *B. megaterium* highly suited to elucidating the anaerobic biosynthesis of the corrin macrocycle, which has remained one of the most challenging problems in chemical biology. Along with investigating the anaerobic pathway, other experiments shall focus on the general microbiology and regulation of vitamin B₁₂ biosynthesis in *B. megaterium* DSM319, whilst other studies shall investigate the limits of genetic manipulation to improve vitamin B₁₂ production within the organism.

The major aims of this thesis are firstly to recombinantly produce and purify many of the cobalamin biosynthetic enzymes from *B. megaterium* DSM509, using homologous production in *B. megaterium* DSM319. Secondly the enzyme CbiH₆₀ will be investigated and an attempt to elucidate the mechanism of ring contraction problem will be sought. Then, if this reaction can be solved, other enzymes from the anaerobic pathway will be investigated with the hope of obtaining some of the long sought after intermediates in the pathway. In addition, a brief study of the regulation of cobalamin biosynthesis shall be presented, along with attempts to engineer new *B. megaterium* strains enhanced in cobalamin overproduction.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Chemicals and Kits

Most chemicals, media, antibiotics were purchased from Sigma-Aldrich, Fluka, and Fisher. Disposable and empty PD-10 columns from Amersham Biosciences; DNeasy® Tissue Kit, MiniElute™ Gel Extraction kit, QIAprep® Spin Miniprep Kit, QIAquick® Gel Extraction Kit, RNeasy® Mini Kit and RNAprotect® Bacteria Reagent from QIAGEN; iScript™ One-Step RT-PCR Kit with SYBR Green, Bradford reagent and 96-well plates from Bio-Rad; restriction enzymes and T₄ DNA ligase from Promega; *Bsr*GI, *Pac*I, *Eco*NI, *Eag*I and *Acc*65I were purchased separately from New England Biolabs; shrimp alkaline phosphatase from Fermentas; Expand High Fidelity^{PLUS} PCR System from Roche Diagnostics GmbH; Immobiline™ DryStrip gels for IEF from GE Healthcare.

2.2 Bacterial strains and plasmids

2.2.1 Bacterial strains

Bacterial strains (Table 2.1) were purchased from Novagen, Promega, Gibco Life Technologies, DSMZ (Braunschweig) or kindly provided by Dr Rebekka Biedendieck (Technical University of Braunschweig) and Dr Evelyne Deery.

Table 2.1: Bacterial Strains

Name	Description	Reference/Source
<i>Bacillus megaterium</i>		
DSM319	Wild-type, vitamin B ₁₂ producer, genome sequenced	DSMZ, Braunschweig
DSM509	Wild-type, vitamin B ₁₂ producer	DSMZ, Braunschweig
<i>Escherichia coli</i>		
DH10B	<i>F</i> ⁻ <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) φ80 <i>dlacZ</i> Δ <i>M15</i> Δ <i>lacX74</i> <i>deoR</i> <i>recA1</i> <i>endA1</i> <i>araD139</i> Δ(<i>ara, leu</i>)7697 <i>galU</i> <i>galK</i> λ ⁻ <i>rpsL</i> <i>nupG</i>	Gibco Life Technologies
JM109	<i>endA1</i> <i>glnV44</i> <i>thi-1</i> <i>relA1</i> <i>gyrA96</i> <i>recA1</i> <i>mcrB</i> ⁺ Δ(<i>lac-proAB</i>) <i>e14-</i> [<i>F'</i> <i>traD36</i> <i>proAB</i> ⁺ <i>lacI</i> ^q <i>lacZ</i> Δ <i>M15</i>] <i>hsdR17</i> (<i>r_K⁻m_K⁺</i>)	Promega
BL21*(DE3)	<i>F</i> ⁻ <i>ompT</i> <i>hsdS_B</i> (<i>r_B⁻, m_B⁻</i>) <i>gal dcm</i> (DE3)	Novagen
<i>Salmonella typhimurium</i>		
AR2680	<i>Leu</i> ⁺ <i>cbiB</i> <i>metE</i> , <i>Sm</i> ^R (Cobinamide detection strain)	(Raux <i>et al.</i> , 1996)
AR3612	<i>Leu</i> ⁺ <i>cysG</i> <i>metE</i> , <i>Sm</i> ^R (Cobyric acid detection strain)	(Raux <i>et al.</i> , 1996)

2.2.2 Plasmids

Plasmids (Table 2.2) were kindly provided by Dr Rebekka Biedendieck, Dr Steffi Frank, Dr Evelyne Deery and Dr Martin Gamer, or prepared in this work.

Table 2.2: Plasmids

Plasmid	Description	Reference/Source
pWH1520	Shuttle vector for cloning in <i>E. coli</i> (<i>colEI</i> , Amp ^R) and gene expression under xylose control in <i>B. megaterium</i> (<i>repU</i> , Tet ^R); P _{<i>xyIA</i>} -MCS-Stop	(Rygus <i>et al.</i> , 1991)
pStop1622	pWH1520 derivative-vector for cloning in <i>E. coli</i> (<i>colEI</i> , Amp ^R) and gene expression under xylose control in <i>B. megaterium</i> (<i>repU</i> , Tet ^R); P _{<i>xyIA</i>} -MCS-Stop	(Biedendieck <i>et al.</i> , 2007b)
pC-His1622	pWH1520 derivative-vector for intracellular production of C-terminal His ₆ -tagged proteins in <i>B. megaterium</i> P _{<i>xyIA</i>} -MCS-His ₆ -tag-Stop	(Biedendieck <i>et al.</i> , 2007a)
pN-His-TEV1622	pWH1520 derivative-vector for intracellular production of N-terminal His ₆ -tagged proteins in <i>B. megaterium</i> P _{<i>xyIA</i>} -His ₆ -tag-MCS-Stop	(Biedendieck <i>et al.</i> , 2007a)
p3Stop1623hp	pStop1622 derivative vector, with optimised promoter elements and RBS. <i>PacI</i> restriction site inserted between -35 and -10 region of the P _{<i>xyIA</i>} promoter. One stop codon per reading frame introduced after MCS. Increased protein production (10-fold)	(Stammen <i>et al.</i> , 2010)
pHBIntE	Shuttle vector for cloning in <i>E. coli</i> (<i>colEI</i> , Amp ^R) and for integration in <i>Bacillus spp</i> using a temperature sensitive origin of replication (pE194) and Ery ^R . Contains a MCS with xylose inducible promoter; P _{<i>xyIA</i>} -MCS-Stop	(Biedendieck <i>et al.</i> , 2010)
pMGBm19	Shuttle vector for cloning in <i>E. coli</i> (<i>colEI</i> , Amp ^R) and gene expression under xylose control in <i>B. megaterium</i> (CIm ^R); P _{<i>xyIA</i>} -MCS-Stop	(Gamer <i>et al.</i> , 2009)
pAR8766	16.3 kb <i>Sau3AI</i> fragment (<i>cobI</i> operon) from <i>B. megaterium</i> DSM509 cloned into <i>BamHI</i> site of pKK223.3	(Raux <i>et al.</i> , 1998b)
pETcoco-2-ABCDC	<i>M. barkeri cobA</i> , <i>M. thermoautotrophicus hemB-sirC</i> , <i>B. megaterium hemC-D</i> genes sub-cloned from pET14b into	(Frank <i>et al.</i> , 2007)

	pETcoco-2. All genes encode a N-terminal His ₆ -tag	
pETcoco-2- <i>ABCDCXL</i>	Derivative of pETcoco-2-ABCDC, with the additional His ₆ -tagged genes <i>cbiX</i> and <i>cbiL</i> from <i>M. thermoautotrophicus</i> . Amp ^R gene replaced with a kanamycin resistance cassette.	Dr S. Frank
pWH1520- <i>cbiX</i>	<i>B. megaterium</i> derived <i>cbiX</i> PCR fragment cloned into <i>SpeI</i> and <i>BamHI</i> of pWH1520 creating P _{<i>xylA</i>} - <i>cbiX</i> -Stop	(Leech <i>et al.</i> , 2003)
pET14b- <i>cobA</i>	<i>M. barkeri</i> derived <i>cobA</i> PCR fragment cloned into <i>NdeI</i> and <i>BamHI</i> of pET14b creating P ₇₇ -His ₆ -tag- <i>cobA</i> -Stop	Dr E. Deery
pSJM007	<i>B. megaterium</i> derived <i>cbiL</i> PCR fragment cloned into <i>BglII</i> and <i>SphI</i> of pC-His1622 creating P _{<i>xylA</i>} - <i>cbiL</i> -His ₆ -tag-Stop	This work
pSJM019	<i>B. megaterium</i> derived <i>cbiL</i> PCR fragment cloned into <i>BglII</i> and <i>SphI</i> of pN-His-TEV-1622 creating P _{<i>xylA</i>} -His ₆ -tag- <i>cbiL</i> -Stop	This work
pSJM188	<i>B. megaterium</i> derived <i>cbiH₆₀</i> PCR fragment cloned into <i>BglII</i> and <i>EagI</i> of pC-His1622 creating P _{<i>xylA</i>} - <i>cbiH₆₀</i> -His ₆ -tag-Stop	This work
pSJM013	<i>B. megaterium</i> derived <i>cbiH₆₀</i> PCR fragment cloned into <i>BglII</i> and <i>EagI</i> pN-His-TEV-1622 creating P _{<i>xylA</i>} -His ₆ -tag- <i>cbiH₆₀</i> -Stop	This work
pSJM037	<i>M. barkeri</i> derived <i>cobA</i> sub-cloned from pET14b into pWH1520 creating P _{<i>xylA</i>} -His ₆ -tag- <i>cobA</i> -Stop	This work
pSJM059	<i>B. megaterium</i> derived <i>cbiH</i> PCR fragment (Truncated, C-terminus extension excluded), cloned into <i>BglII</i> and <i>EagI</i> of pC-His1622 creating P _{<i>xylA</i>} - <i>cbiH</i> -His ₆ -tag-Stop	This work
pSJM064	<i>B. megaterium</i> derived <i>cbiX</i> PCR fragment cloned into <i>SpeI</i> and <i>SphI</i> of pSJM037 by link-and-lock method (Helen Leech Thesis), creating P _{<i>xylA</i>} -His ₆ -tag- <i>cobA</i> - <i>cbiX</i>	This work
pSJM073	<i>B. megaterium</i> derived <i>cbiG</i> PCR fragment, cloned into <i>BglII</i> and <i>SphI</i> of pC-His1622 creating P _{<i>xylA</i>} - <i>cbiG</i> -His ₆ -tag-Stop	This work
pSJM078	<i>B. megaterium</i> derived <i>cbiJ</i> PCR fragment, cloned into <i>BglII</i> and <i>SphI</i> of pC-His1622 creating P _{<i>xylA</i>} - <i>cbiJ</i> -His ₆ -tag-Stop	This work

pSJM105	<i>cbiX-J-C-D-E-T-L</i> fragment sub-cloned from pAR8766 with <i>PacI</i> and <i>BglIII</i> into pWH1520- <i>cbiX</i> , creating pWH1520- <i>cbiX-J-C-D-E-T-L</i>	This work
pSJM108	<i>cbiH₆₀</i> sub-cloned from pSJM188 with <i>BglIII</i> and <i>EagI</i> into pHBIntE, creating pHBIntE- <i>cbiH₆₀</i>	This work
pSJM115	<i>B. megaterium</i> derived <i>cbiW</i> PCR fragment, cloned into <i>BglIII</i> and <i>SphI</i> of pN-His-TEV1622 creating <i>P_{xyIA}</i> -His ₆ -tag-TEV- <i>cbiW</i> -Stop	This work
pSJM116	<i>B. megaterium</i> derived <i>cbiF</i> PCR fragment, cloned into <i>BsrGI</i> and <i>BamHI</i> of pC-His1622 creating <i>P_{xyIA}</i> - <i>cbiF</i> -His ₆ -tag-Stop	This work
pSJM117	<i>cbiH₆₀-X-J-C-D-E-T-L-F-G-A-cysG^A-cbiY-btuR</i> fragment sub-cloned from pAR8766 with <i>NcoI</i> and <i>EagI</i> into pSJM108, creating pHBIntE- <i>cbiH₆₀-X-J-C-D-E-T-L-F-G-A-cysG^A-cbiY-btuR</i>	This work
pSJM129	<i>cbiH₆₀-X-J-C-D-E-T-L-F-G-A-cysG^A-cbiY-btuR</i> fragment sub-cloned from pSJM117 with <i>SacI</i> and <i>XbaI</i> into pStop1622 with <i>SacI</i> and <i>SpeI</i> , creating pStop1622- <i>cbiH₆₀-X-J-C-D-E-T-L-F-G-A-cysG^A-cbiY-btuR</i> . No <i>P_{xyIA}</i> promoter, expression constitutive	This work
pSJM152	<i>B. megaterium</i> derived <i>cbiW</i> PCR fragment, cloned into <i>BglIII</i> and <i>SphI</i> of pC-His1622 creating <i>P_{xyIA}</i> - <i>cbiW</i> -His ₆ -tag-Stop	This work
pSJM161	<i>P_{xyIA}-cbiH₃₀</i> sub-fragment cloned from pSJM059 into <i>EcoNI</i> and <i>EagI</i> of pSJM064, creating pWH1520- <i>cbiH₃₀-cbiX</i>	This work
pSJM164	<i>P_{xyIA}-cbiH₆₀</i> sub-fragment cloned from pSJM013 into <i>EcoNI</i> and <i>EagI</i> of pSJM064, creating pWH1520- <i>cbiH₆₀-cbiX</i>	This work
pSJM166	<i>cbiX-J-C-D-E-T-L-F-G-A-cysG^A-cbiY-btuR</i> fragment sub-cloned from pAR8766 with <i>PacI</i> and <i>SacI</i> into pSJM161 giving <i>P_{xyIA}-cbiH₃₀-cbiX-J-C-D-E-T-L-F-G-A-cysG^A-cbiY-btuR</i>	This work
pSJM167	<i>cbiX-J-C-D-E-T-L-F-G-A-cysG^A-cbiY-btuR</i> fragment sub-cloned from pAR8766 with <i>PacI</i> and <i>SacI</i> into pWH1520- <i>cbiX</i> , giving <i>P_{xyIA}-cbiX-J-C-D-E-T-L-F-G-A-cysG^A-cbiY-btuR</i>	This work

pSJM169	<i>cbiX-J-C-D-E-T-L-F-G-A-cysG^A-cbiY-btuR</i> fragment sub-cloned from pAR8766 with <i>PacI</i> and <i>SacI</i> into pSJM164 giving <i>P_{xylA}-cbiH₆₀-cbiX-J-C-D-E-T-L-F-G-A-cysG^A-cbiY-btuR</i>	This work
pSJM173	pStop31623hp promoter sub-cloned from pKMBm19 into <i>BglI</i> and <i>SpeI</i> of pSJM007 creating pStop31623hp- <i>cbiL</i> -His ₆	This work
pSJM175	pStop31623hp promoter sub-cloned with <i>BglI</i> and <i>SpeI</i> into pSJM188 creating pStop31623hp- <i>cbiH₆₀</i> -His ₆	This work
pSJM182	<i>B. megaterium</i> derived <i>hemCDB</i> (2865 bp) PCR fragment, cloned into pGEM [®] -T Easy. Sequencing showed five mutations; two silent, one frameshift (<i>hemC</i>) and two point mutations (<i>hemB</i>). Negative orientation	This work
pSJM189	<i>B. megaterium</i> derived <i>hemCDB</i> (2865 bp) PCR fragment, cloned into pGEM [®] -T Easy. Sequencing showed four mutations; one silent, and three point mutations (one each per gene). Negative orientation (SP6- <i>hemCDB</i> -T7)	This work
pSJM192	<i>XbaI</i> - <i>BamHI</i> <i>hemB</i> fragment in pSJM189 sub-cloned into pSJM183 to repair <i>hemB</i> mutation	This work
pSJM195	<i>B. megaterium</i> derived <i>cbiD</i> PCR fragment, cloned into <i>BglII</i> and <i>SphI</i> of pC-His1622 creating <i>P_{xylA}-cbiD</i> -His ₆ -tag-Stop	This work
pSJM198	<i>B. megaterium</i> derived <i>cbiWH₆₀</i> PCR fragment, cloned into <i>BglII</i> and <i>EagI</i> of pN-His-TEV1622 creating <i>P_{xylA}-His₆-tag-cbiW</i> -Stop- <i>cbiH₆₀</i> -Stop	This work
pSJM203	<i>B. megaterium</i> derived <i>hemCDB</i> PCR fragment, cloned from pSJM189 into <i>SpeI</i> and <i>BamHI</i> of pWH1520- <i>cbiX</i> , creating pWH1520- <i>hemCDB</i>	This work
pSJM209	<i>B. megaterium</i> derived <i>hemCDB</i> PCR fragment, cloned from pSJM203 into <i>EcoNI</i> and <i>BamHI</i> of pMGBm19, creating pMGBm19- <i>hemCDB</i>	This work
pSJM216	<i>B. megaterium</i> derived <i>cbiC</i> PCR fragment, cloned into pGEM [®] -T Easy. Positive orientation (T7- <i>cbiC</i> -SP6)	This work
pSJM220	<i>B. megaterium</i> derived <i>sirABC</i> (2270 bp) PCR fragment, cloned into pGEM [®] -T Easy. Positive orientation (T7- <i>sirABC</i> -SP6). One silent mutation. Small region	This work

	in <i>sirA</i> not sequenced, incomplete	
pSJM222	<i>B. megaterium</i> derived <i>cbiP</i> (1565 bp) PCR fragment, cloned into pGEM [®] -T Easy. Negative orientation (SP6- <i>hemCDB</i> -T7). No mutations	This work
pSJM231	<i>B. megaterium</i> derived <i>cbiC</i> sub-cloned from pSJM216 into <i>Bgl</i> II and <i>Eag</i> I of pN-His-TEV1622 creating P _{<i>xyIA</i>} -His ₆ -tag- <i>cbiC</i> -Stop	This work
pSJM233	<i>B. megaterium</i> derived <i>BMD0328</i> PCR fragment cloned into <i>Bam</i> HI and <i>Sph</i> I of pWH1520- <i>cbiX</i> creating pWH1520- <i>cbiX</i> - <i>BMD0328</i>	This work
pSJM234	<i>B. megaterium</i> derived <i>cbiP</i> fragment cloned from pSJM222 into <i>Spe</i> I and <i>Sph</i> I of pWH1520- <i>cbiX</i> creating pWH1520- <i>cbiP</i>	This work
pSJM236	<i>B. megaterium</i> derived <i>cbiX</i> - <i>BMD0328</i> fragment cloned from pSJM233 into <i>Eco</i> NI and <i>Acc</i> 65I of pSJM209 creating pMGBm19- <i>cbiX</i> - <i>BMD0328</i> - <i>hemCDB</i>	This work
pSJM237	<i>B. megaterium</i> derived <i>cbiET</i> PCR fragment cloned into <i>Bgl</i> II and <i>Sph</i> I of pC-His1622 creating P _{<i>xyIA</i>} - <i>cbiET</i> -His ₆ -tag-Stop	This work

2.3 Microbiological techniques

2.3.1 Sterilisation

Unless noted otherwise, all media was vapour sterilised at 121 °C and 1 bar positive pressure for 20 min. Other substances and solutions were either vapour sterilised or, if temperature sensitive, sterilised by filtration (pore width of the filter was 0.2 µm). Antibiotics that were dissolved in ethanol were assumed to be sterile and stored at -20 °C.

2.3.2 Media for growth of bacterial strains

As a standard media for routine growth all bacterial strains, Luria Bertani (LB) medium (Sambrook & Frisch, 1989) was used as a broth or for solid media 1.5% agar (w/v) was added. Terrific broth was also used as an alternative rich medium for *B. megaterium*.

<u>LB Medium</u>	Tryptone	10 g
	Yeast extract	5 g
	NaCl	5 g
	Fill to 1 L with ddH ₂ O	
<u>Terrific broth</u>	Tryptone	12 g
	Yeast extract	24 g
	Glycerol	4 ml
	Fill to 0.9 l L with ddH ₂ O	

Sterilize by autoclaving and adjust volume to 1 L with filter sterilised 0.17 M KH₂PO₄ and 0.72 M K₂HPO₄.

For expression experiments (Western blotting, vitamin B₁₂ and sqRT-PCR analysis) growth of *B. megaterium* was performed under minimal conditions with M9 medium.

<u>10 × M9 Salts</u>	Na ₂ HPO ₄	60 g
	KHPO ₄	30 g
	NH ₄ Cl	10 g
	NaCl	5 g
	Fill to 1 L with ddH ₂ O	

<u>M9 medium (<i>Bacillus</i> modified)</u>	10 × M9 salts	100 ml
	Glycerol [100% (v/v)]	2 ml
	MgSO ₄ (1 M)	1 ml
	CaCl ₂ (0.1 M)	1 ml
	FeSO ₄ (0.1 M)	1 ml
	MnCl ₂ × 6H ₂ O (25 mM)	40 µl
	Fill to 895 ml with ddH ₂ O	

2.3.3 Liquid Cultures of *E. coli*

Aerobic liquid cultures were inoculated using a single colony from a LB agar plate. The medium was supplemented with the appropriate antibiotics when required. Cultures were shaken at 200 rpm in test tubes or baffled flasks at 37 °C for a maximum of 16 hrs.

2.3.4 Liquid Cultures of *B. megaterium*

Aerobic liquid cultures were inoculated using a single colony from a LB agar plate for pre-cultures. The medium was supplemented with the appropriate antibiotics when required. Incubation occurred at 30-37 °C in baffled flasks at 100 rpm for 14 h. Following overnight growth, main cultures were inoculated at a ratio of 1: 50 to 1: 100 from the pre-cultures. The culture volume varied from 50 ml to 1 L in baffled flasks at 150-250 rpm and 30 °C or 37 °C. The incubation times varied according to the desired optical densities. For anaerobic growth reinforced glass bottles sealed with rubber stoppers and aluminium caps were used with 10 mM NaNO₃ added as an electron acceptor.

2.3.5. Antibiotics and additives

Additives and antibiotics were filter sterilised aseptically (Pore size 0.2 µm) before addition. Antibiotics and additives were made as concentrated stock solutions, filter sterilised and added to the medium (Table 2.3). For antibiotics containing ethanol, these were simply dissolved and stored at -20 °C.

Table 2.3. Antibiotics and additives

(A)

Antibiotic	Solute	Stock Solution	Working Solution
Ampicillin	ddH ₂ O	100 mg ml ⁻¹	100 µg ml ⁻¹ (<i>E. coli</i>)
Chloramphenicol	Ethanol 70% (v/v)	34 mg ml ⁻¹	34 µg ml ⁻¹ (<i>E. coli</i>)
		4.5 mg ml ⁻¹	4.5 µg ml ⁻¹ (<i>B. meg</i>)
Erythromycin	Ethanol 70% (v/v)	10 mg ml ⁻¹	5 µg ml ⁻¹ (<i>B. meg</i>)
Kanamycin	ddH ₂ O	50 mg ml ⁻¹	30 µg ml ⁻¹ (<i>E. coli</i>)
Tetracycline	Ethanol 100% (v/v)	5 mg ml ⁻¹	10 µg ml ⁻¹ (<i>B. meg</i>)

(B)

Additive	Solute	Stock Solution	Working Solution
ALA	ddH ₂ O	1 M	0.05-5 mM
L-arabinose	ddH ₂ O	2% (v/v)	0.02% (v/v)
Betaine	ddH ₂ O	1 M	0.1-10 mM
IPTG	ddH ₂ O	1 M	400 µM
CoCl ₂ × 6H ₂ O	ddH ₂ O	25 mM	1-250 µM
Glutathione	ddH ₂ O	1 M	0.1-1 mM
FeSO ₄	ddH ₂ O (Dilute HCl)	100 mM	100-500 µM
IPTG	ddH ₂ O	1 M	400 µM
MnCl ₂ × 6H ₂ O	ddH ₂ O	25 mM	1 µM
X-Gal	DMF	50 mg ml ⁻¹	40 µg ml ⁻¹
D-Xylose	ddH ₂ O	50% (w/v)	0.5% (w/v)

2.3.6 Spectrometric determination of cell density

For routine estimation of cell concentration the OD_{578nm} was measured to estimate the cell density of *B. megaterium* using the assumption that an OD_{578nm} of 1.0 is equivalent to 1×10^9 cells ml⁻¹.

2.3.7 Storage of bacteria

For long-term storage of bacteria, 650 µl of pre-culture was mixed with 350 µl of 87.5% (v/v) glycerol [Final concentration 30% (v/v) glycerol], kept on ice for 15 min and then stored at -80 °C indefinitely.

2.3.8 Bioassay of vitamin B₁₂ production

An indicator strain of *S. typhimurium* AR2680 or AR3612 was used to quantify cobinamide and cobyrinic acid, respectively, as described by (Raux et al, 1996). Production of these intermediates is an estimate of vitamin B₁₂ production and was quantified from cell lysates of *B. megaterium* and repeated for more than three independent measurements and given as an average.

2.4 Molecular biology techniques

2.4.1 Isolation of genomic DNA from *B. megaterium*

From an overnight culture, approximately 2×10^9 cells were centrifuged at 7500 rpm for 10 min. The cells were then re-suspended in 180 μ l of lysis buffer [20 mM Tris-HCl pH 8.0, 2 mM EDTA, 1.2% (v/v) Triton X-100, 20 mg ml⁻¹ lysosyme], followed by incubation at 37 °C for 30 min. Genomic DNA was extracted then by following the instructions from the DNeasy tissue kit (QIAGEN).

2.4.2 Isolation of total RNA from *B. megaterium*

Approximately 1×10^9 cells were removed from cultures growing in minimal media when an OD_{578nm} of 1.0 was reached. Each condition was performed as triplicate biological repeat and total RNA was extracted from *B. megaterium* cells using the RNeasy extraction kit with the RNAProtect® Bacteria Reagent and on column DNase I treatment (QIAGEN Ltd, Crawley, UK). Total RNA was quantified spectrophotometrically and used for cDNA synthesis.

2.4.3 Determination of DNA concentration

For estimation of the prepared plasmid DNA concentration, the plasmid was enzymatically linearised and visualised on an agarose gel and compared to the DNA standards. The concentration and purity of prepared genomic DNA was determined by measuring the absorbance at 260 nm (A_{260}) and additionally at 280 nm (A_{280}) to account for protein impurities. For a pure DNA solution, an A_{260} of 1 unit corresponded to a concentration of dsDNA of 50 μ g ml⁻¹. The quality of the DNA solution is deduced from the ratio of A_{260} to A_{280} . With an A_{260}/A_{280} ratio of 1.8-2.0, the DNA can be considered as pure.

2.4.4 Determination of RNA concentration

For RNA quantification a nanodrop reader (ThermoFisher) was used to quantify samples of RNA prepared from *B. megaterium* using the RNeasy kit (QIAGEN). For a pure RNA solution, an A_{260} of 1 unit corresponded to a concentration of ssRNA of 44 μ g ml⁻¹. The quality of the RNA solution is deduced from the ratio of A_{260} to A_{280} . With an A_{260}/A_{280} ratio of 1.9-2.1, the RNA can be considered as pure.

2.4.5 Plasmid purification from *E. coli*

Plasmid DNA was prepared from 5-10 ml of liquid overnight LB culture of *E. coli* DH10B carrying the plasmid, using the mini Prep Kit from QIAGEN (QIAGEN) according to the manufacturer's instructions. For protoplast transformation approximately 5 µg was required and was eluted in ddH₂O for the primary protoplast transformation protocol and in 10 mM Tris-HCl pH 8.0 for the secondary protoplast transformation protocol.

2.4.6 Agarose gel electrophoresis

For analytical separation of DNA fragments, agarose gels consisting of 0.7 to 3% (w/v) agarose (according to the expected fragment size) in TAE buffer (40 mM Tris-HCl pH 8.0, 10 mM acetic acid, 1 mM EDTA) containing ethidium bromide (0.5 µg ml⁻¹) was prepared. The DNA samples were mixed with 5 x DNA loading dye (Bioline) to facilitate loading and to indicate the progress of the samples in the gel. Hyperladder I (Promega) was used as a DNA standard marker. Depending on the size of the gel, a voltage of 80–100 V was applied. The DNA fragments migrate towards the anode with a velocity that is proportional to the negative logarithm of their length. After electrophoresis, The DNA was detected via its fluorescence under UV light and a photo was taken.

2.4.7 Amplification of DNA by polymerase chain reaction (PCR)

For amplification of DNA by PCR, oligonucleotide primers for each DNA fragment of interest were designed. Recognition sequences for restriction endonucleases were inserted via these primers at both ends of the corresponding fragment. Also additional restriction sites were added for a multi-step cloning procedure. All oligonucleotide PCR primers are listed in the Table 2.4. Recognition sequences of restriction endonucleases are coloured in red or blue. Oligonucleotides were purchased from Fisher Scientific, Leicestershire.

2.4.8 Oligonucleotide primers for PCR

Oligonucleotides were designed and obtained from Fisher. They were designed to incorporate one or more restriction sites for subsequent cloning steps. For incorporation of a C-terminal His₆-tag, the stop codon on the reverse primer was removed and extra bases were added to keep the gene in frame.

Table 2.4. Oligonucleotides for PCR. Primary restriction enzyme sites are coloured in red, whilst secondary restriction enzyme sites are coloured in blue

Primer Name	Sequence (5'-3')	Restriction Site
<i>Bm cbiH 5</i>	CATAGATCTGGAAAGGTAAACTGTTAGTTAT TGG	<i>BglII</i>
<i>Bm cbiH 3</i>	CATCGGCCGATCTCCGCACGGAGCTG	<i>EagI</i>
<i>Bm cbiL 5</i>	CATAGATCTGGAACATGATTGGCAACATTG	<i>BglII</i>
<i>Bm cbiL 3</i>	CATGCATGCAATTTTCGCACCACCATTAG	<i>SphI</i>
<i>Bm cbiX 5</i>	ATGTCTAGAGGAGATTAAGAAATGGGAGG	<i>XbaI</i>
<i>Bm cbiX 3</i>	ATGGGTACCAGTATCATTTTAGCTCGCCGA CC	<i>SpeI, KpnI</i>
<i>Bm sirC 5</i>	CATACTAGTGAAGGAGATATACCATGGGC	<i>SpeI</i>
<i>Bm sirC 3</i>	ATGGCATGCTCCAGCGGTATTCGAAGC	<i>SphI</i>
<i>Bm sirA 5</i>	CATAGATCTCTATGGGGAAAGTATATCTAGT C	<i>SpeI</i>
<i>Bm sirC 3</i>	ATGAGATCTTCCAGCGGTATTCGAAGC	<i>BglII</i>
<i>Bm cbiX 3</i>	ATGGCATGCACTAGTCAATTTAGCTCGCCGA CC	<i>SphI</i> <i>SpeI</i>
<i>Bm cbiG 5</i>	CAC AGA TCT GGA TGA TTC AAC TCG AAG AAG	<i>BglII</i>
<i>Bm cbiG 3</i>	CATGCATGCTTGTTCATATGGAATAAGTGC	<i>SphI</i>
<i>Bm cbiJ 5</i>	CACAGATCTGGATGATTTTATTGTTAGCTGG	<i>BglII</i>
<i>Bm cbiJ 3</i>	CATGCATGCTTACTCCTGGTTTGG	<i>SphI</i>
<i>Bm cbiW 5</i>	CATAGATCTTAATGAGTGTGACGACG	<i>BglII</i>
<i>Bm cbiW 3</i>	ATGGCATGCTTATGCGTTACTTTTAC	<i>SphI</i>
<i>Bm cbiF 5</i>	GCGTGTACAATGAAGTTATACATAATCGGA GCTGG	<i>BsrGI</i>
<i>Bm cbiF 3</i>	CATGGATCCTTTCCGATTTCACTCC	<i>BamHI</i>
<i>Bm cbiD 5</i>	CACAGATCTGGATGAAGGAAGTCGCAAAAG AACC	<i>BglII</i>
<i>Bm cbiD 3</i>	GTGGCATGCTTATTGCCATGTTGCACC	<i>SphI</i>
<i>Bm cbiP 5</i>	CATGAGCTCGTATGGAAAAAGTGAAAAAAG G	<i>SacI</i>
<i>Bm cbiP 3</i>	ATGGCATGCAATACTTTTTGTGTTTGCTCC	<i>SphI</i>
<i>Bm hemC 5</i>	CATGGTACCCGTAATTAGGAGGACAACATGC G	<i>KpnI</i>
<i>Bm hemB 3</i>	ATGGGATCCATATCTACTGATTTAAATGCAC G	<i>BamHI</i>
<i>Bm hemL 5</i>	CATGGATCCTATTTTTATGGAACGAGGAAAA GG	<i>BamHI</i>
<i>Bm hemL 3</i>	ATGGCATGCGCTGTTTATCCACGTAACCTCG	<i>SphI</i>
<i>Bm 0328 3</i>	CGCGGATCCGGTAAAAAAGGAGTGTAGGA AATG	<i>BamHI</i>
<i>Bm 0328 3</i>	ATGGCATGCGGTACCCAACAGATTAAGCTT TCG	<i>SphI</i> <i>Acc65I</i>
<i>Bm cbiET 5</i>	CACAGATCTGGATGGCAATTAATTAATTGG	<i>BglII</i>
<i>Bm cbiET 3</i>	GTGGCATGCTTTTCTCTCTTTTGGCTG	<i>SphI</i>

For amplification of the DNA of interest, PCR reactions of a total volume of 50 μl were prepared. For amplification of DNA fragments between 500-3000 bp, the FastStart High Fidelity PCR SystemTM (Roche) was used which consists of a blend of the FastStart thermostable Taq DNA polymerase (*Thermus aquaticus*) and Tgo DNA polymerase (*T. aquaticus*) with proofreading activity. PCR reactions were prepared as described in Table 2.5.

Table 2.5. PCR reaction mix

ddH ₂ O	35.5 μl
10 \times PCR Buffer (Roche)	5 μl
DNA template (100-500 ng)	3 μl
Forward primer (10 μM)	2 μl
Reverse primer (10 μM)	2 μl
dNTP's (5 mM)	2 μl
Taq/Tgo DNA polymerase (3.5 U/ μl)	0.5 μl

For standard PCR amplifications up to 2500 bp the following protocol was followed (Table 2.6).

Table 2.6. PCR protocol for standard amplifications (Up to 2500 bp)

	Temp	Time	Cycles
Denaturation	95 °C	2 min	1
Denaturation	95 °C	30 sec	
Annealing	55-59 °C	30 sec	30-45
Elongation	72 °C	0.5-2 min	
Final elongation	72 °C	5 min	1
Cooling	4 °C	Indefinite	

2.4.9 Digestion of plasmid DNA with restriction endonucleases

Digestion of plasmid DNA was performed with restriction endonucleases purchased from Promega (Southampton, UK) or New England Biolabs (NEB, Ipswich, USA). Reaction conditions including buffer, temperature and concentration of enzyme and DNA was followed as according to the manufacturer's instructions. Digested DNA was separately by agarose gel electrophoresis.

2.4.10 Purification of PCR products and linear plasmid fragments

Following agarose gel electrophoresis, DNA fragments of the correct size were excised from the gel and purified using the QIAquick Gel Extraction Kit (QIAGEN).

2.4.11 Ligation of DNA – direct cloning

Purified DNA fragments with compatible ends were ligated using the Promega Rapid Ligation buffer and T₄ DNA ligase. In general a 10 µl mix was prepared by combining an excess of insert (0.5-1 µg) with vector (50-250 ng) in the 2 × ligation buffer (Promega) and 1-5 U of T₄ DNA ligase. A control ligation, where the parent vector and no insert was also prepared in parallel. In general if there were more colonies on the vector plus insert ligation than vector alone, then it was assumed the ligation was successful, so colonies could be picked for preparation of plasmid DNA.

2.4.12 Cloning into pGEM[®]-T Easy – indirect cloning

Genes of interest obtained from PCR reaction were cloned using the pGEM[®]-T Easy vector system from Promega. 2 µl of PCR product was mixed with 1 µl of pGEM[®]-T Easy vector and ligated overnight. Commercial super competent *E. coli* JM109 were transformed with the ligation mixture and recombinants were identified by blue/white screening when plating the cells on LB plates containing ampicillin, 0.5 mM IPTG and 80 µg ml⁻¹ X-Gal. Please refer to the pGEM[®]-T Easy technical manual for more detail. The gene(s) were then subsequently sequenced before further subcloning.

2.4.13 Preparation of competent cells

From an overnight pre-culture of *E. coli* (DH10B and BL21* derivative strains) was prepared and sub-cultured (1:100) dilution into 50 ml of LB medium. This was grown at 37 °C, 200 rpm until an OD_{600nm} of 0.3 was reached. The culture was then placed on ice for 30 min, before centrifugation for 10 min, 4 °C, 4000 rpm. The cells were then resuspended in ice-cold freshly filter sterilised 0.1 M CaCl₂. The cells were left on ice for 30 min before centrifugation again. Cells were then resuspended in 2 ml of ice cold 0.1 M CaCl₂, 15% glycerol, before aliquots of 30-50 µl were prepared in sterile eppendorfs and stored at -80 °C.

2.4.14 Transformation of *E. coli*

Competent cells were defrosted on ice for 10 min. 0.5 µl of plasmid DNA, or 3-5 µl of ligation reaction was mixed with 30-50 µl of competent cells and left on ice for 20 min. These were heat shocked at 42 °C for 50 seconds. 150 µl of LB broth was added and incubated at 37 °C for 60 min. 100 µl aliquots were spread plated onto LB agar with appropriate antibiotics and incubated at 37 °C overnight.

2.4.15 DNA sequencing

After successful cloning and confirmation by restriction digest analysis, all DNA prepared from PCR was sequenced to check for mutations. Sequencing was performed with an ABI 3730xl (Applied Biosystems, Perkin Elmer, Boston, USA) by either GATC Biotech (Konstanz, Germany) or Beckman Coulter Genomics (Takeley, UK).

2.4.16 Oligonucleotide primers used for sequencing

Oligonucleotide primers for sequencing were either provided by Dr Rebekka Biedendieck or purchased from Fisher Scientific. Sequencing oligonucleotides are shown in Table 2.7.

Table 2.7. Sequencing oligonucleotides

Primer Name	Sequence (5'-3')	Primer used for
SeqpWH1520_for	ATGATGAGATAAAGTTAGTTTATTGG	pWH1520 derivatives (cloning in MCS)
SeqpWH1520_rev	GTTTGCGCATTACAGTTCTCC	pWH1520 derivatives (cloning in MCS)
T7 universal primer	TAATACGACTCACTATAGGG	pGEM-T Easy derivatives
SP6 universal primer	ATTTAGGTGACACTATAG	pGEM-T Easy derivatives
hemCDBseq1	CACATTACGGGAAAAGACCCCATTCG	<i>B. megaterium</i> hemD
hemCDBseq2	CATGAAAGATTTACAATTTACACGTC ATCG	<i>B. megaterium</i> hemB

sirBseq1	TGGGGAACATGTGCGGACC	<i>B. megaterium</i> <i>sirB</i>
nhlFseq1	CCATGACCATGATCACG	<i>B. megaterium</i> <i>nhlF</i>

2.4.17 Semi-quantitative reverse transcription-PCR

A control without reverse transcription was prepared to make sure no PCR products were formed from genomic DNA contamination to validate results. An additional control with no template RNA was also performed. Reactions of each biological repeat were performed in duplicate. The expression level of the genes were determined by semi-quantitative reverse transcription-PCR (sqRT-PCR) using the iScript One-Step RT-PCR Kit (Bio-Rad Laboratories, Inc., Hemel-Hempstead, UK) with SYBR green label on a Chromo4 Real-Time PCR Detection System (Bio-Rad Laboratories, Inc.). Each sqRT-PCR reaction was carried out in a total volume of 25 μ l and contained 12.5 μ l of 2 \times SYBR green sqRT-PCR reaction mix, 100 nM of the appropriate primers (Fisher) and 50 μ g of total RNA template. Reverse transcription was achieved by inclusion of an initial incubation at 50 $^{\circ}$ C for 10 min. The amplification protocol consisted of an initial denaturation of 95 $^{\circ}$ C for 2 min, followed by 40 cycles of 95 $^{\circ}$ C denaturation for 30 sec and 55 $^{\circ}$ C annealing for 30 seconds and 72 $^{\circ}$ C elongation for 15 sec. Subsequently, dissociation melt curves were performed by stepwise increment of the temperature from 55 $^{\circ}$ C to 95 $^{\circ}$ C. Crossing point, quantification, and melting curve analysis were undertaken using the Opticon Monitor software (Bio-Rad Laboratories, Inc.). The chosen housekeeping gene was glyceraldehyde-3-phosphate dehydrogenase (*gap*). Primers are listed in Table 2.8.

2.4.18 Oligonucleotide primers used for sqRT-PCR

Oligonucleotides were carefully designed and analysed for secondary structure, homodimers, heterodimer formation. Oligonucleotides used for sqRT-PCR are shown in Table 2.8.

Table 2.8. Oligonucleotide primers used for sqRT-PCR

Primer Name	Sequence (5'-3')
<i>GAPDH_for</i>	TGCTAACATGCTTGCTCACC
<i>GAPDH_rev</i>	GCTCAATTATCTTGGGGCAA
<i>cbiX_for</i>	CAGTTTTATTTGTTCGGTCATGG
<i>cbiX_rev</i>	GAACGCCCAAATGTAAGTCAAG
<i>btuR_for</i>	CAACCGCTGCTTTAGGTCTTG
<i>btuR_rev</i>	GATTCAAACAGGAGTCGGCTTC

2.4.19 Protoplast transformation

Two methods were used for the transformation of *B. megaterium* DSM319 during this PhD thesis. The original method as described in (Kunkel, 2004) was replaced with a modified method that was more successful during this work. A negative control transformation was also prepared in parallel and grown without antibiotics, to ascertain the viability and efficiency of the protoplast regeneration.

2.4.20 Preparation of protoplasts

Based on the method described (Kunkel, 2004), LB and SMMP mediums are replaced with Hyp medium and growth was performed at 30 °C. Hyp medium was pre-incubated on ice and all centrifugations were done at 4 °C.

A pre-culture of *B. megaterium* DSM319 was grown overnight in 50 ml Hyp medium, 100 rpm, 30 °C, for a maximum of 16 hrs. 1 ml of pre-culture was then inoculated into 50 ml of Hyp medium and incubated at 30 °C, 200 rpm until an OD₅₇₈ of 1.0 was reached. Cells were centrifuged at 4000 rpm, 4 °C for 15 min. Supernatant was removed and cells were re-suspended in 5 ml of Hyp medium. 100 µl of lysosyme (5 mg ml⁻¹ dissolved in Hyp medium and filter sterilised) was added and incubated 30 °C, 100 rpm for 10-20 min, until at least 50% of cells were protoplasts. Cells were incubated on ice for 30 min and then spun at 2500 rpm, 4 °C for 10 min. Supernatant was removed and cells were re-suspended in 5 ml of Hyp medium. This was repeated three times to remove lysosyme. After final re-suspension in 5 ml of Hyp medium, 652.5 µl of 100% glycerol was added and

mixed. 500 μ l aliquots were prepared and either used fresh or stored at $-80\text{ }^{\circ}\text{C}$ (Viable for 2 months).

2.4.21 Transformation

DNA ($\sim 5\text{ }\mu\text{g}$) was eluted in 50 μ l of 5 mM Tris-HCl pH 8.0 and mixed with 500 μ l of protoplasts. This was added to 1.5 ml of 40% PEG-P solution, mixed carefully and incubated at room temperature for 4 min. To this, 5 ml of pre-chilled Hyp medium was added, mixed carefully and protoplasts harvested by centrifugation (4 min, $1400\times g$, $4\text{ }^{\circ}\text{C}$). Supernatant was discarded carefully and the pellet was re-suspended in 1 ml of Hyp-medium. Protoplasts were regenerated by incubating for 2-3 hr at $30\text{ }^{\circ}\text{C}$ and 100 rpm. After regeneration protoplast transformation aliquots (100 μ l and 900 μ l) were mixed with 7 ml of pre-warmed Hyp top-agar ($42\text{ }^{\circ}\text{C}$). This was then poured onto pre-warmed ($37\text{ }^{\circ}\text{C}$) antibiotic Hyp agar plates, dispersed evenly and incubated at $37\text{ }^{\circ}\text{C}$ for 24 hrs.

2.4.22 Media

<u>Prot-Medium (pH 7.5)</u>	NH ₄ Cl	1 g
	Tris-HCl	12 g
	KCl	35 mg
	NaCl	58 mg
	MgSO ₄ x 7 H ₂ O	267 mg
	MgCl ₂ x 6 H ₂ O	4.67 g
	Fill to 864 ml with ddH ₂ O	
<u>Hyp-Medium</u>	Prot-Medium	864 ml
	Sucrose [50% (w/v)]	136 ml
	Glucose [20% (w/v)]	10 ml
	Yeast extract [10% (w/v)]	6 ml
	KH ₂ PO ₄ (0.5 M)	2 ml
	MnSO ₄ (13.25 mM)	100 μ l
<u>PEG-P solution</u>	PEG-6000	40% (w/v)
	Dissolved in Hyp-Medium	
<u>Hyp agar plates</u>	Agar	1.7% (w/v)
	Dissolved in Prot-Medium, autoclaved, then prepared with Hyp-Medium components aseptically	

Hyp top-agar

Agar 0.9% (w/v)
Dissolved in Prot-Medium, autoclaved, then prepared
with Hyp-Medium components aseptically

2.5 Protein Biochemistry

2.5.1 Recombinant protein production in *E. coli*

An overnight pre-culture was started as previously described. The pre-culture was inoculated 1:50 into 50-1000 ml of rich medium (typically LB) and grown at 37 °C, 160 rpm for 2-4 hr, then induced with 0.4 mM IPTG and left overnight at 16 °C and 160 rpm.

2.5.2 Recombinant protein production in *B. megaterium*

An overnight pre-culture was started as previously described. The pre-culture was inoculated 1:50 into 50-1000 ml of rich medium (typically LB) and grown at 30 °C, 160 rpm for 2-2.5 hr, then induced with 0.16-0.5% (w/v) D-xylose and left overnight at 30 °C and 160 rpm.

2.5.3 Preparation of cell-lysates for analysis of intracellular protein

For SDS-PAGE analysis, $3 \times OD_{578}$ equivalents of cells were centrifuged at 4000 rpm, 4 °C for 10 min and the cells re-suspended in 30 μ l lysis buffer (50 mM Tris-HCl, pH 8.0, 1 mg ml⁻¹ lysosyme). The re-suspended cells were incubated at 37 °C for 45 min, with rapid mixing. The cell debris was centrifuged at 13,500 rpm, 4 °C for 20 min by centrifugation. Protein concentration was estimated by an A₂₈₀ measurement or by Bradford assay. For SDS-PAGE analysis and Western blotting, protein samples were diluted in SDS-loading buffer (Sigma) and boiled for 5 min. Approximately 50 μ g protein was loaded per sample.

2.5.4 His₆-bind column purification

All end cultures were centrifuged at 4000-5000 rpm, 4 °C for 20 min. Cells were then re-suspended in 15 ml of binding buffer (20 mM Tris-HCl, 500 mM NaCl, 5 mM Imidazole), followed by sonication on ice for 5 minutes. Cell lysates were centrifuged at 17500 rpm, 4 °C for 20 min. The supernatant was then purified using 5 ml of pre-charged nickel chelated sepharose (GE Healthcare). This was washed with 50 ml of binding buffer, followed by increasing concentrations of imidazole, 25 ml wash buffer I (50 mM imidazole) and 25 ml of wash buffer II (100 mM imidazole). Elution was performed with buffer containing 400 mM imidazole.

Fractions from each purification step were analysed by SDS-PAGE and stained with colloidal Coomassie Blue G-250 stain.

2.5.5 Anaerobic purification of proteins

Proteins containing Fe-S clusters and all *in vitro* studies concerning incubation of intermediates with purified and crude enzymes were prepared in a glovebox (Belle Technologies), with O₂ levels at less than 2 ppm. All buffers and solutions were purged with argon prior to being placed inside the glovebox.

2.5.6 Buffer exchange (Desalting)

Crude cell lysates and purified protein were desalted on a pre-packed Sephadex G25 column (PD-10, bed volume 8.3 ml). The column was pre-equilibrated in 25 ml of buffer into which the protein was to be exchanged. 2.5 ml of protein was applied onto the column, before eluting in 3.5 ml of buffer. Column was washed with ddH₂O.

2.5.7 A₂₈₀ protein concentration estimation

Protein concentrations were estimated by the method described by Warburg and Christian (Warburg, 1942). The absorbance at 280 nm was measured and the concentration estimated from the calculated extinction coefficient for the protein using the following equations.

$$A_{280} = \text{Concentration (M)} \times \text{Extinction coefficient (M}^{-1} \text{ cm}^{-1}) \times \text{Path Length (cm)}$$

2.5.8 Gel filtration

An Amersham Biosciences P-920 FPLC chromatography system was used for gel filtration purification. A SuperdexTM G200 column (GE Healthcare) was pre-equilibrated with 1.2 column volumes of the component buffer of which the protein of interest was stable in. Purified protein was filtered prior to injection (1-2 ml) onto the column. The protein was eluted at a flow rate of 0.5 ml min⁻¹ into 1 ml fractions. An online detector measuring the absorbance at 280 nm, was used to detect protein elution from the column. Fractions were further analysed by SDS-PAGE to determine purity.

2.5.9 SDS-PAGE

Solutions for protein acrylamide gels

<u>10 × running buffer</u>	Tris-HCl	30 g L ⁻¹
	Glycine	144 g L ⁻¹
<u>2 × SDS loading buffer</u>	0.5 M Tris-HCl, pH 6.8	6.0 ml
	Glycerol	4.8 ml
	SDS 10% (w/v)	9.6 ml
	Bromophenol blue 0.05% (w/v)	1.2 ml
	ddH ₂ O	24.0 ml

Add 14 µl β-mercaptoethanol per ml of 2 x SDS loading buffer.

<u>Fixing solution</u>	Ethanol	300 ml
	Acetic acid	100 ml

Made up to 1 L with ddH₂O.

<u>Native sample buffer</u>	0.5 M Tris-HCl, pH 6.8	6.0 ml
	Glycerol	4.8 ml
	Bromophenol blue 0.05% (w/v)	1.2 ml
	ddH ₂ O	24.0 ml

Table 2.9. Denaturing molecular mass marker Dalton VII

Protein	Molecular weight (Da)
Bovine serum albumin	66,000
Ovalbumin	45,000
G-3-P dehydrogenase	36,000
Carbonic anhydrase	29,000
Trypsinogen	24,000
Trypsin inhibitor	20,100
α -Lactalbumin	14,200

<u>Coomassie blue stain:</u>	Trichloroacetic acid (100%)	250 ml
	Coomassie blue R250	0.60 g
	SDS	0.10 g
	Tris-HCl	0.25 g
	Glycine	0.15 g

Made up to 500 ml

Table 2.10. Recipe for SDS gels

SDS GELS						
Running gels	8%	10%	12.5%	15%	Stacking gel	5%
ddH ₂ O (ml)	5.7	4.7	3.4	2.2	ddH ₂ O (ml)	3.4
30% Acrylamide (ml)	4	5	6.3	7.5	30% Acrylamide (ml)	1.5
1.5 M Tris-HCl, pH 8.8 (ml)	3.8	3.8	3.8	3.8	0.5 M Tris-HCl, pH 6.8 (ml)	1.9
10% (w/v) SDS (ml)	1.5	1.5	1.5	1.5	10% (w/v) SDS (ml)	0.75
10% (w/v) APS (ml)	0.15	0.15	0.15	0.15	10% (w/v) APS (ml)	0.075
TEMED (ml)	0.01	0.01	0.01	0.01	TEMED (ml)	0.01
NATIVE GELS						
Running gels	3%	5%	7%	9%	Stacking gel	5%
ddH ₂ O (ml)	9.7	8.7	7.8	6.7	ddH ₂ O (ml)	4
30% Acrylamide (ml)	1.52	2.55	3.45	4.56	30% Acrylamide (ml)	1.5
1.5 M Tris-HCl, pH 8.8 (ml)	3.8	3.8	3.8	3.8	0.5 M Tris-HCl, pH 6.8 (ml)	1.9
10% (w/v) APS (ml)	0.15	0.15	0.15	0.15	10% (w/v) APS (ml)	0.075
TEMED (ml)	0.01	0.01	0.01	0.01	TEMED (ml)	0.01

2.5.10 2D-PAGE

50 ml of cell culture was sedimented by centrifugation and re-suspended in 2 ml of lysis buffer (7 M urea, 2 M thiourea, 4% CHAPS, 1% DTT, 0.8% Pharmalyte™ pH 3–10, and 5 mM Pefabloc). Cells were disrupted by ultra-sonication at 4 °C for 8 min with 45 sec pulses and 30 sec rests. The first-dimensional isoelectric focusing was performed with IPGPhor IEF at 21 °C (GE Healthcare) following the method in Table 2.11. 100 µg of protein extract was loaded onto Immobiline DryStrip gels pH 3-10 (GE Healthcare) and IEF/SDS-PAGE was performed as previously described, with slight modifications (Wang *et al.*, 2005).

Table 2.11. IEF method

Step	Transition	Tension (V)	Time (hr)
1	Step and hold	30	14
2	Gradient	200	0.45
3	Step and hold	500	0.45
4	Step and hold	1000	0.45
5	Gradient	8000	1
6	Step and hold	8000	< 9
Finish 21 °C, 49 µA, 1 W, Stop between 8000-10000 Vhr			

2.5.11 Bradford assay for 2D-PAGE

Protein sample (1 µl) was added to 9 µl of 2D lysis buffer. Then 10 µl of 0.1 M HCl and 80 µl of ddH₂O was added. This mixture was then added to the Bradford reagent (1:4 diluted with ddH₂O) before mixing. This was left for 5 min, before the absorbance at 595 nm was measured. BSA (5-50 µg) was used as a standard to estimate protein concentration.

2.5.12 MALDI-TOF

Gel pieces were cut and subjected to alkylation and reduction, before in-gel tryptic digestion, as described before (Wang *et al.*, 2005). Samples were then analysed by a Bruker Ultraflex time-of-flight mass spectrometer and the peptide mass fingerprint was submitted to the Mascot search engine and database. Scores above 82 are considered statistically significant (Perkins *et al.*, 1999).

2.5.13 Western blotting

Protein samples (Prepared and boiled in $2 \times$ SDS loading buffer) were loaded onto SDS-PAGE and run at 200 V for 1.5-2 hrs. When running a Western blot 7.5 μ l of rainbow marker was loaded. For Western blotting two replica SDS-PAGE gels are required for comparing. The filter paper, fibre pads and gel are pre-soaked in transfer buffer ($1 \times$ SDS running buffer, 20% methanol) whilst the nitrocellulose membrane is equilibrated in methanol for 10 seconds, rinsed in ddH₂O and equilibrated in transfer buffer. A fibre pad is aligned on grey panel of opened cassette holder and saturated with transfer buffer. Then sequentially filter paper, gel, nitrocellulose membrane, filter paper and fibre pad are aligned, whilst taking care to remove all bubbles with transfer buffer. The cassette is closed and loaded into the tank, which is filled with transfer buffer and a Bio-Ice cooling unit. It was run for 1 hr at 100 V. For Western blot analysis, proteins from a SDS-PAGE gel were transferred to a polyvinylidene fluoride (PVDF) membrane. A HisProbeTM-horseradish peroxidase (HRP) kit (Thermo Fisher) was used which detects polyhistidine-tagged fusion proteins via one step incubation. All incubations were prepared with Tris buffered saline-Tween[®]-20 (TBST, Thermo Fisher) composed of 25 mM Tris, 0.15 M NaCl, pH 7.6, 0.05% Tween[®]-20, and carried out at room temperature with gentle agitation. For high sensitivity enhanced chemiluminescence was used following instructions provided by Thermo Fisher. Alternatively colorimetric detection was used after the wash-steps. For this the membrane was incubated with the detection reagent [Prepared in TBST: 0.05% (w/v) 4-chloro-1-naphthol, 16.7% methanol (v/v), 0.015% (v/v) H₂O₂] in the dark for a maximum of 30 min and then washed with ddH₂O.

2.6 Biochemical Techniques

2.6.1 Anaerobic techniques

All buffers and solutions were prepared by bubbling with argon for 40 min prior to being placed inside the glovebox port and purged with nitrogen. For all solutions not requiring pH adjustment, these were taken into the glovebox as powders, before dissolving in anaerobic water. The glovebox was constantly maintained at or below 2 ppm of O₂.

2.6.2 Preparation of cobalt(II)-factor III

The vector pET-coco-2-*cobA-hemB-hemC-hemD-sirC* was provided by Dr Stefanie Frank. This was modified from pETcoco-2-ABCDC, which was originally constructed by Dr Evelyne Deery (Frank *et al.*, 2007). In addition it contains the *cbiX^S* and *cbiL* from *M. thermoautotrophicus*, fused to N-terminal His₆-tags. After production of the proteins in *E. coli*, these were purified on a single His₆-bind column. Then after transfer to the glovebox, the proteins were buffer exchanged in Buffer A (20 mM Tris-HCl, 100 mM NaCl). These proteins were then incubated with ALA, SAM, NAD⁺ and cobalt, at a temperature of 37 °C (please see Appendix V for further details).

2.6.3 Ultraviolet-visible (UV-Vis) spectrophotometry

All UV-Vis spectra were recorded on a Varian Cary 50 Bio UV-Vis spectrophotometer over a range of 280-700 nm. For UV/Vis spectra obtained in tandem with mass spectra, please see section 2.6.7.

2.6.4 Electron paramagnetic resonance (EPR)

EPR spectra were recorded in collaboration with Dr Steve Rigby (Manchester University). Samples were prepared as described in the text and then frozen in liquid nitrogen. EPR experiments were performed on a Bruker ELEXSYS E500 spectrometer operating at X-band, employing a Super High Q cylindrical cavity (Q factor ~ 16,000) equipped with an Oxford Instruments ESR900 liquid helium cryostat linked to an ITC503 temperature controller. The iron-sulphur spectra of CbiH₆₀ was recorded at 12 K employing a microwave power of 1 mW, modulation

frequency of 100 kHz and modulation amplitude of 5 G. Spectra of CbiH₆₀ was also recorded at 70 K, which resulted in a loss of signal. This indicated the iron-sulphur centre was of the 4Fe-4S class.

For calculation of the g-value (g_e) the following equation is equation (1) is used. Where h is Plancks constant (6.626×10^{-34} J.s), ν is the electromagnetic radiation frequency (X-band 9.4 GHz), μ_β is the Bohr magneton constant ($9.27400915 \times 10^{-24}$ J.T⁻¹) and B_0 is the magnetic field expressed as Tesla (1 T = 1000 Gauss). The g-value is a dimensionless unit that relates the orientation of the magnetic field. The relationship between g_e and B_0 is inversely proportional and represented by following equation.

$$g_e = h\nu / \mu_\beta B_0$$

2.6.5 Anion-exchange chromatography

5 ml of diethylaminethyl (DEAE)-sepharose solution (Sigma) was loaded onto an empty PD-10 column and washed with 10 ml of wash solution I (20 mM Tris-HCl pH 8.0). Next the supernatant was loaded onto the column and the flow-through collected. The column was subsequently washed with 25 ml of wash solutions with an increasing salt concentration (50 mM - 1 M), and flow-through fractions were collected.

2.6.6 RP-18 chromatography

LiChroprep® RP-18 (Merck) was used routinely for preparative reverse phase chromatography. 5 g of LiChroprep® RP-18 was placed inside an empty PD-10 column, before washing in the following order: 25 ml hexane, 25 ml ethanol, 25 ml 1% acetic acid (v/v). Alternatively for smaller quantities of tetrapyrroles, mini Sep-Pak^(R) C₁₈ cartridge (Waters) were used and prepared using the same method. Before purification, the tetrapyrroles were acidified with acetic acid and centrifuged to remove precipitated material. The supernatant was applied to the media and then washed with 1% acetic acid (v/v). The intermediate was eluted in 40-50% ethanol, before drying and re-suspension in 1% (v/v) acetic acid. The sample was then ready for HPLC-MS analysis.

2.6.7 Reverse phase HPLC

10-100 μl of sample was injected on to an Ace 5 AQ column (2.1 x 150 mm, 5 μm , Advanced Chromatography Technologies) that was attached to an Agilent 1100 series HPLC coupled to a microTOF-Q (Bruker) mass spectrometer and equipped with online diode array and fluorescence detectors and run at a flow rate of 0.2 ml min^{-1} . Tetrapyrroles were routinely separated with a linear gradient of acetonitrile in 1% (v/v) acetic acid or 0.1% TFA. A diagram of the LC timetable for separating tetrapyrroles (Method A) is shown in Figure 2.1.

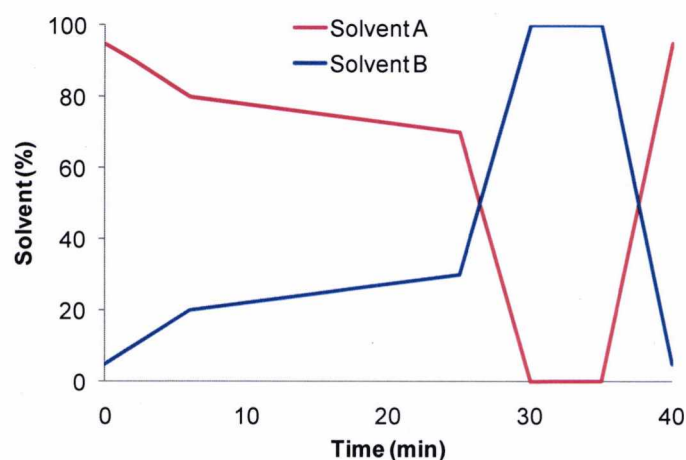


Figure 2.1. Method A - LC timetable for routine separation of tetrapyrroles using 0.1% TFA (solvent A) and acetonitrile (solvent B). Method designed by Dr Andrew Lawrence.

2.6.8 Electrospray ionisation mass spectrometry (ESI-MS)

Mass spectra were obtained using an Agilent 1100 liquid chromatography system connected to a Bruker microTOF II MS, using electro-spray ionisation in positive mode. UV-Vis spectra were monitored by DAD-UV detection.

2.6.9 Inductively coupled plasma mass spectrometry (ICP-MS)

ICP-MS was carried out through our collaboration with Professor Nigel Robinson and Dr Kevin Waldron (Newcastle University), for method details please see (Waldron *et al.*, 2009).

2.6.10 Iron determination by colorimetric assay

Following a protocol described by (Riemer *et al.*, 2004), iron in CbiH₆₀ was quantified by a ferrozine-based colorimetric assay, as follows.

Prepare iron releasing agent by mixing 1.4 M HCl with 4.5% (w/v) KMnO_4 (50:50), preparing under a fume hood (Cl_2 gas evolved). Incubate 100 μl of purified protein with 100 μl of iron releasing reagent and 100 μl of 10 mM HCl at 60 °C for 3 hrs under a fume hood. Cool to room temperature and add 30 μl of iron detection reagent (6.5 mM ferrozine, 2.5 M ammonium acetate, 1 M ascorbic acid). Dilute to 1 ml with 10 mM HCl and measure the absorbance at 550 nm. In order to quantify iron content, FeCl_3 (0-300 μM) dissolved in 10 mM HCl is used as a standard.

Chapter 3

The Regulation of Vitamin B₁₂ Biosynthesis in *B. megaterium*

3.1 Introduction

Vitamin B₁₂ (cobalamin), the anti-pernicious anaemia factor, is a highly complex tetrapyrrole that can only be synthesised by microorganisms. This requires approximately 30 enzymatic steps and can comprise up to 1% of an entire genome (Roth *et al.*, 1993). Many microorganisms encode coenzyme B₁₂ dependent enzymes including the methionine synthase (MetH), methylmalonyl-CoA mutase (MutAB), class II ribonucleotide reductase (NrdJ), 1,2-propanediol dehydratase (PduCDE) and ethanolamine ammonia lyase (EutBC). Cobalamin-dependent enzymes offer the advantage of high efficiency for their methylation and isomerisation reactions.

In prokaryotes, gene expression is predominantly controlled at the level of transcription initiation by a *trans*-regulatory element (Schilling *et al.*, 2004). This involves the interaction of a specific DNA sequence with a regulatory protein, which is often controlled by a metabolite or sensory mechanism. However, a second regulatory mechanism for gene expression is modulated by RNA control elements (riboswitches), which are *cis*-regulatory elements that can control up to 2% of all genes in some microorganisms (Serganov *et al.*, 2006). This class of regulation is controlled at the level of transcription elongation or translation initiation [for a review see (Mandal & Breaker, 2004)]. Translational regulators predominate in Gram-negative bacteria, whilst transcription regulators are most commonly found in Gram-positive bacteria (Borovok *et al.*, 2006). At first it was thought that only genes involved in coenzyme biosynthesis or transport were regulated by these RNA sequences. However, now seven metabolites (vitamin B₁₂, TPP, SAM, FMN, lysine, guanine and adenine) and recently Mg²⁺ have been shown to regulate numerous genes involved in central metabolism (Dann *et al.*, 2007; Winkler *et al.*, 2002).

The cobalamin riboswitches (B₁₂-box) are composed of an aptamer region for binding vitamin B₁₂ and an expression platform domain. Binding results in changes in either transcription termination or translation initiation (Mandal & Breaker, 2004). The cobalamin riboswitch is found in 5' regions of untranslated mRNA (Nahvi *et al.*, 2004). It is present to regulate genes related to cobalamin

biosynthesis, transport and coenzyme B₁₂-independent enzymes (Nahvi *et al.*, 2002). In most cases it appears that the binding of cobalamin regulates the interaction of a translation enhancer element with a left-stem of a RNA hairpin. This then forms and sequesters the RBS, preventing translation (Santillan & Mackey, 2005). This mode of regulation appears more common in Gram-negative bacteria, such as *S. typhimurium*. In this organism a large 25-gene operon encodes most of the cobalamin biosynthetic enzymes. This was characterised as a translationally controlled operon, whereby the first gene *cbiA* is regulated by translation initiation (Ravnum & Andersson, 1997). In addition RNA hairpins are also located throughout the transcript to control further translation. Alternatively, a more efficient mechanism is transcriptional termination, whereby the binding of cobalamin promotes formation of an intrinsic terminator stem to destabilise transcription elongation. This method of regulation is more predominate for all riboswitches found in Gram-positive bacteria such as *Bacillus*, *Streptomyces* and *Clostridium* species (Borovok *et al.*, 2006; Kazanov *et al.*, 2007; Serganov *et al.*, 2006; Sudarsan *et al.*, 2005). However, no publication has provided experimental data to confirm the activity of a transcriptional cobalamin riboswitch.

It has previously been reported that cobalt has a positive effect on vitamin B₁₂ production in *B. megaterium* (Biedendieck *et al.*, 2010). The cobaltochelataase CbiX^L has an unusual polyhistidine rich-motif and a single [4Fe-4S] cluster housed within a MXCXXC motif (Leech *et al.*, 2003). The functional requirement of these two motifs is unknown. However, the polyhistidine rich-motif allows for the detection of this protein by Western blotting. Many riboswitches are located within the genome of *B. megaterium* DSM319 (Table 3.1). In particular, two are associated with the *cobI* operon. By measuring the response of CbiX^L to cobalt and vitamin B₁₂, this may serve as a starting point for understanding the regulation of this pathway. In addition, with the recent advances molecular biology of *B. megaterium*, a plasmid based strategy to enhance vitamin B₁₂ biosynthesis was attempted. Using the *cobI* operon isolated in pAR8766 (Raux *et al.*, 1998b), this was overexpressed on a plasmid in an attempt to increase vitamin B₁₂ production.

Table 3.1: Cobalamin riboswitches in *B. megaterium* DSM319*

Cobalamin Riboswitch	Genes controlled
1	<i>cobI</i> operon (14 genes)
2	<i>cobI</i> operon (14 genes)
3	<i>cobII</i> operon (9 genes)
4	<i>cbiP</i>
5	BMD_0512 (Putative metal binding protein)
6	BMD_0328 (Putative cobalt transporter)
7	B ₁₂ independent ribonucleotide reductase (<i>nrDEF</i>)
8	B ₁₂ independent methionine synthase (<i>metE</i>)

*Cobalamin riboswitches were found in the *B. megaterium* DSM319 genome by a nucleotide BLAST on <http://megabac.tu-bs.de/> using the *cobI* operon cobalamin riboswitch sequence.

3.2 Results

3.2.1 Cobalt stimulates growth and vitamin B₁₂ production

Previously, it had been reported that cobalt (250 μM) had a positive effect on vitamin B₁₂ production in *B. megaterium* DSM509 when grown in rich-media (Biedendieck *et al.*, 2010; Kunkel, 2004). Cobalt is inserted at an early-stage into sirohydrochlorin by the cobaltochelatase CbiX^L, during the anaerobic biosynthesis of vitamin B₁₂ (Leech *et al.*, 2003). This finding was investigated further in the plasmidless *B. megaterium* DSM319 wild-type strain. Triplicate biological repeats were prepared for each condition and the average final OD_{578nm} and standard deviation are provided as a summary in Table 3.2. Using minimal media, growth studies indicated cobalt is stimulatory at low concentrations (1-10 μM), but becomes inhibitory at higher levels (>100 μM). No growth was observed if cobalt >250 μM was added to cells with a low starting OD_{578nm} < 0.05. To ensure cells were able to grow, a larger culture (500 ml) was first grown to an OD_{578nm} of 0.4 and then subsequently split into separate 50 ml cultures. At this point, cobalt was added at different concentrations. The final OD_{578nm} (24 hrs) was highest with cells grown in 1 μM or 10 μM cobalt with an OD_{578nm} of 2.71 and 2.64, respectively. Cells grown without cobalt had a lower OD_{578nm} of 2.3, whilst cultures with 100 μM and 250 μM of cobalt grew slower during the exponential growth phase (data not shown) and reached an OD_{578nm} of 2.18 and 0.58, respectively.

Vitamin B₁₂ production was also monitored in parallel with growth studies and estimated by a cobyrinic acid bioassay (see Section 2.3.8), which is an indicator of vitamin B₁₂ production (Raux *et al.*, 1996). These results suggested that the rate of vitamin B₁₂ production was highest at non growth inhibitory levels of cobalt (1-10 μM), whilst in the absence of cobalt, vitamin B₁₂ was barely detectable (Table 3.2). Maximum levels of vitamin B₁₂ produced were recorded after 24 hours of growth post cobalt addition. This showed that although high concentrations of cobalt do enhance vitamin B₁₂ production, lower concentrations of 1-10 μM cobalt are optimal and do not inhibit the growth of the cells. This was repeated using rich media and with *B. megaterium* DSM509, with similar results recorded (data not shown). The maximum production of vitamin B₁₂ by the wild-type *B. megaterium*

DSM319 strain under minimal conditions was approximately $7 \mu\text{g L}^{-1}$ when grown in the presence of $10 \mu\text{M}$ cobalt. Higher concentrations of cobalt inhibited growth and decreased vitamin B₁₂ production. Triplicate biological repeats were prepared for each condition and the average and standard deviation for the production of vitamin B₁₂ is provided as a summary in Table 3.2.

Table 3.2: Effect of cobalt on growth and vitamin B₁₂ production

Cobalt (μM)	OD _{578nm}	Vitamin B ₁₂ ($\mu\text{g L}^{-1}$)
0	2.47 ± 0.09	0.14 ± 0.03
1	2.81 ± 0.14	3.35 ± 2.29
10	2.85 ± 0.33	6.93 ± 3.80
100	2.18 ± 0.09	2.66 ± 1.94
250	0.58 ± 0.17	0.13 ± 0.11

3.2.2 CbiX^L is regulated by vitamin B₁₂ feedback

The *cobI* (*cbiW-H₆₀-X-J-C-D-ET-L-F-G-A-cysG^A-cbiY-btuR*) operon encodes for the enzymes that convert uroporphyrinogen III into adenosylcobyrinic acid *a,c*-diamide (Raux *et al.*, 1998b). The intergenic region upstream of *cbiW* contains two cobalamin riboswitches. The first riboswitch is located with a putative RBS, although no start codon is present. The second riboswitch is found with a hairpin loop and six uridine residues, a classical feature of a transcriptional terminator (Christie *et al.*, 1981). Based on the initial findings that cobalt increases vitamin B₁₂ production, the relative levels of CbiX^L were monitored. To do this an easy and reliable method to detect the cobaltochelatase (CbiX^L) by Western blotting was established. CbiX^L contains a polyhistidine-rich motif that allows for the overproduced protein to be purified using a His₆-bind column (Leech *et al.*, 2003). Using the same method CbiX^L could be purified from wild-type *B. megaterium* DSM319, grown under minimal conditions. This purification yielded approximately $0.5\text{-}1 \text{ mg L}^{-1}$ of CbiX^L, which was detected as a single protein at 35 kDa on SDS-PAGE (Figure 3.1) and confirmed by MALDI-TOF. In addition, CbiX^L could be detected by Western blotting using the HisProbe[®] (ThermoFisher) chemiluminescent system. This confirmed that CbiX^L could be detected in the purified fractions and the cell-lysate. Additionally a second protein (~ 20 kDa) was detected in the cell-lysate by Western blotting, but not the purified fractions. Unfortunately, this protein was not identified, but in later expression experiments

the relative levels of this protein did not alter significantly. This unknown protein was renamed as Protein Y, as it was used as a suitable control for the Western blotting experiments.

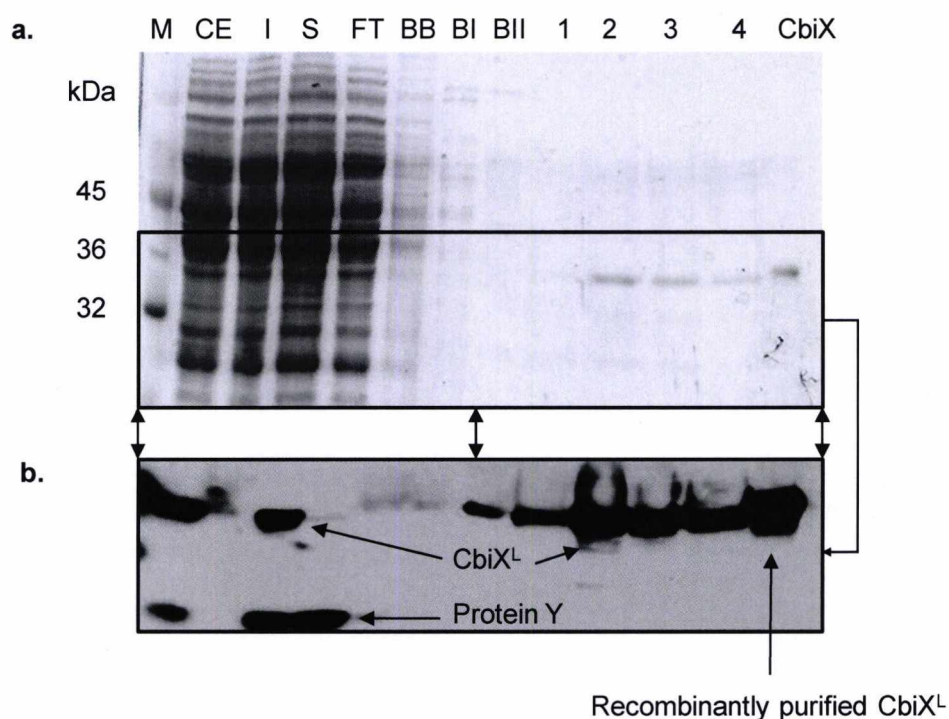


Figure 3.1. Nickel affinity purification and Western blot of *B. megaterium* DSM319 wild-type cells grown under minimal conditions. **A.** SDS-PAGE and colloidal Coomassie blue staining of the purification. M: low molecular weight marker, CE: crude cell extract, I: insoluble protein fraction, S: soluble protein fraction, FT: flow-through, BB: binding buffer, BI: wash-buffer I, BII: wash buffer II, 1-4: elution fractions, CbiX: recombinantly purified CbiX^L is also shown as a positive control, with an extra six amino acids (1-2 kDa larger). **B.** Western blot of a replica of the SDS-PAGE purification. Arrows indicate the proteins in relation to the stage of purification. CbiX^L and Protein Y are detected in the soluble protein fraction as indicated. Only CbiX^L binds to column and is eluted at a high concentration (500 mM) of imidazole. The band for CbiX^L was extracted, digested by trypsin and analysed by MALDI-TOF to confirm its identity.

To begin with, CbiX^L was monitored in response to cobalt and vitamin B₁₂, which were added to early-stage cultures. Cobalt (1-10 μM) decreased the level of CbiX^L in comparison to the condition without cobalt (Figure 3.2). If cobalt levels were then increased up to 100-250 μM, the amount of CbiX^L appeared to increase (Figure 3.2). The level of Protein Y remained constant throughout these conditions. Initially it was uncertain if CbiX^L was under some form of metalloregulation by cobalt, considering its function as a cobaltochelator, or whether it was regulated by

the cobalamin riboswitch in response to the elevated levels of vitamin B₁₂ production observed.

To distinguish between these two possibilities, a range of metals and cobalamins were tested. It was noted that for optimal growth under minimal media conditions, both iron (100 μM) and manganese (1 μM) were required (data not shown). A range of conditions with these metals were tested and relative levels of CbiX^L did not change in response to either metal (data not shown). Subsequently other transition metals were tested including nickel, zinc and copper. Only cobalt seemed to elicit a response (Figure 3.2). To investigate if a cobalamin riboswitch was controlling the level of CbiX^L various amount of vitamin B₁₂ were added. However, it should also be noted that vitamin B₁₂ transport is likely to be regulated. This has been shown with the vitamin B₁₂ transporter, BtuB, which is regulated by a cobalamin riboswitch (Nahvi *et al.*, 2002). Regardless of this, vitamin B₁₂ was added in a range of concentrations to *B. megaterium* DSM319. It was found that the level of CbiX^L was effectively switched off at low concentrations (<10 nM) of exogenous B₁₂. Despite the regulation of vitamin B₁₂ transport being an unknown factor, this provided evidence that CbiX^L was most likely regulated by a cobalamin riboswitch.

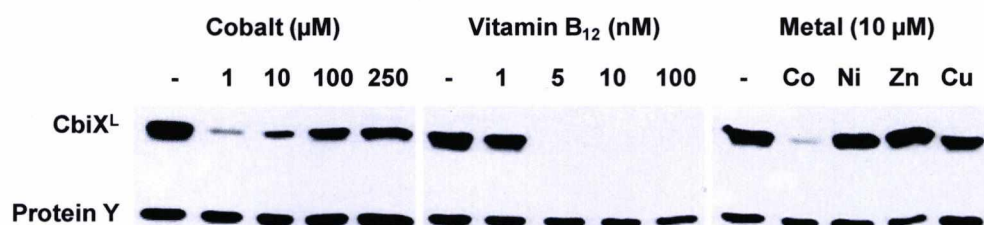


Figure 3.2. Western blots of the CbiX^L protein in response to cobalt, vitamin B₁₂ and transition metals. 50 μg of protein for soluble cell-lysate was loaded and detected with the HisProbe[®] (ThermoFisher). Protein Y is the loading control, as its levels did not alter significantly. Experiments were repeated at least twice to confirm repeatability.

3.2.3 The *cobI* operon is regulated by a cobalamin riboswitch

To determine whether the amount of CbiX^L is regulated at the transcriptional or translational level, expression analysis of the mRNA transcript was required. To address this, the levels of transcript were analysed by semi-quantitative reverse

transcription PCR (sqRT-PCR). Primers were designed to amplify a small region (~150 bp) in the glyceraldehyde-3-phosphate dehydrogenase (*gap*) gene, *cbiX^L* and the last gene in the *cobI* operon, *btuR*. *btuR* was chosen as it would determine whether the entire mRNA transcript for the *cobI* operon was under similar control to *cbiX^L*. RNA extraction was performed during mid-logarithmic growth and stabilised (see Section 2.4.2). This was performed as the half-life of mRNA in bacteria is low, when compared to eukaryotes (Condon, 2007). After extraction of total RNA, relative transcript levels were calculated in comparison to the *gap* gene. The *gap* gene increased slightly in expression with vitamin B₁₂ addition. In contrast, the expression of *cbiX^L* and *btuR* decreased 13- and 4.5-fold, respectively, when vitamin B₁₂ was added (Table 3.3). Additionally, a similar test was performed with the addition of 10 µM exogenous cobalt. The expression of both the *cbiX^L* and *btuR* was downregulated to similar levels observed with the addition of vitamin B₁₂. However the *gap* gene also decreased significantly, therefore the results of this experiment could not be compared to the vitamin B₁₂ test condition. It should also be noted that the overall RNA yield decreased significantly from cultures grown with cobalt. Although it is not understood why this occurred, it is known cobalt can elicit multiple toxicity effects on central metabolism (Thorgersen & Downs, 2007).

Table 3.3. Transcript levels in *B. megaterium* DSM319

RNA ^b	mRNA expression level ^a (Relative to <i>gap</i> mRNA, control condition)	
	Vitamin B ₁₂ (nM)	
	0	100
<i>gap</i>	1.00 ± 0.12	1.12 ± 0.35
<i>cbiX</i>	0.19 ± 0.01	0.01 ± 0.01
<i>btuR</i>	0.11 ± 0.01	0.03 ± 0.01
RNA ^b	Cobalt (µM)	
	0	10
	<i>gap</i>	1.00 ± 0.10
<i>cbiX</i>	0.14 ± 0.02	0.01 ± 0.01
<i>btuR</i>	0.08 ± 0.01	0.02 ± 0.01

^a*B. megaterium* DSM319 grown in M9 minimal medium with either exogenous vitamin B₁₂ (100 nM) or cobalt (10 µM). A control condition without these additives was also grown in parallel. Total RNA was extracted when cells had reached an OD_{578nm} of 1.0. mRNA levels was measured by sqRT-PCR, with expression level relative to the house-keeping gene *gap*. Each condition was conducted as triplicate biological repeat. In addition each RNA sample was

quantified by three reactions during the sqRT-PCR run. ^bThe vitamin B₁₂ and cobalt experiments were performed in separate sqRT-PCR runs. However the RNA from the control condition (no cobalt or vitamin B₁₂) was used in both experiments.

3.2.4 Development of plasmids to overproduce the *cobI* operon

The previous results are consistent with the *cobI* operon being under the control of a cobalamin riboswitch. Removing this regulation could therefore increase vitamin B₁₂ production. Homologous recombination is a common technique in metabolic engineering to produce stable strains without requiring antibiotic selection. However, in *B. megaterium* this technique is challenging, with only a few successful examples (Vary *et al.*, 2007). Instead, a different strategy was employed. Using expression plasmids, the *B. megaterium* DSM509 *cobI* operon (obtained from the plasmid pAR8766) was cloned under the control of the xylose inducible promoter in a derivative of the plasmid pWH1520. This plasmid contained the *cbiX^L-J-C-D-ET-L-F-G-A-cysG^A-Y-btuR* genes (pSJM167). Two variants of this plasmid were created where either the *cbiH₆₀* (pSJM169) or *cbiH₃₀* (pSJM166) genes were placed upstream of the *cbiX^L* gene. In addition, another plasmid (pSJM129), with *cbiH₆₀-cbiX^L-J-C-D-ET-L-F-G-A-cysG^A-Y-btuR*, was also cloned under the control of a constitutive promoter derived from pHBIntE. Due to the cloning procedure, none of the plasmids contained *cbiW*.

After transformation of *B. megaterium* DSM319 with these plasmids, the resultant strains were grown in LB media and analysed for protein production using either SDS-PAGE or 2D-PAGE. Initial analysis of the crude extract revealed a number of the proteins were successfully overproduced. To investigate this in detail, 2D-PAGE was used to analyse the intracellular proteome. Plasmid strains with the xylose inducible promoter were analysed first, as these strains showed the largest bands with SDS-PAGE. After staining, gels were scanned and overproduced proteins were identified by direct visual comparison. Candidate proteins were selected and analysed by mass spectrometry. Of the 25 selected, 24 were successfully identified as cobalamin biosynthetic proteins, derived from the plasmid (Figure 3.3). CbiD was the least overproduced with only a small but repeatable spot detected and confirmed by mass spectrometry. In summary, all the cobalamin biosynthetic proteins were produced from these plasmids. This was repeatable, with little variation in the abundance of the overproduced proteins.

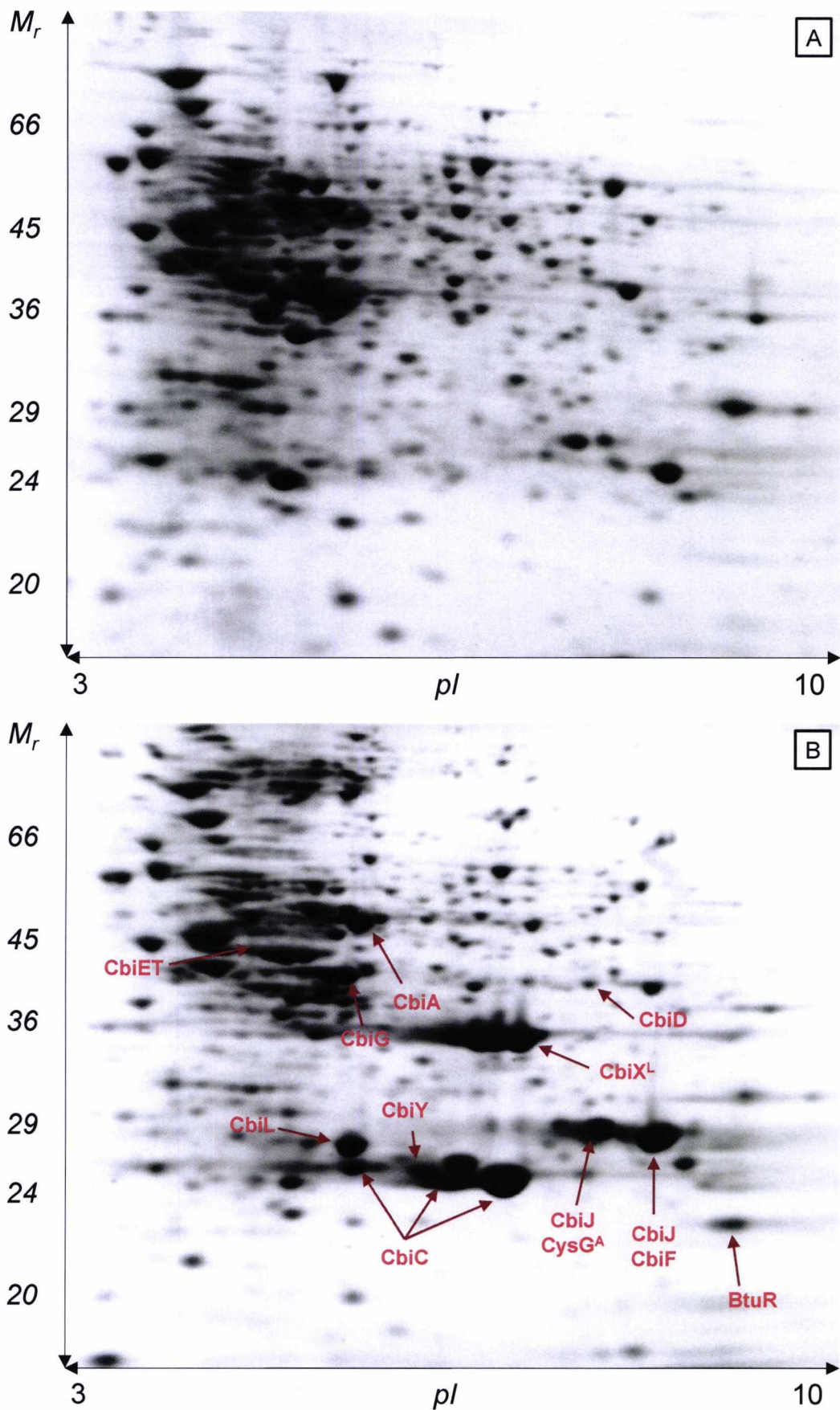


Figure 3.3. 2D-PAGE of recombinant strains of *B. megaterium* DSM319. Continued on next page.

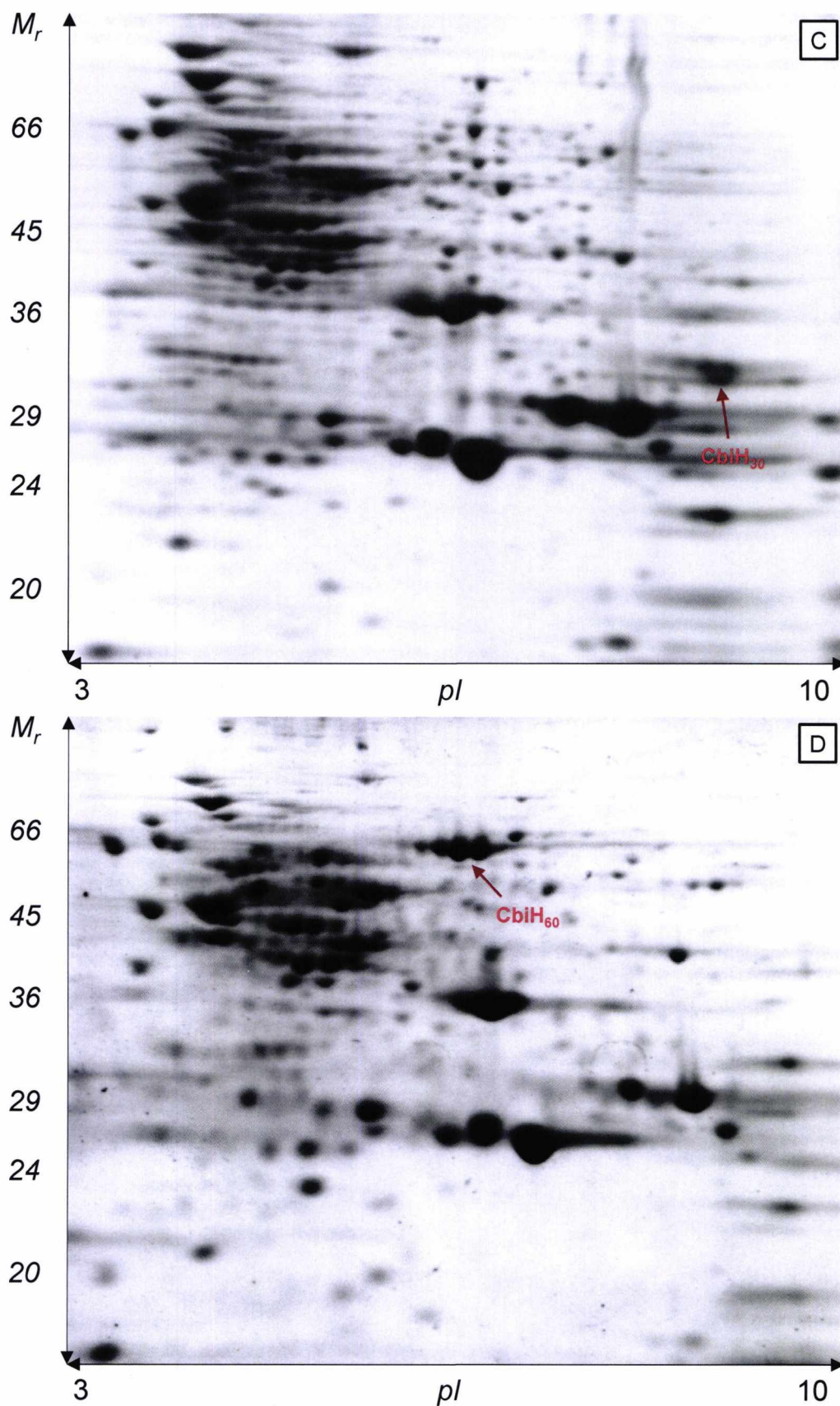
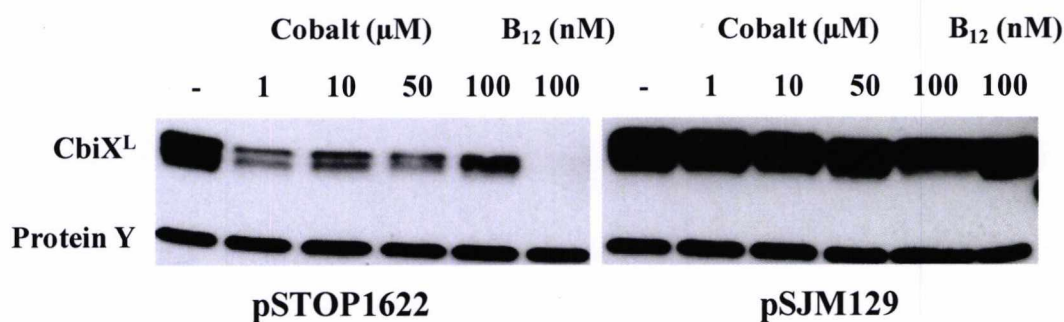


Figure 3.3. 2D-PAGE of recombinant strains of *B. megaterium* DSM319. 2-D IEF/SDS-PAGE gel electrophoretic separation of 100 μ g of cell lysate from *B. megaterium* carrying the pStop1622 (A), pSJM167 (B), pSJM166 (C) or pSJM169 (D) plasmid. Cells were grown in LB broth at 37 $^{\circ}$ C, induced with 0.5% (w/v) D-

xylose and grown for 24 hrs. Protein spots (red arrows) were identified by mass spectrometry and identified using the Mascot search engine.

3.2.5 Vitamin B₁₂ production in recombinant strains

Vitamin B₁₂ production was estimated in these recombinant plasmid strains of *B. megaterium* DSM319 using the *S. typhimurium* bioassay method. All plasmid strains were grown in M9 minimal media (*B. megaterium* modified) conditions. At first different cobalt concentrations were varied as before, with high concentrations of cobalt (>100 μ M) producing low levels of vitamin B₁₂, whilst low levels (1-10 μ M) were found to be stimulatory. Most plasmid strains showed increases in vitamin B₁₂ production, with only *B. megaterium* pSJM167 (no CbiH on plasmid) producing similar levels to the control. Unexpectedly the strongest increase in vitamin B₁₂ levels was observed with *B. megaterium* DSM319 pSJM129 (constitutive expression of the *cobI* operon), which showed a 25-fold increase over the control (Table 3.4). This strain was analysed by 2D-PAGE (see appendix), but it was not clear what proteins were derived from the plasmid. As a more sensitive method, Western blotting was used to detect the level of CbiX^L in this strain. This revealed that CbiX^L was produced only at a slightly elevated level when compared to the control strain (Figure 3.4). Thus the addition of cobalt or vitamin B₁₂ did not affect the level of CbiX^L in the strain containing pSJM129 and it appears likely that the whole *cobI* operon is expressed constitutively and is not regulated by vitamin B₁₂ feedback. This is consistent with the lack of a riboswitch on pSJM129.



B. megaterium DSM319

Figure 3.4. Western blotting measuring the levels of CbiX^L from *B. megaterium* pSJM129. Cells were grown in M9 media, 0.2% glycerol, 100 μM reduced glutathione and 10 μM cobalt. These were incubated at 37 °C, 200 rpm for 24 hrs. 50 μg of protein for soluble cell-lysate was loaded and detected with the HisProbe[®] (ThermoFisher). Protein Y is the loading control, as its levels did not alter significantly.

Table 3.4: Vitamin B₁₂ produced in plasmid strains of *B. megaterium* DSM319

Plasmid	6 hrs induction		24 hrs induction	
	ng ml ⁻¹ OD ⁻¹ ₅₇₈	μg L ⁻¹	ng ml ⁻¹ OD ⁻¹ ₅₇₈	μg L ⁻¹
pStop1622 Empty vector	1	6	1	8
pSJM129 <i>cbiH₆₀-X-J-C-D-ET-L-F-G-A-cysG^A-Y-btuR</i> Constitutive low expression	17	41	33	221
pSJM166 <i>cbiH₃₀-X-J-C-D-ET-L-F-G-A-cysG^A-Y-btuR</i> Inducible high expression	1	2	6	35
pSJM167 <i>cbiX-J-C-D-ET-L-F-G-A-cysG^A-Y-btuR</i> Inducible high expression	1	1	2	14
pSJM169 <i>cbiH₆₀-cbiX-J-C-D-ET-L-F-G-A-cysG^A-Y-btuR</i> Inducible high expression	ND	ND	8	39

3.2.6 Isolation and characterisation of novel zinc tetrapyrrole

Overexpression of the *cobI* operon did not lead to high levels of vitamin B₁₂. However, after a few hours of growth and induction the recombinant strains overproducing the *cobI* operon would appear pink and show red fluorescence when exposed to UV light. Furthermore, the addition of cobalt had no effect on this observation. It was expected that this fluorescence was due to the over-activity of CysG^A which can produce the fluorescent by-product trimethylpyrrocorphin (Warren *et al.*, 1990b). To isolate this unknown intermediate from a crude extract a combination of anion-exchange and RP-18 purification was used. Analysis by UV-Vis spectroscopy, LC-MS and ICP-MS revealed the intermediate contained zinc instead of cobalt and was most likely a derailed product of the vitamin B₁₂ pathway. The UV-Vis spectrum revealed two peaks at 362 (100) and 535 (30) nm, which did not shift upon reduction with sodium dithionite. MS revealed a series of peaks at *m/z* 955, 953 and 951, with an isotope pattern consistent with the presence of zinc, which was later confirmed by ICP-MS. Although the structure of this intermediate was not solved, it is most likely derived from uroporphyrinogen III and possesses four methyl groups with a zinc ion centrally chelated within the macrocycle. This data is summarised in Figure 3.5.

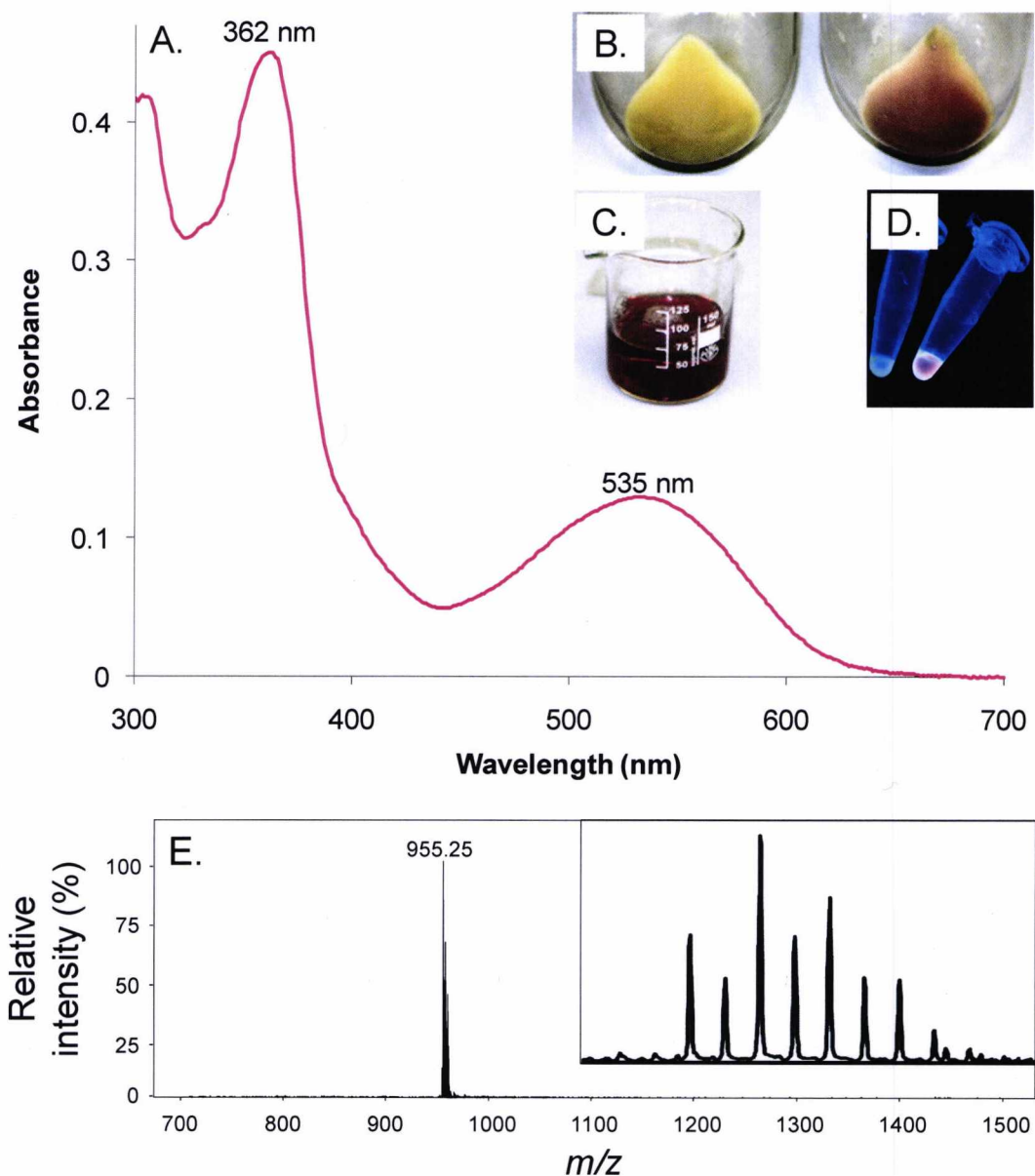


Figure 3.5. Characterisation of the zinc-tetrapyrrole accumulated in *B. megaterium* DSM319 overexpressing the *cobI* operon (pSJM169). A. UV-Vis spectrum of purified intermediate. B. Colour of cells, with a suitable control on the left. C. Picture of the intermediate in the crude extract. D. Fluorescence of cells, with a control on the left. E. MS of purified intermediate with inset diagram of isotope pattern. This pattern suggested the intermediate contained zinc, which was then confirmed by ICP-MS (performed by Professor Nigel Robinson and Dr Kevin Waldron).

3.2.7 Identification of BMD_0328 as a putative cobalt/nickel transporter

The accumulation of the zinc compound was unexpected and could possibly indicate limitations in cobalt transport and chelation. To investigate this problem, this section focuses on attempts to increase cobalt transport into the cell. *B. megaterium* DSM319 contains a two component ABC type cobalt transporter CbiO₁ and CbiO₂. This is believed to transport cobalt in a similar manner to the NikABCDE system (Navarro *et al.*, 1993). As cobalt transport has not been researched in detail, a simpler approach was chosen. In the knowledge that cobalamin riboswitches are often located with genes required for vitamin B₁₂, this was used to locate putative cobalt transporters. This search revealed two new hypothetical proteins that had been annotated as a nickel transporter (BMD_0328) and a nickel binding protein (BMD_0512).

BMD_0328 was chosen as a candidate for cobalt transport as it was predicted to be a membrane protein involved in nickel transport (Figure 3.6). However, it shares closest similarity with the high affinity cobalt transporter NhlF, previously characterised from *Rhodococcus rhodochrous* J1 (Komeda *et al.*, 1997). Here, this transporter is required to deliver cobalt inside the cell for the low molecular-mass nitrile hydratase NhlAB, which uses it as a metallocofactor (Komeda *et al.*, 1997). Similarities also exist with HoxN from *Alcaligenes eutrophus*, a previously classified high affinity nickel transporter (Wolfram *et al.*, 1995). However, no structural or mechanistic information is available regarding the mode of transport. Pairwise sequence alignment revealed a number of conserved residues between the three transporters (Figure 3.7). In particular, three histidine residues are found near the N-terminus, which could potentially be involved in metal binding.

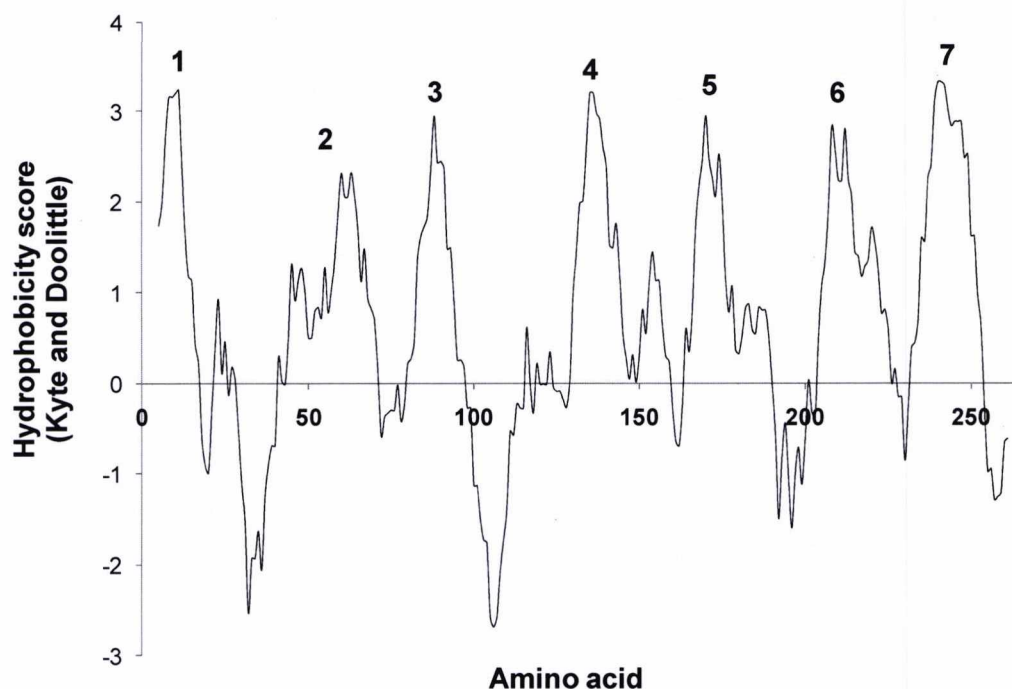


Figure 3.6. Hydrophobicity plot (Kyte and Doolittle) of BMD_0328. Amino acid sequence of BMD_0328 was analysed using <http://expasy.org/tools/protscale.html>. Putative transmembrane domains are indicated as 1-7.

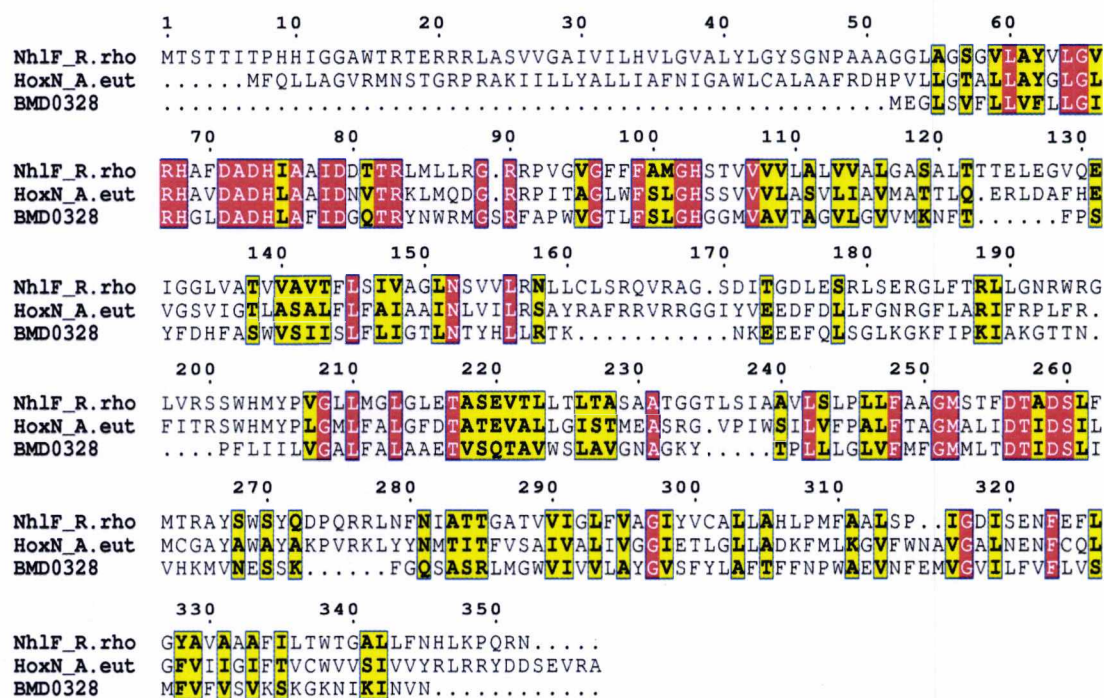


Figure 3.7. Pairwise alignment of BMD_0328, NhlF and HoxN. Sequences for specific cobalt (NhlF) and nickel (HoxN) transporters were aligned with BMD_0328, using ClustalW2 (www.ebi.ac.uk/Tools/msa/clustalw2) and graphically visualised on ESPript 2.2 (<http://esprict.ibcp.fr/ESPript/cgi-bin/ESPript.cgi>). Species abbreviations R.rho (*R. rhodochrous* J1), A.eut (*A. eutrophu*). Sequence identity is shown in red and sequence similarity is yellow.

3.2.8 Investigation into the role of BMD_0328 as a cobalt transporter

The putative nickel transporter BMD_0328 was PCR amplified, cloned downstream of *cbiX^L* in pWH1520-*cbiX^L* and transformed into *B. megaterium* DSM319. Attempts to clone BMD_0328 without *cbiX^L* were made, but these were unsuccessful. After transformation, strains were grown and protein production was induced. Unfortunately, although CbiX^L was overproduced, SDS-PAGE could not detect overproduction of BMD_328 (data not shown). As BMD_0328 is a membrane protein, its solubility and expression may be low. Instead, a simple growth inhibitory test was devised, based on the fact that cobalt is known to be toxic and growth inhibitory. Using a plate growth method, *B. megaterium* DSM319 pWH1520-*cbiX^L*-BMD_0328 was grown up to an OD_{578nm} of 1.0. For a control strain, *B. megaterium* DSM319 pWH1520-*cbiX^L* was also grown in parallel. A 10-fold dilution series was then prepared and aliquots were incubated on a LB plate with and without 0.5% (w/v) D-xylose. The strains were then tested for growth in increasing concentrations of cobalt. This revealed that without xylose induction high concentrations of cobalt are inhibitory to growth, with no differences between the strains (Figure 3.8A). However, with xylose induction, a lower concentration of cobalt (250 μM) was inhibitory only to cells expressing BMD_0328 (Figure 3.8B). This was also repeated in the presence of nickel. Higher concentrations (>1 mM) were required to inhibit growth, with only a small effect observed with BMD_0328 overproduced (data not shown). This evidence suggests BMD_0328 could be involved in cobalt and nickel transport. ICP-MS was also used to analyse cobalt and nickel content in the two strains used. This revealed that if either cobalt or nickel were supplemented to cultures overexpressing BMD_0328, this increase of both metals by approximately 3-fold over the control (Table 3.5). This provided evidence that the transporter could transport both cobalt and nickel.

The next phase of this work was to develop new strains with this cobalt transporter overexpressed in combination with the *cobI* operon. Nonetheless, although double plasmid strains were obtained, time limited any further investigations (see appendix for 2D-PAGE of a double plasmid strain).

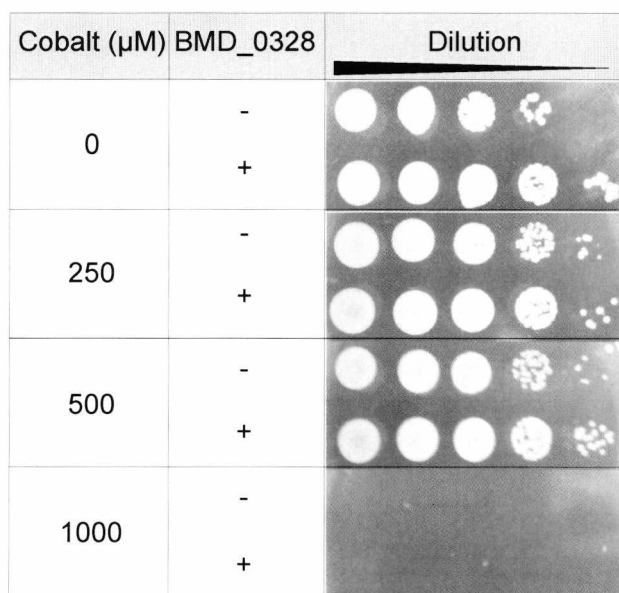
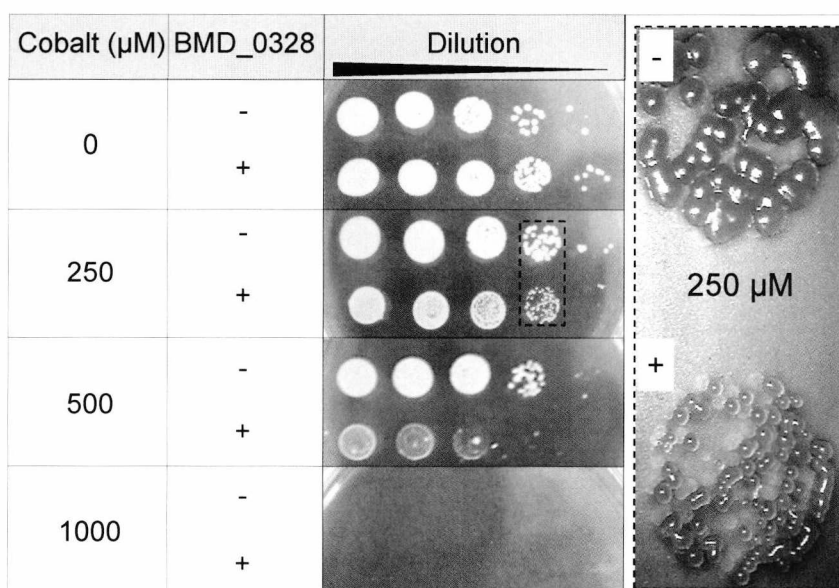
(A) No induction**(B) 0.5% (w/v) D-xylose induction**

Figure 3.8. Effect of BMD_0328 on cobalt minimal toxicity concentration. For comparison *B. megaterium* DSM319 pWH1520-*cbiX* (-) was grown in parallel with *B. megaterium* DSM319 pWH1520-*cbiX*-BMD_0328 (+). Cultures were grown to an $\text{OD}_{578\text{nm}}$ of 1.0 in LB medium, before a 10-fold dilution series was prepared. 20 μl aliquots were spotted onto dry LB plates containing 0.0% (A) or 0.5% (w/v) (B) D-xylose, tetracycline ($10 \mu\text{g ml}^{-1}$) and increasing concentrations of cobalt. These were incubated overnight at 37 °C. A control condition without xylose induction was also prepared. No difference in growth was observed between the two strains. Enlarged inset picture of highlighted dotted box (250 μM cobalt). With BMD_0328 colonies are smaller, indicating inhibited growth.

Table 3.5. Analysis of cobalt and nickel levels in *B. megaterium* DSM319 plasmid strains

A. Cobalt analysis

Plasmid	Metal (50 μ M)	Mass Co / Mass Dry Cell (ng/mg)
pWH1520- <i>cbiX</i>	-	1.1
pWH1520- <i>cbiX</i>	Cobalt	35.9
pWH1520- <i>cbiX</i>	Nickel	3.2
pWH1520- <i>cbiX-BMD0328</i>	-	1.2
pWH1520- <i>cbiX-BMD0328</i>	Cobalt	89.2
pWH1520- <i>cbiX-BMD0328</i>	Nickel	2.5

B. Nickel analysis

Plasmid	Metal (50 μ M)	Mass Ni /Mass Dry Cell (ng/mg)
pWH1520- <i>cbiX</i>	-	0.4
pWH1520- <i>cbiX</i>	Cobalt	0.5
pWH1520- <i>cbiX</i>	Nickel	24.8
pWH1520- <i>cbiX-BMD0328</i>	-	1.0
pWH1520- <i>cbiX-BMD0328</i>	Cobalt	0.5
pWH1520- <i>cbiX-BMD0328</i>	Nickel	105.4

Plasmid strains were grown in M9 media (*B. megaterium* modified) at 37 °C and induced with 0.5% (w/v) D-xylose when an OD_{578nm} of 0.3 was reached. In addition cobalt (50 μ M) and nickel (50 μ M) were added to separate cultures, with a control condition also grown. Cultures were grown overnight. Then 2 ml of cells were centrifuged. The supernatant was removed, before cells were washed with 0.9% (w/v) NaCl. Cells were centrifuged again to remove supernatant, before freeze-drying and weighing of the sample. Samples were analysed by ICP-MS in collaboration with Professor Nigel Robinson from University of Newcastle.

3.3 Discussion

The biosynthesis of vitamin B₁₂ involves a complex metabolic pathway requiring approximately 30 enzymatic steps. It is regulated at several key enzymatic steps, as well as being globally controlled at the transcription-translation level. Moreover, bioavailability of important precursors such as uroporphyrinogen III, SAM, DMB and cobalt can limit the production of vitamin B₁₂. Using *B. megaterium* DSM319, this chapter has focused on cobalt availability, riboswitch regulation and strategies to overcome these limitations to increase the yield of vitamin B₁₂.

In most natural environments only trace quantities of cobalt are available. However, excess cobalt can elicit multiple effects on central metabolism, involving disruption of iron homeostasis, iron-sulphur biosynthesis and sulphur metabolism (Thorgersen & Downs, 2007). In particular, the *Isc* and *Suf* systems that are required for iron-sulphur cluster assembly are upregulated in response to cobalt (Fantino *et al.*, 2010). Previous results showed 250-500 μM cobalt increased vitamin B₁₂ production in *B. megaterium* DSM509 (Biedendieck *et al.*, 2010; Kunkel, 2004; PhD Thesis: Barg, 2003). In *B. megaterium* DSM319 although slight increases were observed at this concentration, growth was inhibited. In addition RT-PCR studies revealed even low concentrations of cobalt (10 μM) would decrease cellular RNA levels. By fine-tuning the balance between the toxic effects of cobalt and its requirement for vitamin B₁₂, an optimum cobalt concentration (1-10 μM) was shown to increase production by 25-fold.

The next major constraint in *B. megaterium* DSM319 is the gene regulatory systems that control global tetrapyrrole biosynthesis. Vitamin B₁₂ biosynthesis is a highly complex process with many of its enzymes encoded within operons. The initial stages leading to the formation of uroporphyrinogen III from glutamyl-tRNA are shared between the haem, sirohaem and vitamin B₁₂ pathways. These steps are encoded within the *hem* operon (*hemAXCDBL*), previously characterised in *B. subtilis*, which is controlled by haem feedback regulation (Levican *et al.*, 2007). For vitamin B₁₂ metabolism there are eight cobalamin riboswitches located within the genome of *B. megaterium* DSM319 (Table 3.1). These are associated with cobalamin biosynthetic operons, coenzyme B₁₂-independent enzymes and putative

cobalt homeostasis genes. In response to elevated vitamin B₁₂ levels these regulatory elements would be predicted to downregulate B₁₂ biosynthetic genes, decrease cobalt import and turn off expression of redundant coenzyme B₁₂-independent enzymes. Initial Western blots showed levels of CbiX^L were decreased when production of vitamin B₁₂ was maximised by availability of cobalt. However, it was thought that CbiX^L could be controlled by metalloregulation as previous studies found CbiX^L in *S. solensis* was upregulated in tandem with the nickel superoxide dismutase (NiSOD), when responding to elevated nickel (Kim *et al.*, 2003). In *B. megaterium* DSM319, nickel had no effect on CbiX^L levels and neither did manganese, iron, zinc or copper. Evidence that this regulation is directly mediated by the riboswitch was obtained by adding vitamin B₁₂ to the culture, which had the effect of decreasing the abundance of CbiX^L. This correlated with the levels of vitamin B₁₂ detected when cobalt was added to cultures. A threshold limit that inhibited further vitamin B₁₂ production was estimated to be between 5-10 nM. Further sqRT-PCR evidence also suggested that the entire *cobI* operon is regulated by a transcriptional cobalamin riboswitch.

To bypass this regulation a number of plasmids containing the *cobI* operon were prepared for overexpression of this operon in *B. megaterium* DSM319. After transformation plasmid strains were tested for their ability to produce cobalamin biosynthetic enzymes and vitamin B₁₂ production was measured. Surprisingly, all genes under the control of xylose inducible promoter were successfully overexpressed and their encoded proteins detected by 2D-PAGE analysis. For vitamin B₁₂ production the plasmid without the ring contraction enzyme (pSJM167) did not show any increase in vitamin B₁₂ levels. However, if CbiH₃₀ or CbiH₆₀ was present, significant increases were observed. Although no inferences for the function of the C-terminal extension of CbiH₆₀ can be made, it seems that the N-terminal domain is functional (CbiH₃₀), despite being truncated. Surprisingly, the largest increase in vitamin B₁₂ production was observed with a plasmid strain that produces the cobalamin enzymes constitutively at a low level. It was unexpected that low levels of expression would achieve the highest levels of vitamin B₁₂. It was assumed that higher levels of enzymes would lead to higher levels of vitamin B₁₂. However overexpression of the *cobI* operon led to approximately 10% (estimated by direct visual comparison) of the entire proteome being dedicated to cobalamin

biosynthesis. This was also detrimental to growth as overproduction of these enzymes demands a high level of the cells resources. A common observation with all the strains discussed was the accumulation of a zinc tetrapyrrole. It is likely that this intermediate is zinc factor S3, previously isolated from strains of *P. shermanii* (Ozaki et al., 1993). This intermediate was shown to be derived from uroporphyrinogen III, with the SUMT enzyme methylating at C-2, C-7 and C-12 forming 2,7,12-trimethylpyrrocorphrin. Then CbiF (C-11 methyltransferase) is shown to incorrectly methylate the C-16 position, followed by non-enzymatic insertion of zinc into the ring macrocycle to form tetramethylated zinc corphinoid factor S3. In the context of this thesis, the accumulation of this intermediate seems to be related to the level of the SUMT enzyme. Only low levels of this intermediate were detected in the plasmid strain that constitutively produces the *cobI* operon at a low level.

Zinc factor S3 is an interesting side-product that seems to form due to the over activity of the SUMT enzyme. In an attempt to reduce the level of this unwanted side-product a putative cobalt transporter was chosen for overexpression to increase cobalt import. As very little is known about the molecular basis for cobalt transport, a simple strategy was chosen to overproduce a single protein aimed to increase cobalt import. The chosen candidate cobalt transporter, the gene product of *BMD_0328*, was briefly characterised using a MTC test and ICP-MS. This provided evidence that *BMD_0328* increases intracellular cobalt and nickel levels. As *B. megaterium* DSM319 has no obvious requirement for nickel, this work proposes the gene product *BMD_0328* can be renamed as a new cobalt transporter CbiV. To develop this idea further double plasmid strains were created. The cobalt transporter gene *cbiV* was transferred into the plasmid pMGBm19 to allow for transformation of this as a separate plasmid into the recombinant strains discussed. Although double plasmid strains of *B. megaterium* DSM319 were obtained, unfortunately time limited a detailed analysis. Follow up experiments shall investigate these strains further with the aim of reducing the level of zinc factor S3 and increasing vitamin B₁₂ production.

This work has explored the regulatory network that controls vitamin B₁₂ biosynthesis in *B. megaterium* DSM319. Cobalt availability is a key limitation in

the pathway, with a fine balance required to limit its adverse effects on central metabolism. However this work has shown by optimising cobalt bioavailability to increase vitamin B₁₂ production (0.2 → 7 μg L⁻¹) and combining this with increasing levels of cobalamin biosynthetic enzymes (7 → 200 μg L⁻¹), a 1000-fold increase in vitamin B₁₂ production can be obtained. This work has explored the molecular biology of *B. megaterium* DSM319, providing new approaches to metabolic engineering. Further developments will focus on fine tuning the balance of cobalamin biosynthetic enzyme levels in the cell, with the aim of achieving even higher levels of vitamin B₁₂.

Chapter 4

Enzyme-capturing of Anaerobic Pathway Intermediates

4.1 Introduction

The anaerobic pathway requires approximately 15 enzymatic steps to convert uroporphyrinogen III to adenosylcobyrinic acid, a precursor of adenosylcobalamin. Recent studies have highlighted a novel method for investigating metabolic pathways (Evelyn Raux, unpublished data; see Section 4.3). This involves using enzymes to capture and retain intermediates from their metabolic pathway. Nevertheless, this relies on a particular enzymatic property that probably has remained undetected in many past studies. This unique property involves the enzyme being able to retain pathway intermediates within its active site. The enzyme can acquire intermediates from its pathway through direct catalysis and retention of the product. For pathways that involve labile intermediates such a property allows for a form of substrate channelling where direct metabolite transfer between consecutive enzymes would permit enhanced flux.

Regarding vitamin B₁₂ biosynthesis, the initial characterisation of CobH provides a fascinating example of enzyme-capturing of pathway intermediates. CobH was originally identified as a hydrogenobyric acid (HBA) binding protein (Blanche *et al.*, 1990). This was obtained by purifying the enzyme from genetically enhanced B₁₂ producing strains of *P. denitrificans*. Upon isolation, CobH was found to be tightly bound to HBA, which does not release its product. This unusual property was overlooked as the enzyme was later characterised as the C-11 → C-12 methyl group rotase (precorrin-8X methylmutase). Later crystallography revealed that CobH tightly binds to HBA, shown in Figure 4.1 (Shipman *et al.*, 2001). This could be a regulatory mechanism to control flux through the rest of the pathway, but nonetheless provides an insight into the findings revealed in this chapter.

Previously, the SirC, CbiX^L and CbiF from *B. megaterium* DSM509 had been heterologously produced in *E. coli* and purified. In particular, both SirC (Schubert *et al.*, 2008) and CbiF (Schubert *et al.*, 1998) have been crystallised, revealing details of NAD⁺ and SAM binding, respectively. However a number of the enzymes from the anaerobic pathway in *B. megaterium* DSM509 were found to be insoluble when produced in *E. coli*, including CbiH₆₀ and CbiW. Attempts to homologously overproduce enzymes in the host *B. megaterium* DSM319 were successfully

demonstrated with CbiX^L, which was obtained in high soluble pure yields (Leech *et al.*, 2003).

As *B. megaterium* DSM319 is a good producer of recombinant proteins, this host was chosen to overproduce many of the anaerobic pathway enzymes from *B. megaterium* DSM509. The main aim of this work was to obtain high soluble yields of protein to allow further characterisation of the anaerobic pathway. As a secondary focus it was suspected that some of these enzymes may be able to capture intermediates from the active vitamin B₁₂ pathway in *B. megaterium* DSM319. This work details the homologous overproduction and purification of CbiL, CbiH₃₀, CbiH₆₀, CbiF, CbiG, CbiD, CbiJ, CbiET, CbiC and CbiW in *B. megaterium* DSM319. Each enzyme was individually characterised and analysed for the presence of bound pathway intermediates.

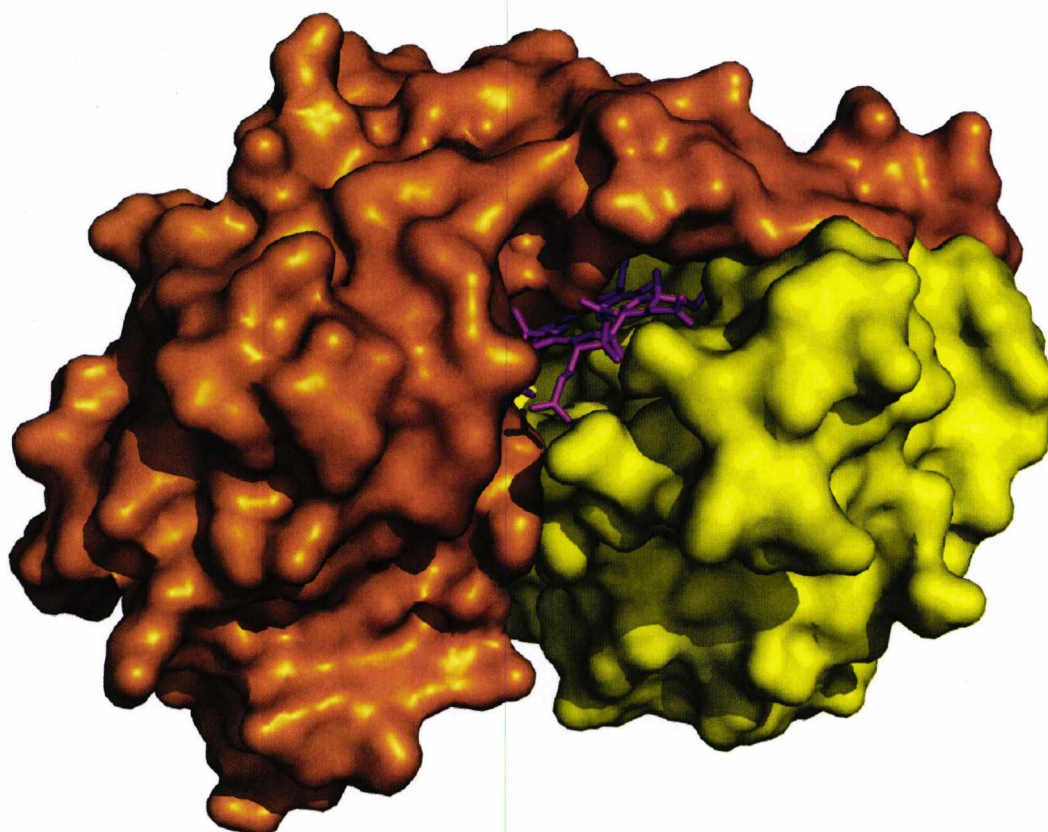


Figure 4.1. HBA bound to CobH (PDB_1I1H) from *P. denitrificans* (Shipman *et al.*, 2001). The enzyme (Chain A orange, Chain B yellow) is displayed with the product HBA (pink, sticks) bound within the active site.

4.2 Results

4.2.1 Cloning of the anaerobic pathway

The *B. megaterium* DSM509 B₁₂ anaerobic genes were all cloned for overexpression in *B. megaterium* DSM319. This included (in order of predicted reaction): *cbiL*, *cbiH₆₀*, *cbiH₃₀*, *cbiF*, *cbiG*, *cbiD*, *cbiJ*, *cbiET* and *cbiC*. Furthermore, *cbiW* was also cloned, which is a putative thioredoxin/ferredoxin of unknown function with similarity to a range of NAD(P)⁺/NAD(P)H oxidoreductases, including ferredoxin hydrogenases. Each gene was cloned with either a C-terminal His₆-tag or N-terminal His₆-tag. The choice of tag position was flexible as solubility and expression levels varied. In general, the N-terminal His₆-tag was produced at very high levels, but the majority of recombinant protein was found as insoluble inclusion bodies. Yet, the C-terminal His₆-tag enzymes were produced at a lower level, but higher yields of purified protein were obtained and thus the C-terminal His₆-tag became the preferred choice. All gene fusions were cloned in frame, and for the C-terminal His₆-tag fusions the natural stop codon of the gene of interest was removed in the primer. After PCR and restriction digest, the gene was ligated into either pN-His1622 or pC-His1622 (Table 4.1). After successful ligation and sequencing, plasmid DNA was prepared and *B. megaterium* DSM319 was transformed.

Table 4.1. Cloning of anaerobic pathway genes into plasmids

Gene	Plasmid	Restriction site
<i>cbiL</i>	pN-His-TEV1622	<i>Bgl</i> II, <i>Sph</i> I
	pC-His1622	<i>Bgl</i> II, <i>Sph</i> I
<i>cbiH₃₀</i>	pC-His1622	<i>Bgl</i> II, <i>Eag</i> I
<i>cbiH₆₀</i>	pN-His-TEV1622	<i>Bgl</i> II, <i>Eag</i> I
	pC-His1622	<i>Bgl</i> II, <i>Eag</i> I
<i>cbiF</i>	pC-His1622	<i>Bsr</i> GI, <i>Bam</i> HI
<i>cbiG</i>	pC-His1622	<i>Bgl</i> II, <i>Sph</i> I
<i>cbiD</i>	pC-His1622	<i>Bgl</i> II, <i>Sph</i> I
<i>cbiJ</i>	pC-His1622	<i>Bgl</i> II, <i>Sph</i> I
<i>cbiET</i>	pC-His1622	<i>Bgl</i> II, <i>Sph</i> I
<i>cbiC</i>	pN-His-TEV1622	<i>Bgl</i> II, <i>Sph</i> I
<i>cbiW</i>	pN-His-TEV1622	<i>Bgl</i> II, <i>Sph</i> I
	pC-His1622	<i>Bgl</i> II, <i>Sph</i> I

4.2.2 Purification of the B₁₂ anaerobic enzymes

After transformation of *B. megaterium* DSM319, glycerol cultures were prepared for long term storage of recombinant strains. For standard production of recombinant proteins, cultures were grown in 1 litre of LB medium. Initially, cultures were grown at 37°C, but the solubilities of CbiG and CbiH₆₀ were low. However, it was found that the solubility could be substantially improved by growing and inducing at 30 °C. This also increased the final cell biomass, which improved yield. The temperature of 30 °C was routinely used for recombinant protein production in *B. megaterium* DSM319 with an induction period of approximately 16-20 hrs. In addition, a lower xylose induction concentration of 0.25% (w/v) was used to improve production and solubility.

After harvesting cultures by centrifugation, the cells were re-suspended in 15 ml of binding buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 mM imidazole) and kept on ice. Following sonication on ice and centrifugation at 4 °C, the soluble protein fraction was passed through a His₆-bind column and purified by washing with increasing concentrations of imidazole (10-500 mM) in the same buffer. Fractions were collected for total cell extract, soluble protein, insoluble protein and all wash-steps of the purification procedure. Proteins purified from cultures were analysed by SDS-PAGE to determine relative levels of production, solubility, relative molecular mass and the purity of isolated His₆-tagged proteins.

Proteins migrated with the correct theoretical molecular mass and with very little proteolysis observed, with the exception of CbiH₆₀. Most proteins were stable after elution but sometimes precipitation was observed to occur. A summary of the protein purifications is shown in Table 4.2 and a profile of the purification procedure is shown in Figure 4.2. In general, all but two of the proteins could be purified reproducibly, with an approximate yield of between 1-50 mg L⁻¹. CbiG was overproduced the most, yielding up to 50 mg L⁻¹, whilst also remaining stable in a range of buffers. Some enzymes were obtained in much lower amounts, with CbiL and CbiW proving the most difficult. CbiL could only be purified in small amounts, and not always reproducibly (< 1 mg L⁻¹). The putative thioredoxin CbiW proved the most problematic, as it remained completely insoluble. The inclusion bodies were dark-brown to orange in colour, indicating the possible presence of a redox

centre. This protein could be purified by resolubilisation in 8 M urea, however, even with refolding no suitable method could reconstitute the anticipated [2Fe-2S] redox centre. A similar problem had previously occurred when the same protein was produced in *E. coli* (Dr Andrew Lawrence, unpublished data). CbiC could be reproducibly purified, but only at about 1 mg L⁻¹. Although highly produced, the majority of the protein was found as inclusion bodies. The rest of the proteins were obtained at reasonable levels and are discussed later in Chapter 6.

For each enzyme purified a wide range of buffers were tested to optimise stability. This was performed as not every enzyme was stable in buffers similar to the standard purification buffer used. For example CbiH₆₀, CbiF and CbiD were not soluble in the standard Buffer A (20 mM Tris-HCl pH 8.0, 100 mM). Instead CbiH₆₀ and CbiF were found to require a high concentration of NaCl to maintain solubility at pH 8.0. CbiD was stable in a wide range of pH buffers, but its solubility was optimised in the presence of 10% glycerol. These findings are summarised in Table 4.2. In addition, a number of these enzymes (CbiH₆₀, CbiF, CbiG, CbiD and CbiJ) were subjected to crystallisation trials. Using the sparse matrix screen from Molecular Dimensions (MD-01), these proteins were incubated with a range of conditions. Unfortunately none of the enzymes showed any potential to crystallise. This may indicate that substrate and/or cofactors are required to reach a crystallisable form of the enzyme.

Table 4.2. Purification summary of the *B. megaterium* DSM509 cobalamin enzymes produced and purified from *B. megaterium* DSM319

Enzyme	Molecular mass with His ₆ -tag (kDa)	Soluble protein (mg/L)	Stability and buffer choice
CbiX ^L	36.0 (Natural His-tag)	20	Soluble in Buffer A at 10 mg/ml. Slight proteolysis. (Leech <i>et al.</i> , 2003).
CbiL	31.7	< 1	Irregularly purifies up to 1 mg L ⁻¹ , but most of the time no protein could be purified.
CbiH ₆₀	64.7	10	Best in 50 mM MOPS pH 7.0, 500 mM NaCl, 100 mM (NH ₄) ₂ SO ₄ at 5 mg/ml but can still precipitate. High levels of proteolysis. Stable in Buffer H in presence of cobalt-factor III.
CbiH ₃₀	31.6	5	Unstable but no proteolysis. Soluble in Buffer H at 10 mg/ml, but would gradually precipitate. Soluble for long periods of time in the presence of cobalt-factor III.
CbiF	32.4	30	Soluble in Buffer A at 10 mg/ml. No proteolysis.
CbiG	45.4	50-60	Soluble in Buffer A at 15 mg/ml. No proteolysis.
CbiD	43.3	5	Soluble in a range of pH buffers (5.5-10), but required 10% glycerol to remain soluble at 5 mg/ml. Best kept at 4 °C. No proteolysis.
CbiJ	31.6	30	Soluble in 20 mM tri-sodium citrate pH 5.5, 50 mM NaCl at 10 mg/ml. No proteolysis.
CbiET	48.3	30-40	Soluble in Buffer H at 10 mg/ml, further characterisation required.
CbiC	29.3	1	Soluble in Buffer A at 0.5 mg/ml. No proteolysis.
CbiW	18.7	10	Denaturing purification from inclusion bodies using 8M urea. After refolding on the column protein stable in Buffer A at 10 mg/ml. No proteolysis.

Buffer A – 20 mM Tris-HCl, 100 mM NaCl, pH 8.0

Buffer H – 50 mM Tris-HCl, 400 mM NaCl, pH 8.0

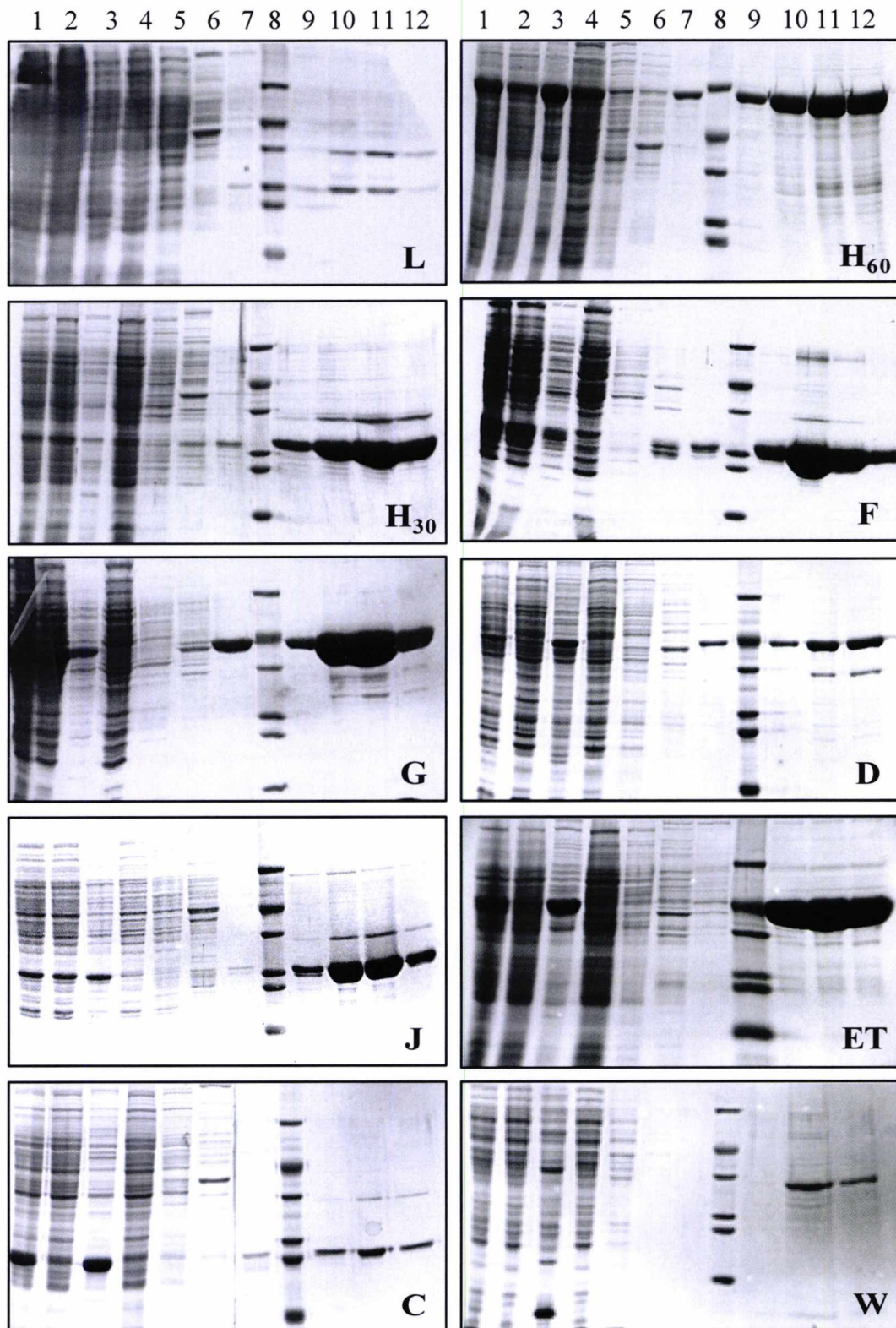


Figure 4.2. SDS-PAGE of purification process of *B. megaterium* DSM509 cobalamin enzymes. Displayed in order of predicted reaction (except CbiW), protein names are abbreviated as a single letter, indicating the Cbi name. Purification as follows (lanes from left-right): (1) crude-lysate, (2) soluble supernatant, (3) insoluble pellet, (4) supernatant flow-through, (5) binding buffer flow-through, (6) wash buffer I flow-through, (7) wash buffer II flow-through, (8) Dalton VII marker and (9-12) final lanes contain elution fractions. From top-bottom the masses (kDa) indicated on the Dalton VII are as follows: 66, 45, 35, 29, 24 and

20 (for further Dalton VII information please see Materials and Methods). Additional notes: The CbiH₆₀ purification is missing the 20 kDa marker and in some of the purifications a single band is observed at 35 kDa indicating the presence of CbiX^L.

4.2.3 Enzyme bound pathway intermediates

CbiL was purified from a 4 litre culture as the protein was only produced in limited quantities. The purification of the protein resulted in an unusual observation. As the protein was purified a bright green band was found to accumulate at the top of the His₆-bind column, and as the purification continued this colour was washed off. This observation was also common to many of the proteins from the anaerobic pathway, with the enzymes found to co-purify with intermediates from the vitamin B₁₂ pathway (Figure 4.3).

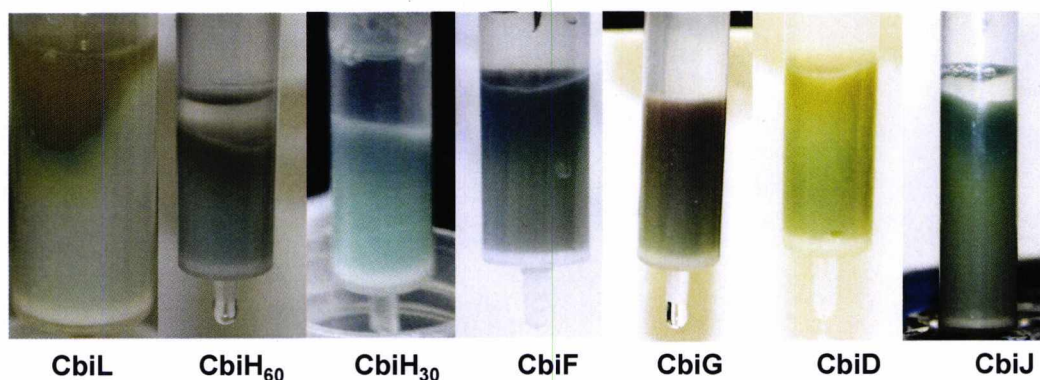


Figure 4.3. Intermediates bound to enzymes from the anaerobic pathway. Summary of intermediates using UV-Vis, MS and picture of purified protein in Table 4.3. Note for CbiH₆₀, the brown colouration at the top of the His₆-bind column is due to the presence of a redox group, presented in chapter 5.

4.2.4 Ring contraction as a bottleneck in the anaerobic pathway

Initially it was established that the proteins would only purify with intermediates from the anaerobic pathway if the cells were grown in the presence of cobalt (10 μ M). Most of the enzymes could be purified with intermediates from the vitamin B₁₂ pathway, with the exception of CbiG and CbiX^L. Instead, CbiG and CbiX^L were found to bind porphyrin intermediates (presumably from the haem pathway), as the UV-Vis spectra of the purified proteins revealed a prominent Soret band at 410 nm, with both α - and β - bands. Unfortunately these intermediates could not be extracted from the enzymes after precipitation. By LC-MS analysis, CbiH₃₀, CbiJ

and CbiF were found to be bound to a mixture of cobalt-factor II (~50%) and cobalt-factor III (~50%). These were detected as di-imidazole complexes, suggesting imidazole acted as both an upper and lower axial ligand. However, no intermediate was found to bind to CbiH₆₀, although during its purification a dark brown-green band would appear at the top of the metal chelate column but slowly wash off during the purification of the protein as the imidazole concentration increased. When this compound was collected, LC-MS analysis revealed that this contained di-imidazole cobalt-factor II. This suggested that CbiH₆₀ could weakly interact with this intermediate, but binding seemed to be displaced by imidazole.

4.2.5 CbiD purifies with cobyrinic acid

During the purification of CbiD a yellow band accumulated at the top of the His₆-bind column. The protein eluted after several washes as a yellow solution with a broad absorbance in the UV-Visible spectrum around 300-600 nm (Figure 4.4). On reducing CbiD by the addition of excess sodium dithionite, a darkening of the colour to a bronze solution was observed. UV-Vis spectral changes were recorded but were very broad. Upon desalting to remove imidazole, CbiD revealed an absorbance peak at 308 nm. Precipitation of the protein by heating resulted in the solution turning pink with absorbance peaks at 355 (100), 408 (20), 507 (30) and 535 (30) nm, shown in Figure 4.4. The UV-Vis spectrum showed characteristic features of later intermediates in the vitamin B₁₂ pathway (Roessner *et al.*, 2005). This colour change could be reversed by the addition of a strong reductant such as sodium dithionite or sodium borohydride, which produced a pale yellow-bronze solution. Exposure to air and rapid shaking reverted this colour change. However, if the enzyme was purified anaerobically these visual and spectral changes were not observed, which provided evidence that the colour change was due to air oxidation.

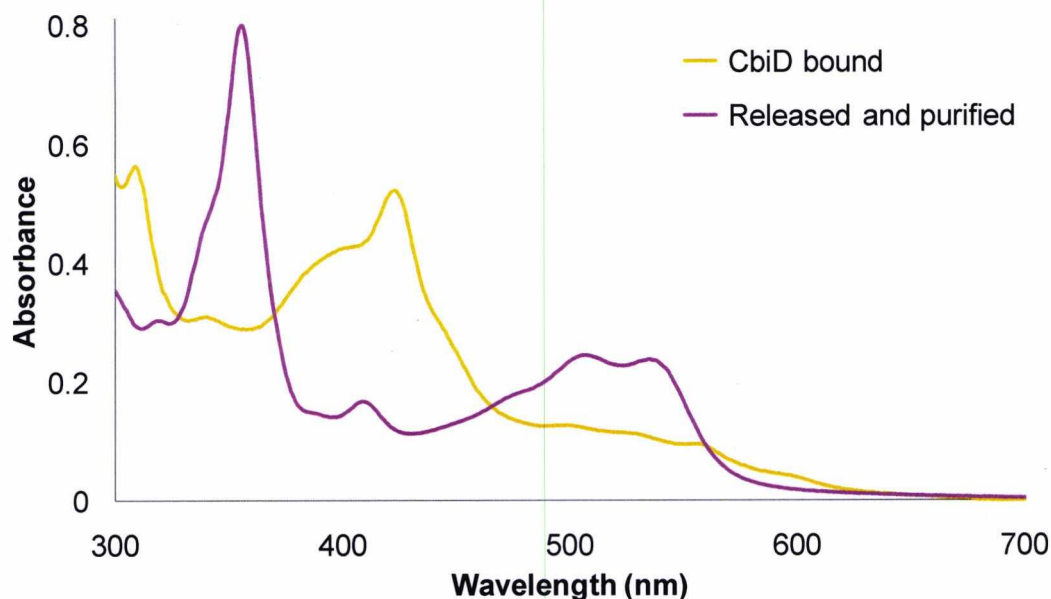


Figure 4.4. UV-Vis spectrum of intermediate isolated from CbiD. The enzyme bound intermediate (orange line), then after purification and air exposure (pink line).

This intermediate was identified as cobyrinic acid, through a combination of bioassay and LC-MS. For bioassay testing, the *S. typhimurium* double mutants AR3612 (*metE*, *cysG*) and AR2680 (*metE*, *cbiB*), which are dependent on cobyrinic acid (and cysteine) and cobinamide (and cysteine) for growth, respectively, were used in a plate assay. The purified intermediate isolated from CbiD was added to these strains and incubated overnight with a cobalamin standard as a positive control and ddH₂O as a negative control (Figure 4.5). As expected, the vitamin B₁₂ standard supported growth of both *S. typhimurium* mutants, whilst the CbiD sample could also support growth on both plates. Interestingly, the growth on the cobyrinic acid plate required a longer period of incubation, as the zone of growth increased for several days, whilst the vitamin B₁₂ positive control remained the same. The growth zone for the cobinamide bioassay remained the same for the CbiD sample even after prolonged incubation. This suggested that the majority of this intermediate contained cobyrinic acid, whilst traces of cobinamide or a later intermediate were also present.



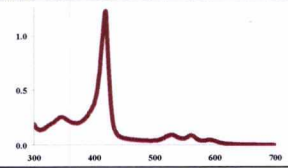


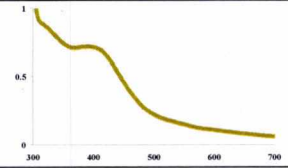

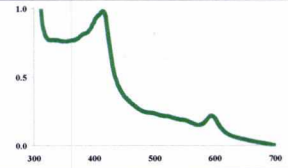

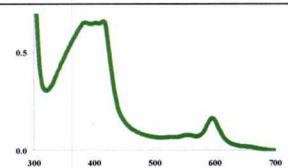

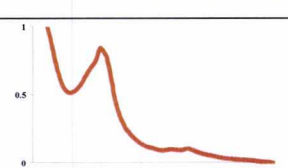

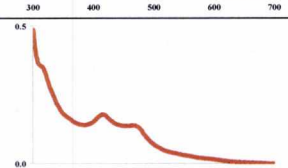

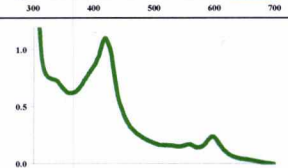

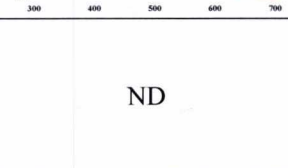
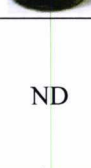
Cobyrinic acid

Cobinamide

Figure 4.5. Bioassay for the vitamin B₁₂ intermediates cobyrinic acid and cobinamide. The *S. typhimurium* mutant strains AR3612 ($\Delta cysG$, $\Delta metE$) and AR2680 ($\Delta cbiB$, $\Delta metE$) were used to test for cobyrinic acid and cobinamide respectively. The indicated intermediate can be transported and converted into active vitamin B₁₂ which is used for the vitamin B₁₂ dependent methionine synthase (MetH) to support growth.

The intermediate could be purified after separation from CbiD by reverse phase chromatography (C-18) using water as the mobile phase and eluted with ethanol. After freeze drying and re-suspension in water, the intermediate was run on LC-MS using 1% acetic acid and acetonitrile solvents, with a linear gradient (30 min runtime). By measuring the absorbance at 350 nm the intermediate eluted between 8-12 min as a series of peaks that seemed to be separated by a single mass unit. The diode array UV-Vis spectrum was identical with the observed spectrum before HPLC. On the mass spectrometer the intermediate was detected as three main species. Masses at m/z 932.40, 933.41 and 934.41 were obtained, along with their more dominant double charged ion. A slight separation of the three species could be observed with the gradient, with the increased masses eluting fractionally later. The mixture of masses was suggestive of the intermediate cobyrinic acid, along with some derivatives presumably lacking some of the amidations. Interestingly, no masses for cobinamide or other later intermediates were detected. In fact it was expected that cobyrinic acid would be adenosylated, as it had previously been shown that the CbiP from *S. enterica* was only active with adenosycobyrinic acid *a,c*-diamide (Fresquet *et al.*, 2007). Even repeating the extraction under low light exposure and anaerobic conditions did not result in the detection of adenosylcobyrinic acid. This raises an interesting question about the timing of the cobalt reduction and adenosylation process in the anaerobic pathway.

Table 4.3. Summary of enzyme-intermediate trapping experiments

Enzyme	UV-Vis spectra	Visual	Mass (Da)	Bound intermediate
CbiX ^L			ND	Unknown
CbiL	ND		ND	Cobalt-factor III?
CbiH ₆₀			917	No intermediate bound. Cobalt-factor II washes off during purification
CbiH ₃₀			931 917	Mixture of cobalt-factor II and cobalt-factor III
CbiF			931 917	Mixture of cobalt-factor II and cobalt-factor III
CbiG			ND	Unknown
CbiD			932 933 934	Mixture of cobyrinic acid derivatives
CbiJ			917 931	Mixture of cobalt-factor II and cobalt-factor III
CbiET	ND	ND	ND	ND
CbiC			ND	No intermediate bound

4.2.6 Cobalamin reduces the levels of protein-bound B₁₂ intermediates

If the work from Chapter 3 is considered, it was anticipated that the end-product of the pathway, adenosylcobalamin, should decrease levels of the B₁₂ biosynthetic enzymes and thus reduce flux through the pathway. Logically, therefore, the presence of exogenous B₁₂ in the medium may be expected to reduce the level of bound intermediates to the pathway enzymes. This was tested with the enzymes CbiF, CbiD and CbiG. It is known that CbiF and CbiD copurify with B₁₂ intermediates, whereas the nature of the intermediate bound to CbiG had not been fully characterised. By growing recombinant strains in parallel with and without cyanocobalamin (100 nM) supplemented, the proteins were purified as before and compared by measuring the UV-Vis spectra with protein concentration for each enzyme standardised.

For CbiF and CbiD, the addition of cyanocobalamin had similar effects. The UV-Vis spectra derived from cultures grown with cyanocobalamin revealed that the level of B₁₂ intermediates bound to CbiF and CbiD had indeed decreased when compared to the control condition (Figure 4.6). For CbiD this was observed with decreased absorbance at 317 and 480 nm. With CbiF, the typical absorbance at 595 nm (indicating a factor intermediate) decreased, with only a single broad absorbance peak found at 380 nm. The nature of this new intermediate was not identified. It should also be mentioned that the majority of purified B₁₂ enzymes from *B. megaterium* show absorbance close to 400 nm. This is particularly true with CbiG. Although addition of cyanocobalamin to CbiG cultures leads to a slight reduction in absorbance at 400 nm, this is most probably due to variation. There are no overall changes in the spectra as observed with CbiD and CbiF, indicating that the CbiG bound intermediate is unlikely to be associated with vitamin B₁₂ biosynthesis.

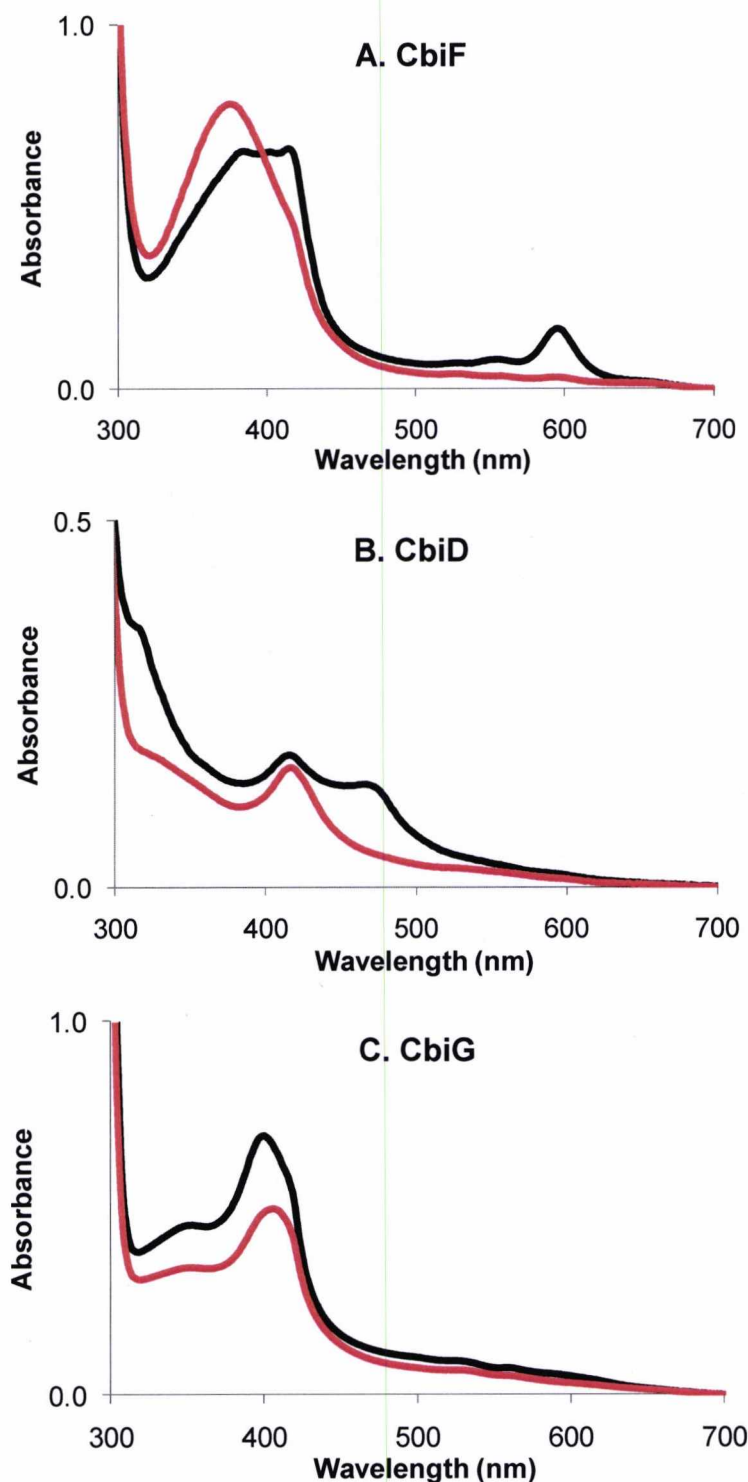


Figure 4.6. The effect of cyanocobalamin on protein-bound B_{12} intermediates. Cultures were grown as before with 100 nM cyanocobalamin (red line) and without (black line). Proteins were purified in parallel under anaerobic conditions and standardised to the same concentration when the UV-Vis spectra was measured. **A.** CbiF. With cyanocobalamin, the protein is purified as a yellow protein with no cobalt-factor II/III present. **B.** CbiD. With cyanocobalamin added the protein purifies colourless with no cobyrinic acid bound. **C.** CbiG. The addition of cyanocobalamin had no effect on the presence of the porphyrin like intermediate bound to the enzyme.

4.3 Discussion

Previous studies highlighted that many of the *B. megaterium* DSM509 B₁₂ biosynthetic enzymes were insoluble when heterologously produced within *E. coli*. This was overcome by overproducing these enzymes homologously within the host *B. megaterium* DSM319. In fact high yields of soluble protein were obtained up to 50 mg L⁻¹. The successful use of *B. megaterium* for recombinant protein production of a wide variety of B₁₂ biosynthetic enzymes provides further evidence of the potential of *B. megaterium* in modern biotechnology. Not only are the proteins produced in a soluble form but they are obtained with very little proteolysis and in high purity. These findings support previous publications that endorse *B. megaterium* as a useful tool in biotechnology (Biedendieck *et al.*, 2007a; Biedendieck *et al.*, 2007b; Biedendieck *et al.*, 2010; Gamer *et al.*, 2009; Vary *et al.*, 2007).

As an interesting side-study many of the enzymes produced and purified from *B. megaterium* DSM319 were found to be bound to pathway intermediates picked up from the cellular metabolic pool. In summary, a few of these enzymes would co-purify with cobalt-factor II/III bound, indicating a lack of specificity with these enzymes. A couple of statements can be made from these simple observations. Firstly, it seems likely that the ring contraction step is a “bottleneck” in the pathway, leading to the accumulation of these early intermediates. Secondly, as these were non-specifically bound to other enzymes, this could be interpreted as a possible regulatory mechanism. This is entirely plausible as only proteolysis could release the intermediates from the enzymes, suggesting that these intermediates were tightly bound and thus could act as inhibitors. An unusual finding was that CbiD purifies bound to cobyrinic acid. This was confusing since CbiD, as will be described later in Chapter 6, is the missing C-1 methyltransferase. Considering, some of the other enzymes were found bound to cobalt-factor II/III, this might provide further evidence for self-regulation in the pathway. Although cobyrinic acid was not found to be as tightly bound, it may also be linked with the finding that CbiD from *S. enterica* seemingly required the presence of CbiA and CbiP for activity (Roessner *et al.*, 2005).

Over 20 years ago the initial characterisation of the CobH protein identified it as a HBA binding protein (Blanche *et al.*, 1990). Through later experiments, CobH was established as the precorrin-8 mutarotase (HBA synthase) (Thibaut *et al.*, 1992). The significance of HBA binding was overlooked for some time. Recently, research has identified that many of the enzymes from the aerobic vitamin B₁₂ pathway can bind tightly to their products (Raux and Warren *et al.*, unpublished data). These findings have shown the aerobic pathway can be reconstructed in *E. coli* to synthesise vitamin B₁₂ *de novo*. To do this, the genes necessary to produce HBA from uroporphyrinogen III (*cobA-I-G-J-M-F-K-L-H*) were cloned into a plasmid and expressed in *E. coli*. This produced high yields of HBA. If CobH was His₆-tagged, this could be purified to homogeneity bound to the product HBA. Indeed, by placing a His₆-tag (*) on the last protein in a reaction sequence, the His₆-tagged protein could be used to capture the last intermediate and allow it to be purified. For example with the proteins CobA-I-G-J-M-F-K* overproduced in *E. coli*, the His₆-tagged protein CobK* could be purified bound to precorrin-6B, the expected intermediate to be produced by these set of proteins. This was repeated for other plasmid constructs, with many of the pathway enzymes found to co-purify bound to the end-product of the constructed pathway. This suggested that the enzymes would bind tightly to their products and would only release upon binding of the next enzyme in the pathway.

Although this chapter did not follow the same approach, it does reveal how an overproduced enzyme in a background of a functional pathway can be used to “fish” out metabolic intermediates. Such a method can provide valuable clues to the overall understanding of the pathway. It is anticipated that the findings in this chapter can be further advanced similarly to the work demonstrated with the aerobic pathway. It may be possible to apply these approaches to other areas such as investigating novel biosynthetic pathways, providing a relatively easy extraction process for valuable secondary metabolites.

Chapter 5

The Characterisation of the Anaerobic Ring Contraction Enzyme CbiH₆₀

5.1 Introduction

Uroporphyrinogen III is the first cyclic intermediate in the modified tetrapyrrole biosynthetic pathway and is composed of four pyrrole groups linked together in an asymmetric form. This provides the molecular frame from which all tetrapyrroles are derived. Central to the chemistry of vitamin B₁₂ is the cobalt ion, which is tightly coordinated within the macrocycle in an octahedral arrangement. However, what distinguishes vitamin B₁₂ from the other modified tetrapyrrole macrocycles is the contracted nature of the molecule. This smaller macrocycle is known as the corrin ring, which is formed from the removal of the methylene bridge between rings A and D, during a process called ring contraction. This allows for a tighter coordination of the cobalt ion, making it very difficult to remove the metal centre (Battersby, 2000). Moreover, another distinguishing feature of vitamin B₁₂ is the presence of the upper and lower ligands, which also bind to the cobalt ion and assist in mediating some of the chemistry associated with the coenzyme.

The mechanism of ring contraction is a key difference between the aerobic and anaerobic pathways of cobalamin biosynthesis. This is reflected at the genetic level, as the aerobic pathway requires two enzymes, CobG (monooxygenase) and CobJ (class III methyltransferase), whilst the anaerobic route only requires a single class III methyltransferase, CbiH (Debussche *et al.*, 1993; Santander *et al.*, 1997). The mechanistic steps of the ring contraction reaction in the aerobic pathway have been recently characterised in detail (Schroeder *et al.*, 2009). CobG is a monooxygenase that contains a [4Fe-4S] cluster and a non-heme iron. The region containing the [4Fe-4S] cluster shares similarity to the NiR/SiR family. CobG utilises molecular oxygen, which is proposed to bind to the non-heme iron. An electron is then provided by the reduced [4Fe-4S] cluster to generate firstly a peroxy intermediate, prior to the formation of a Fe³⁺-O⁺ species. This intermediate is then used to hydroxylate at position C-20 on precorrin-3A, whilst the acetate side chain at C-2 forms a γ -lactone linked to C-1, thus producing precorrin-3B. Subsequently, CobJ methylates at C-17 and ring contraction occurs, forming precorrin-4.

The anaerobic pathway requires only the CbiH enzyme to catalyse ring contraction. CbiH is 40-60% identical to CobJ from the aerobic pathway, implying the

mechanism could be very similar. This single typical class III methyltransferase can methylate the C-17 position, form a δ -lactone ring and mediate ring contraction, producing cobalt-precorrin-4 from cobalt-precorrin-3. The order of this reaction is not known and seems to require only SAM as a methyl donor. In addition cobalt-factor III can be used as a substrate and is converted to cobalt-factor IV. As the anaerobic pathway does not require oxygen for ring contraction, the key difference is the presence of the cobalt ion. It has previously been speculated that the cobalt ion could play a key role in the mechanism of the anaerobic ring contraction, by providing a one or two electron redox reaction (Scott, 2003). Most CbiH enzymes are similar in sequence to all the class III methyltransferases found in both aerobic and anaerobic pathways.

The CbiH₆₀ from *B. megaterium* is a 540 amino acid protein with a predicted molecular mass of 60 kDa. It is unusual because most CbiH enzymes are only 250 amino acids in length and CbiH₆₀ appears to have been formed as a result of a gene fusion. The N-terminus (amino acids 1-247) encodes the cobalt-precorrin-3 methyltransferase domain, whilst a small region at the C-terminus contains a region with similarity to the NiR/SiR family. This is also found in a number of related *Bacilli* species, including *Geobacillus denitrificans* and *Anoxybacillus flavithermus*. There is some degree of sequence conservation with CobG although this is mainly with the region surrounding the cysteine residues that form the Fe-S centre. The NiR/SiR domain shares closest similarity to NirB from *Paenibacillus* spp with 30% identity and 55% similarity. The NiR/SiR family of proteins catalyse the 6 electron reduction of their substrates to provide reduced nitrogen or sulphur for assimilatory processes or dissimilatory systems involved in energy conservation. This class of reaction involves a range of cofactors including 4Fe-4S clusters, FAD, sirohaem, molybdenum and *c*-type haem, with either a reduced ferredoxin or NAD(P)H as an electron source (Butler *et al.*, 1999; Einsle *et al.*, 2002).

The CbiH₆₀ contains a number of cysteine residues throughout the protein sequence. Four of these cysteines are located in the NiR/SiR domain in two separate motifs (C₄₀₂XXC₄₀₅ and C₄₃₉XXXC₄₄₃). In addition, the CbiH from the archaeon *Methanobacterium thermoautotrophicus* also contains an extended C-terminus with multiple cysteines that coordinate two [4Fe-4S] clusters (A. Brindley, unpublished

data). A similar arrangement is found in various closely related archaeans. The role of these C-terminal extensions remains to be understood and is an integral aim of this thesis.

Only limited turnover of the anaerobic ring contraction step has ever been shown (<5%). The product, cobalt-precorrin-4, has a mass of 950 Da and is described as a green pigment (Santander *et al.*, 1997). The oxidised cobalt-factor IV has a mass spectrum that is two mass units less than predicted and appears as a brown coloured intermediate (PhD Thesis: Frank, 2007). Unfortunately, the synthesis of cobalt-factor IV was only been achieved using crude cell lysates of *E. coli* overexpressing the CbiH from *S. enterica* or *M. thermoautophicus* and no direct activity with CbiH in its pure state has been shown. Moreover, the oxidation state of the cobalt ion was not monitored during these reactions. Therefore, it is still not known if the cobalt ion plays a role during the ring contraction process (Spencer *et al.*, 1998). In addition as the enzymes recognise both cobalt-precorrin-3 and cobalt-factor III as substrates, the process of converting factor intermediates back to the reduced precorrin state is poorly understood. The work in this chapter aims to address both these concerns by studying the function of the *B. megaterium* CbiH₆₀ *in vitro* and providing a clear answer to both the oxidation state of the cobalt ion and the substrate.

5.2 Results

5.2.1 CbiH₆₀ and CbiH₃₀

This project was initiated by over-producing the *B. megaterium* DSM509 CbiH₆₀ in *B. megaterium* DSM319. It was produced both as a full-length version and as a truncated variant (with the C-terminal extension removed), which was named CbiH₃₀. The latter was achieved by cloning the codons corresponding to amino acids 1-247. This deleted the region corresponding to the linker and NiR/SiR domains. The truncation site was chosen after Tyr²⁴⁷, which corresponds to a region where an alignment of ring contraction enzymes including CbiH₆₀, CbiH and CobJ suggested an end to the protein (Figure 5.2). CbiH₆₀ was fused to either an N-terminal (CbiH₆₀^{N-His}) or C-terminal His₆-tag (CbiH₆₀^{C-His}), whilst CbiH₃₀ was fused to a C-terminal His₆-tag (CbiH₃₀^{C-His}). Details of the primary structure and fusion tags are outlined in Figure 5.1.

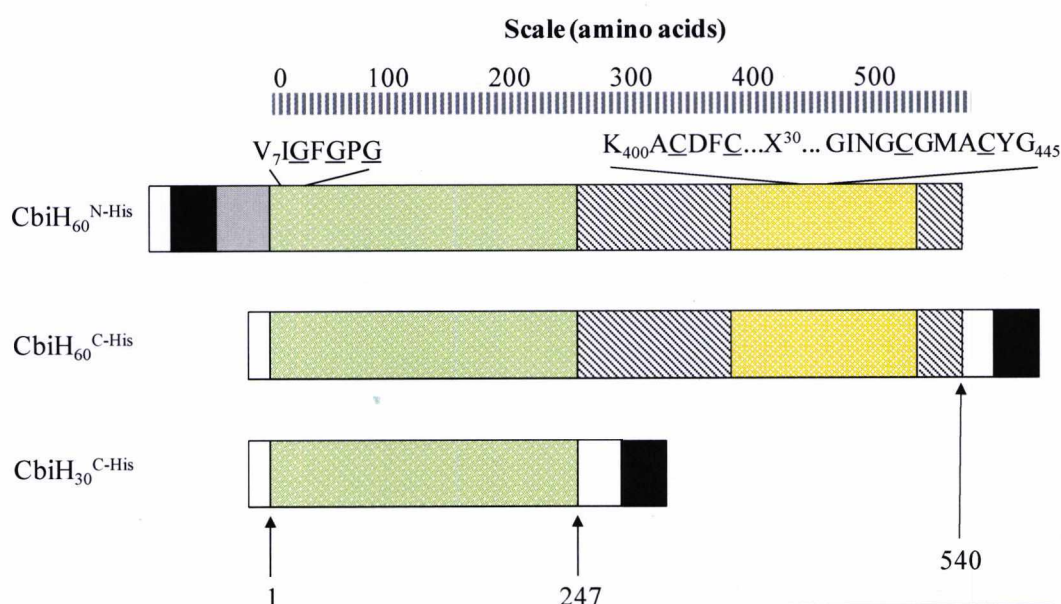
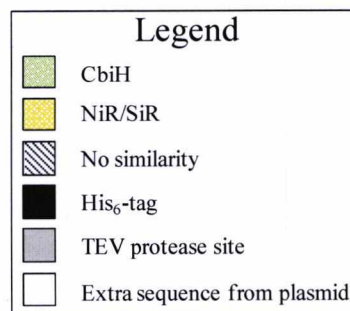
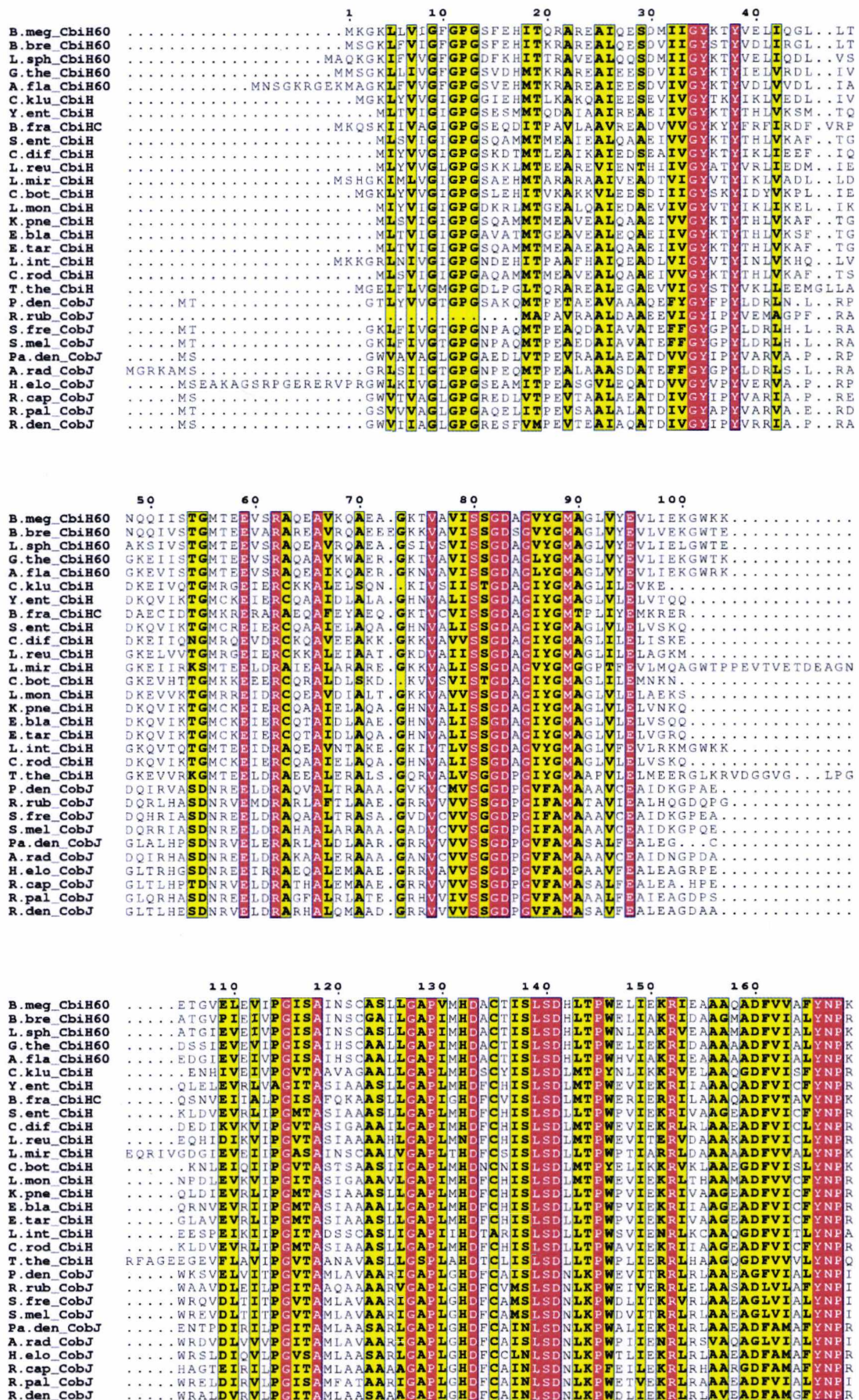
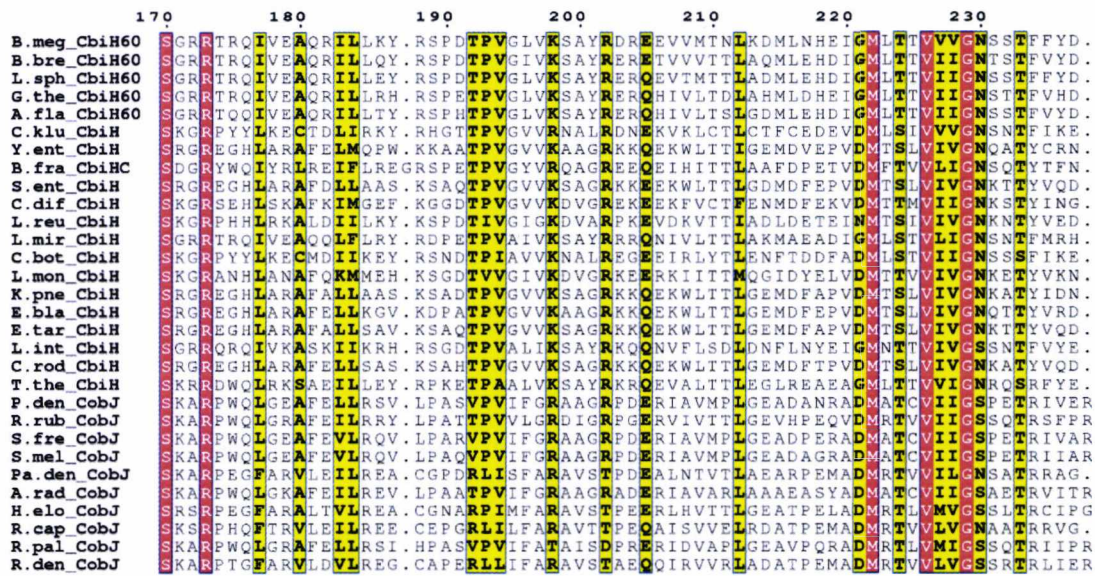


Figure 5.1: Cloning of CbiH₆₀ and truncated CbiH₃₀ from *B. megaterium*. Legend details positions of the CbiH and NiR/SiR domains, regions of no similarity, His₆-tag, TEV cleavage site and extra sequences derived from plasmid. Glycines that are underlined outline a predicted SAM binding domain. Conserved cysteines in the NiR/SiR domain are underlined. Restriction sites are indicated by small arrows. Numbers with arrows indicate amino acid in CbiH₆₀.







Truncation of CbiH₆₀ after Tyr²⁴⁷

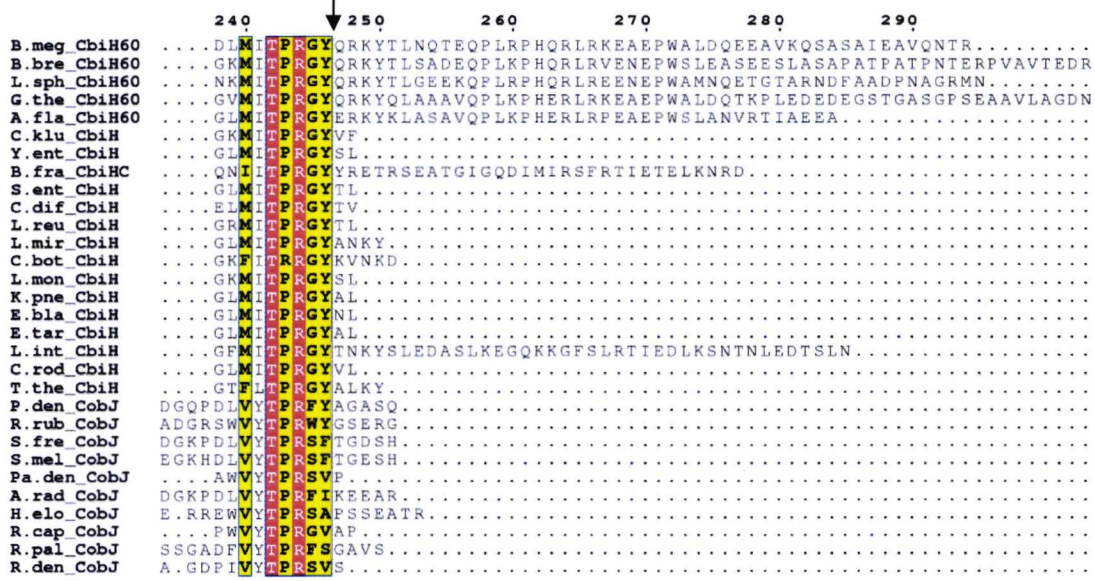


Figure 5.2. Alignment of CbiH₆₀, CbiH and CobJ sequences. Sequences of CbiH₆₀, CbiH and CobJ were aligned with ClustalW2 and graphically visualised on ESPrnt 2.2. The point of truncation is indicated with an arrow. Species abbreviations B.meg (*B. megaterium*), B.bre (*Brevibacillus brevis*), L.sph (*Lysinibacillus sphaericus*), G.the (*G. thermodenitrificans*), A fla (*A. flavithermus*), C.klu (*Clostridium kluveri*), Y.ent (*Yersinia enterocolitica*), B.fra (*Bacteroides fragilis*), S.ent (*S. enterica*), C.dif (*C. difficile*), L.reu (*L. reuteri*), L.mir (*Lautropia mirabilis*), C.bot (*Clostridium botulinum*), L.mon (*Listeria monocytogenes*), K.pne (*Klebsiella pneumonia*), E.bla (*Escherichia blattae*), E.tar (*Edwardsiella tarda*), L.int (*Leptospira interrogans*), C.rod (*Citrobacter rodentium*), T.the (*Thermus thermophiles*), P.den (*P. denitrificans*), R.rub (*Rhodospirillum rubrum*), S.fre (*Sinorhizobium fredii*), S.mel (*Sinorhizobium meliloti*), Pa.den (*Paracoccus denitrificans*), A.rad (*Agrobacterium radiobacter*), H.elo (*Halomonas elongata*), R.cap (*R. capsulatus*), R.pal (*Rhodopseudomonas palustris*), R.den (*Roseobacter denitrificans*).

5.2.2 Purification of CbiH₆₀

Recombinant strains of *B. megaterium* DSM319 containing the relevant CbiH-encoding plasmids were pre-cultured in LB medium at 37 °C for 16 hrs. These were then sub-cultured into 1 litre of LB medium and incubated at 37 °C until an OD_{578nm} of 0.3 was reached. D-xylose was added to 0.5% (w/v) to induce protein production and the culture was left overnight. After harvesting cells and sonication, the initial purification was performed under aerobic conditions with a His₆-bind column. It was noted that the solubility of the C-terminal His₆-tag CbiH₆₀ (CbiH₆₀^{C-His}) was higher than the N-terminal His₆-tag (CbiH₆₀^{N-His}), as judged by SDS-PAGE (see Figure 5.5). Both enzymes migrated on SDS-PAGE with an approximate molecular mass of 65 kDa (5 kDa added from the His₆-tag and *xyIA* gene fusion). Although more highly produced, the CbiH₆₀^{N-His} protein was primarily found in the insoluble protein fraction, indicating that the majority of the protein had formed inclusion bodies. Initially, growth conditions were optimised to improve the solubility and yield. This established the yield of soluble protein could be improved for both proteins by lowering the growth temperature to 30 °C and reducing the concentration of D-xylose to 0.25% (w/v). The CbiH₆₀^{N-His} enzyme yielded between 10-30 mg L⁻¹ whilst the CbiH₆₀^{C-His} was produced at between 5-10 mg L⁻¹. In addition, it was observed that CbiH₆₀ had a tendency to aggregate and precipitate during the purification and post-purification stages. This occurred with both His₆-tagged forms and especially if the enzyme concentration was kept above 5-10 mg ml⁻¹ for more than a couple of hours at 4 °C. This could be prevented by using MOPS, pH 7.0, and L-arginine in the purification buffers. L-arginine has affinity for chelated nickel and can only be used at low concentrations (Abe *et al.*, 2009). After buffer exchange into 20 mM MOPS, pH 7.0, 500 mM NaCl, 50 mM L-arginine and 100 mM (NH₄)SO₄ the protein can be repeatedly freeze-thawed at -80 °C without any precipitation. Unfortunately, the protein was found to be unstable at lower NaCl concentrations (< 400 mM) and no other suitable buffer has been identified.

During purification of CbiH₆₀^{C-His}, the protein appeared as a dark coloured band at the top of the His₆-bind column. After washing the column with increasing concentration of imidazole this changed to a gold colour that diminished in intensity with time. The protein eluted as a pale golden colour and was analysed by UV-Vis spectrometry. It was found to have a spectrum with peaks at 340 nm and 410 nm

(Figure 5.3), with an A_{280}/A_{410} ratio of approximately 4:1. This finding was consistent for the presence of an iron-sulphur centre (Ollagnier-De Choudens *et al.*, 2000). If the protein was left exposed to air, the overall absorbance decreased with time (3-4 hrs), with the A_{280} remaining constant (Figure 5.4). However, if the protein was purified under anaerobic conditions the colouration of the protein remained stable for several days. Although the UV-Vis spectrum was far less defined, the addition of excess dithionite resulted in a decrease in absorbance in the UV-Vis region, with the A_{280} remaining constant. This effect could also be achieved by using DTT, although the latter required a longer incubation period to achieve this effect.

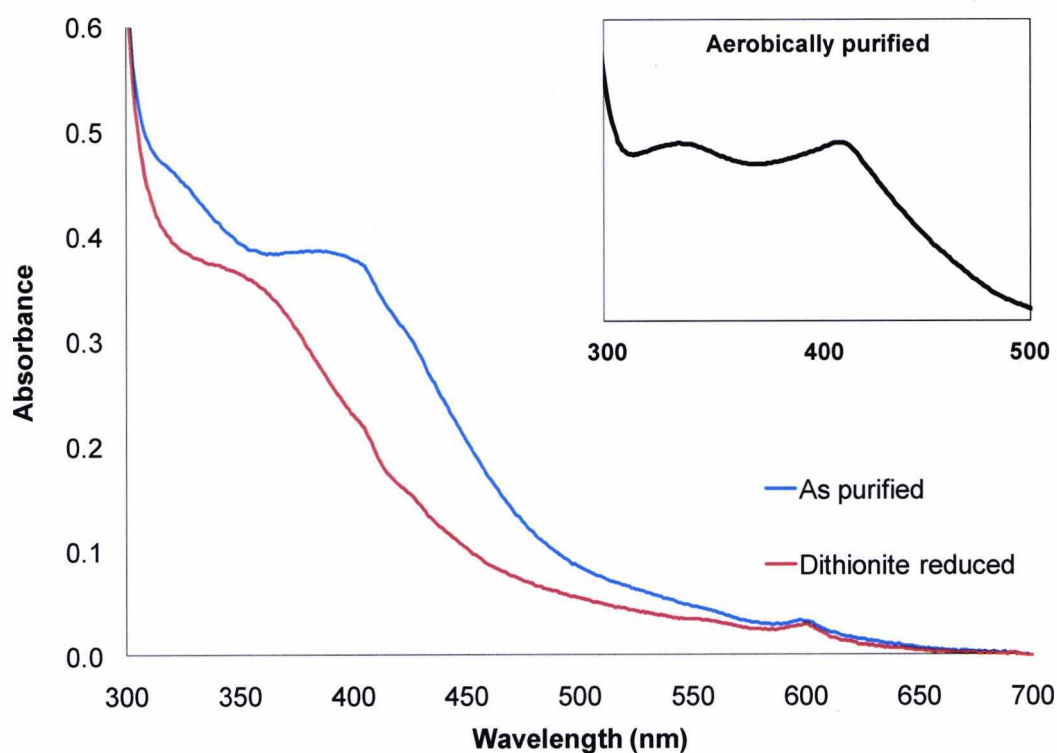


Figure 5.3. UV-Vis spectra of purified $\text{CbiH}_{60}^{\text{C-His}}$. Purified protein as purified anaerobically and buffer exchanged (blue line). Reduced protein (red line) after buffer exchange to remove excess dithionite. Reduction indicated as broad decrease in absorbance in the UV-Vis region, with the A_{280} remaining unchanged. Inset diagram is the UV-Vis spectrum of aerobically purified $\text{CbiH}_{60}^{\text{N-His}}$, with the absorbance at 350 and 420 nm more prominent.

5.2.3 CbiH₆₀ contains a [4Fe-4S] cluster

From the UV-Vis spectrum it was thought that CbiH₆₀ could contain an iron-sulphur centre. In addition, the enzyme has four conserved cysteines composed from two spatially close C₄₀₂XXC₄₀₄ and C₄₃₉XXXC₄₃₄ motifs. A common feature of iron-sulphur proteins is the loss of the iron-sulphur cofactor in the presence of oxygen and quenching of the UV-visible spectrum. However, this could be prevented if the enzyme was purified anaerobically. This revealed the enzyme remained the same colour for several days. However, there was a slight difference in the UV-visible spectrum of the anaerobically purified CbiH₆₀ when compared to aerobic purification. This difference in spectrum is likely due to a deterioration of the Fe-S centre under aerobic conditions. To confirm the identity of the redox group, EPR spectroscopy was performed in collaboration with Dr Steve Rigby, Manchester University. The protein was first reduced with excess dithionite (E° -0.66 V at pH 7), which leads to a bleaching of its colour and a shift in the UV-visible spectrum with a broad decrease in the shoulder between 350-410 nm (Figure 5.3). The EPR spectrum at 15K provided evidence of a reduced [4Fe-4S] centre with g values at 2.04, 1.94 and 1.90 (Figure 5.4). Further evidence of a [4Fe-4S] centre was obtained when it was observed that the spectrum was lost when the temperature is raised to 70K. This is a distinguishing test to differentiate [4Fe-4S] and [2Fe-2S] centres (Schroeder *et al.*, 2009). In addition, the iron content of the anaerobically purified protein was determined and estimated to contain 4.8 mol of iron per mol of CbiH₆₀. An attempt to measure the redox potential of the protein was also attempted in collaboration with Professor Andy Munro and Dr Kirsty McClean at the University of Manchester. However, the midpoint redox potential for the [4Fe-4S] cluster could not be fully determined, due to the instability of the protein during the assay. Nonetheless, from the data that was recorded it appears that the midpoint redox potential might be quite high, with an estimate of -50 mV (Dr Kirsty McClean, personal communication).

5.2.4 Purification of CbiH₃₀

After purification of CbiH₃₀ on a His₆-bind column, SDS-PAGE showed the enzyme migrated with an approximate mass of 31 kDa. Interestingly, unlike the full-length CbiH₆₀, no degradation was observed. Furthermore an additional band was observed at 35 kDa, which was thought to be CbiX^L derived from genomic expression of the

cobI operon. Both bands were extracted, digested with trypsin and analysed by MALDI to confirm their identity. After several purifications it was established the yield of enzyme varied considerably from 1-20 mg L⁻¹. This inconsistency could not be explained and sometimes no enzyme was purified. CbiH₃₀^{C-His} purifies as a colourless enzyme with no absorbance features between 350-420 nm. This indicated that the truncation had removed the [4Fe-4S] centre.

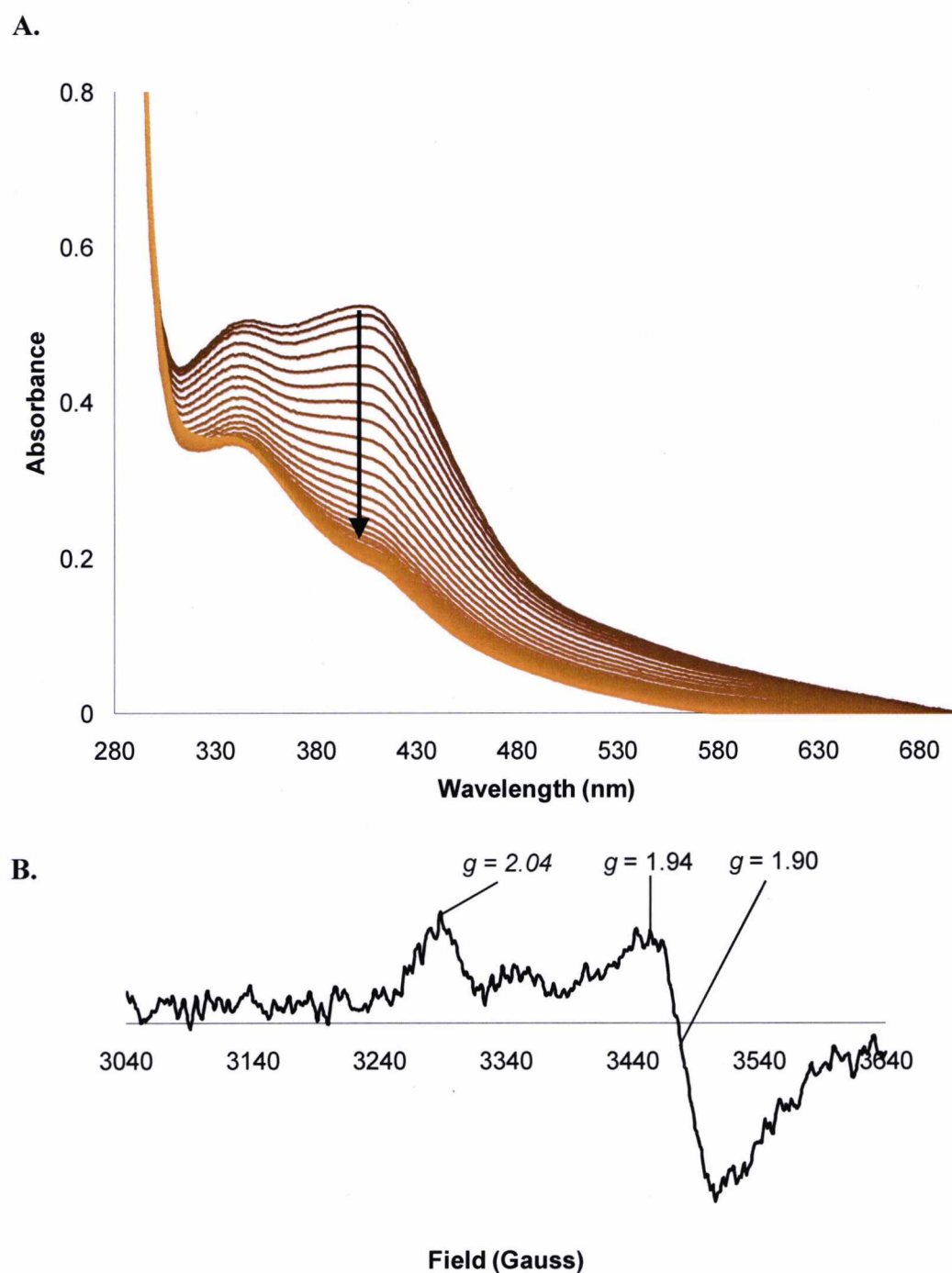


Figure 5.4. CbiH₆₀ contains a [4Fe-4S] centre. **A.** 50 μM CbiH₆₀^{C-His} (3 mg ml⁻¹) of enzyme was buffer exchanged into Buffer H. The protein was purified anaerobically and then exposed to oxygen before scanning every 15 minutes. **B.** 200 μM (12 mg ml⁻¹) CbiH₆₀^{C-His} (buffer-exchanged in Buffer A) was reduced with excess dithionite (10 mM) and frozen in liquid nitrogen. Samples were analysed at 15K on a Bruker ELEXSYS E500 spectrometer operating at X-band, employing a Super High Q cylindrical cavity (Q factor \sim 16,000) equipped with an Oxford Instruments ESR900 liquid helium cryostat linked to an ITC503 temperature controller. The iron-sulphur spectra of CbiH₆₀^{C-His} was recorded at 12 K employing a microwave power of 1 mW, modulation frequency of 100 kHz and modulation amplitude of 5 G. Spectra of CbiH₆₀^{C-His} was also recorded at 70 K, which resulted in a loss of signal. This indicated the iron-sulphur centre was of the [4Fe-4S] class.

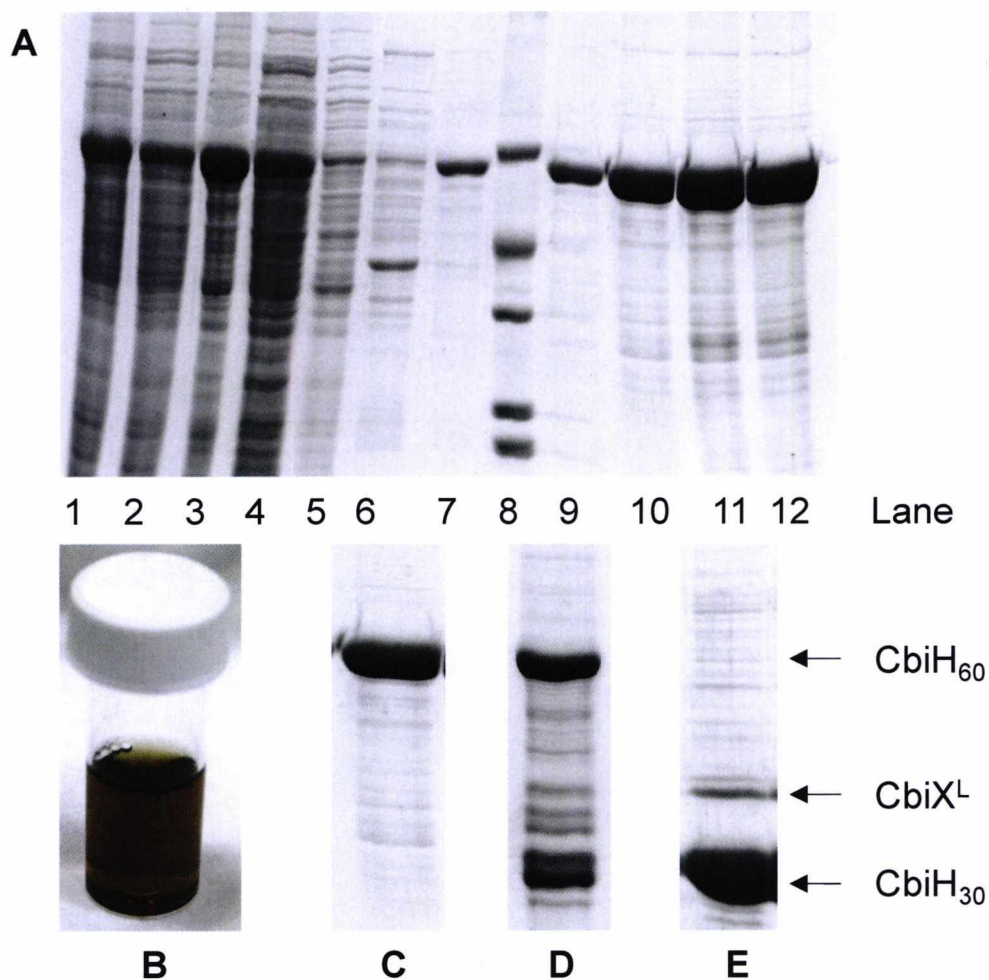


Figure 5.5. Purification of *B. megaterium* CbiH. **A.** SDS-PAGE of CbiH₆₀^{N-His} purification. Lanes as follows: 1- crude extract, 2- soluble supernatant, 3- insoluble fraction, 4- supernatant flow-through, 5- binding buffer flow-through, 6- wash buffer I flow-through, 7- wash buffer II flow-through, 8- Dalton VII marker, 9, 10, 11, 12- elution fractions. **B.** purified CbiH₆₀^{N-His} 10 mg ml⁻¹. Elution fractions of **C.** CbiH₆₀^{N-His}, **D.** CbiH₃₀^{C-His} and **E.** CbiH₃₀^{C-His}. An additional band was observed during the purification of CbiH₃₀^{C-His}. Both bands were extracted, trypsin digested and analysed by MALDI-TOF. This confirmed the proteins as CbiX^L and CbiH₃₀. After several His₆-tag purifications it was noted CbiX^L was a common contaminant for all purifications. Its presence could be omitted by simply adding 100 nM cyanocobalamin to the culture.

5.2.5 Preparation of the substrate, cobalt-factor III

Cobalt-factor III was prepared using a multi-enzyme incubation. To assist in this the vector pET-coco-2-*cobA-hemB-hemC-hemD-sirC-cbiX^S-cbiL* (constructed by Dr Evelyne Deery and Dr Steffi Frank) was provided (all genes fused to N-terminal His₆-tags). This plasmid contains the genes encoding for the enzymes that collectively are able to transform ALA to cobalt-factor III. These were all recombinantly produced in *E. coli* BL21Star(DE3) pLysS. Following His₆-bind purification, the purified enzymes were then incubated under anaerobic conditions with 5-aminolevulinic acid, SAM, NAD⁺ and CoCl₂×6H₂O in Buffer A at 37 °C (see Appendix V for a detailed protocol). The substrate cobalt(II)-factor III is not air stable and even under anaerobic conditions is slowly oxidised to cobalt(III)-factor III, as observed by EPR (PhD Thesis: Frank, 2007; PhD Thesis: Leech, 2004). This can also be observed as a shift in the main UV-Vis maximum absorbance at 392 nm to two new peaks at 414 and 425 nm. To keep the intermediate in the reduced Co(II) state, an excess of DTT has to be added. The UV-Vis spectra of cobalt(II)-factor III and the oxidised cobalt(III)-factor III are shown in Figure 5.6.

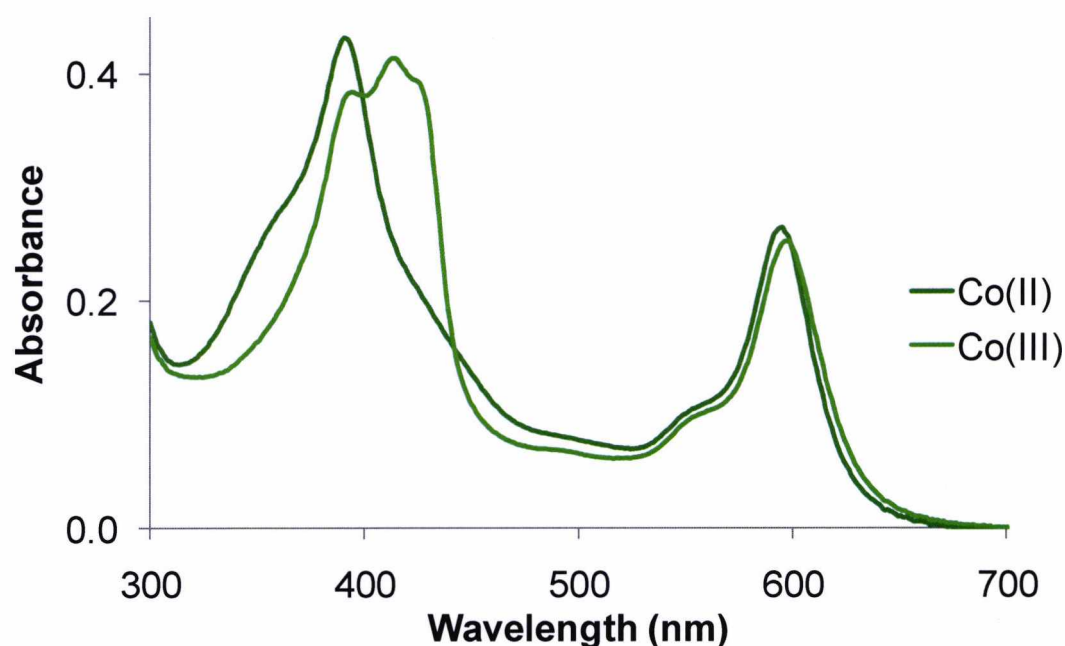


Figure 5.6. UV-Vis spectra of reduced and oxidised cobalt-factor III. Cobalt(II)-factor III (dark green) and cobalt(III)-factor III (light green). Using the UV-Vis maximum at 392 nm, the extinction coefficient for cobalt(II)-factor III is estimated as $1.6 \times 10^5 \text{ M}^{-1} \text{ cm}^{-1}$.

5.2.6 UV-Vis of CbiH₆₀ and CbiH₃₀ titrations with cobalt(II)-factor III

When CbiH₆₀^{C-His} was incubated with cobalt(II)-factor III the UV-Vis spectrum shifted with a decrease in the absorbance peak at 392 nm, with a slight shoulder forming between 400-425 nm. If an excess of CbiH₆₀ was added to cobalt(II)-factor III, the UV-Vis spectrum shifted to two separate peaks at 414 and 425 nm. This indicated that CbiH₆₀ oxidised the intermediate to cob(III)-factor III. This dependence could be demonstrated further by titrating CbiH₆₀ with cobalt(II)-factor III, with the shift dependent on enzyme concentration (Figure 5.7). This was not stoichiometric, as a full shift required an excess of CbiH₆₀. In addition the shift was not dependent on the presence of SAM. To investigate this further a negative control was also prepared, where buffer was titrated with cobalt-factor III. This did not reveal any changes in the spectrum and suggested that the oxidation shift was CbiH₆₀ dependent. If the [4Fe-4S] cluster of CbiH₆₀ was first reduced with dithionite and then buffer exchanged, a titration of the reduced enzyme with cobalt(II)-factor III did not give rise to any spectral shifts. If the truncated CbiH₃₀ was also titrated with cobalt(II)-factor III, this did not reveal any significant changes.

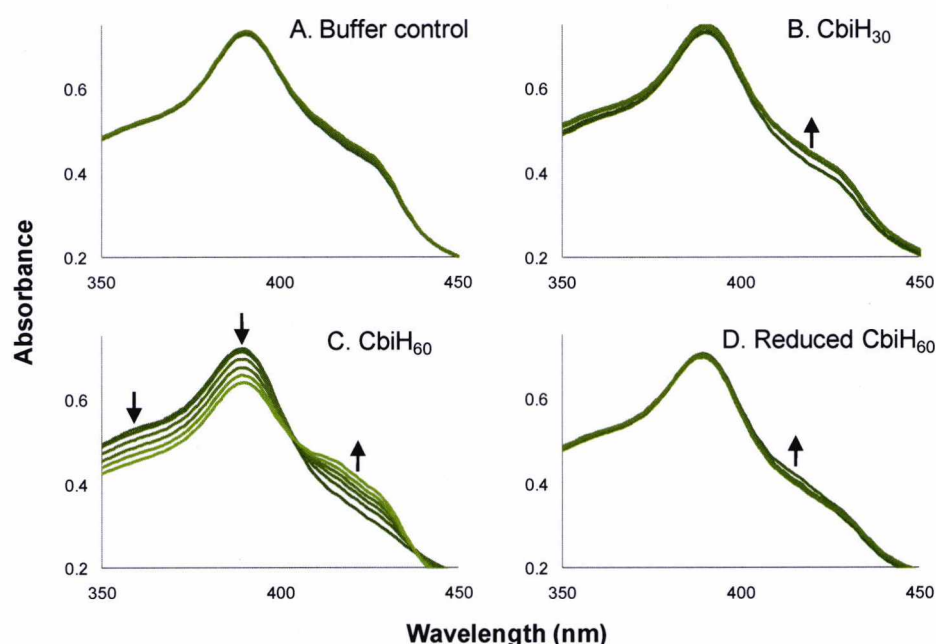


Figure 5.7. UV-Vis of 5 μ M cobalt(II)-factor III incubated with increasing CbiH concentration. A. Buffer control. Buffer H (see table 4.2) with no enzyme was added. B. As purified CbiH₃₀^{C-His}. C. As purified (oxidised) CbiH₆₀^{C-His}. D. Reduced CbiH₆₀^{C-His}. CbiH₆₀^{C-His} was reduced with excess dithionite and buffer exchanged into Buffer H. All enzyme concentrations were titrated from 0-10 μ M

5.2.7 CbiH₆₀ catalyses the C-17 methylation and ring contraction reaction

A strategy to solve the anaerobic ring contraction reaction began by linking a couple of simple observations. As mentioned in Chapter 4, during the purification of CbiH₆₀^{C-His}, a dark band was observed at the top of the His₆-bind column, which washed off as a green-brown coloured solution. LC-MS analysis identified this as cobalt-factor II. Ring contraction is likely to be a bottleneck in the B₁₂ pathway and the absence of any cobalt-factor III suggests that CbiH₆₀^{C-His} is active as any converted product will be passed on through the pathway. In addition, the spectral shifts observed with oxidised CbiH₆₀^{C-His} and cobalt-factor III indicated the enzyme was binding the substrate or oxidising the cobalt ion. A simple test to show that the enzyme was bound to its substrate was developed by mixing them together and passing through a desalting PD-10 column. This demonstrated the enzyme would co-elute with cobalt(III)-factor III. However, if only substrate was added the majority was retained on the column. This suggested the CbiH₆₀^{C-His} enzyme is able to bind its substrate with high affinity. In addition the only suitable buffer that could prevent precipitation was a high salt buffer, such as 50 mM Tris-HCl, pH 8.0, containing 400 mM NaCl (Buffer H). Interestingly the mixture of CbiH₆₀^{C-His} and cobalt(II)-factor III was found to remain stable in Buffer H for several days at 37 °C, whilst CbiH₆₀^{C-His} on its own in Buffer H would slowly precipitate within a few hours post-purification.

Using purified CbiH₆₀^{C-His} and CbiH₆₀^{N-His}, enzyme-substrate mixtures were prepared by incubating the enzyme variants with cobalt(II)-factor III, before exchanging into Buffer H. These mixtures were then incubated with SAM (1 mM) both in the absence and presence of reducing agent (10 mM DTT). After overnight incubation at 37 °C, a new intermediate formed but only in the reaction containing cobalt(II)-factor III, CbiH₆₀^{C-His}, SAM and DTT. This new intermediate at first appeared brown in colour. The UV-Vis spectrum of this reaction was compared to the reaction without CbiH₆₀^{C-His}. This revealed that the main peaks at 392 and 595 nm for cobalt(II)-factor III had reduced in size. In addition, a new broad absorbance peak appeared between 300-350 nm. From the UV-Vis spectrum it appears that this brown intermediate is a mixture of a new species and cobalt(II)-factor III. This new intermediate was purified by first acidifying the solution with 1% (v/v) acetic acid, to remove proteins, followed by centrifugation. This step removed any residual

cobalt-factor III, which precipitated with CbiH₆₀^{C-His} leaving a red coloured supernatant. This was then bound to an RP-18 column, washed and eluted in 50% (v/v) ethanol. After freeze-drying, samples were analysed by LC-MS using Method A (see materials and methods). The intermediate eluted at 18 min with a peak at m/z 948.24 (Figure 5.8). This was compared to previous UV-Vis and MS spectra, with the intermediate identified as cobalt-factor IV (PhD Thesis: Frank, 2007). After several repeats it was found that the yield of the reaction could be improved by increasing the enzyme to substrate ratio. This indicated that only a fraction of the purified enzyme was active. Monitoring the reaction by continuous scanning of the UV-Vis spectrum, demonstrated that a full conversion of cobalt-factor III to cobalt-factor IV could be obtained (Figure 5.9). This revealed a defined peak at 335 nm. However, if cobalt-factor IV was purified the UV-Vis spectrum became broader with no easily distinguishable maximum peaks in the UV-Vis range (Figure 5.10). By air oxidising the pure cobalt-factor IV and then rescanning the UV-Vis spectrum, a shift was observed with new absorbance maxima at 346, 533 and 585 nm. This spectrum resembles the spectra obtained during LC-MS.

These reactions were repeated several times with other pure CbiH enzymes. The *B. megaterium* CbiH₆₀^{N-His}, CbiH₃₀^{C-His} and *S. enterica* CbiH were inactive under the conditions tested. Due to time constraints, other parameters such as buffer composition were only briefly investigated. In summary, no reaction was observed if the NaCl concentration was below 400 mM. In addition, the CbiH₆₀^{C-His} enzyme was active in the temperature range of 4 to 37 °C although this was not investigated in detail.

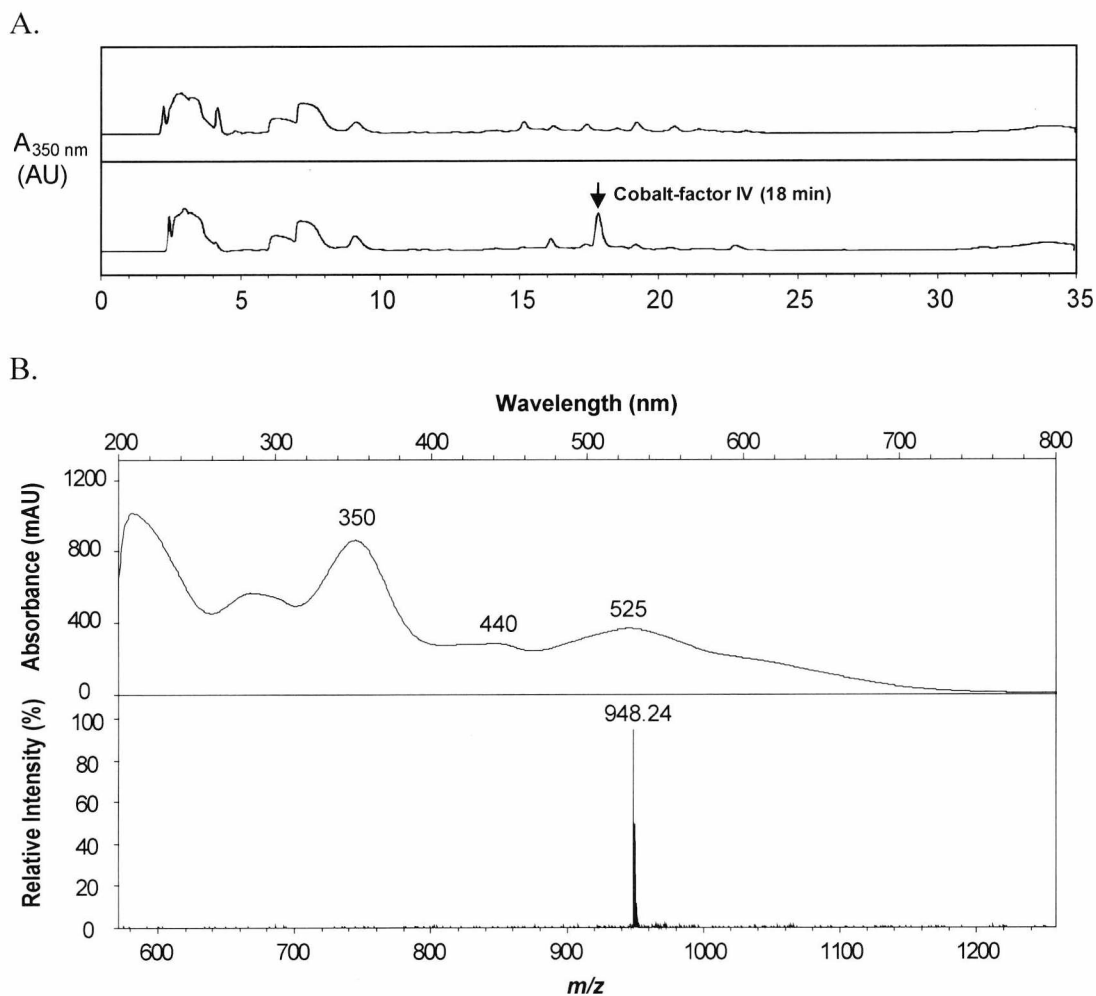


Figure 5.8. LC-MS of cobalt-factor IV. A. LC chromatogram of sample derived from reactions containing CbiH₆₀^{C-His}, cobalt-factor III, SAM and in absence (top panel) and presence (bottom panel) of DTT. Due to acidification most of the cobalt-factor III precipitated with the enzyme. B. UV-Vis and MS of cobalt-factor IV.

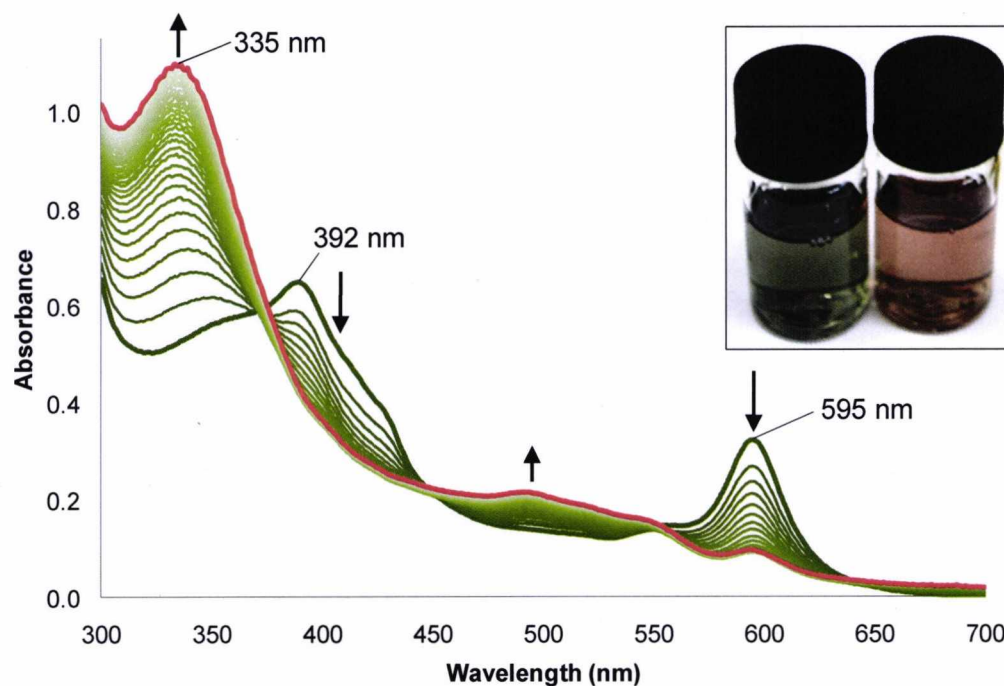


Figure 5.9. Conversion of cobalt(II)-factor III to cobalt-factor IV by UV-Vis spectroscopy. The reaction was performed at 37 °C, with 5 μM cob(II)alt-factor and 20 μM CbiH₆₀^{C-His} in the presence of 1 mM SAM and 10 mM DTT. UV-Vis scans were measured every 15 min. Inset picture shows a similar reaction with 30 μM cobalt-factor III, 30 μM CbiH₆₀^{C-His}, 10 mM DTT in the absence (left vial) and presence (right vial) of 1 mM SAM, after incubation at 37 °C for 24 hr.

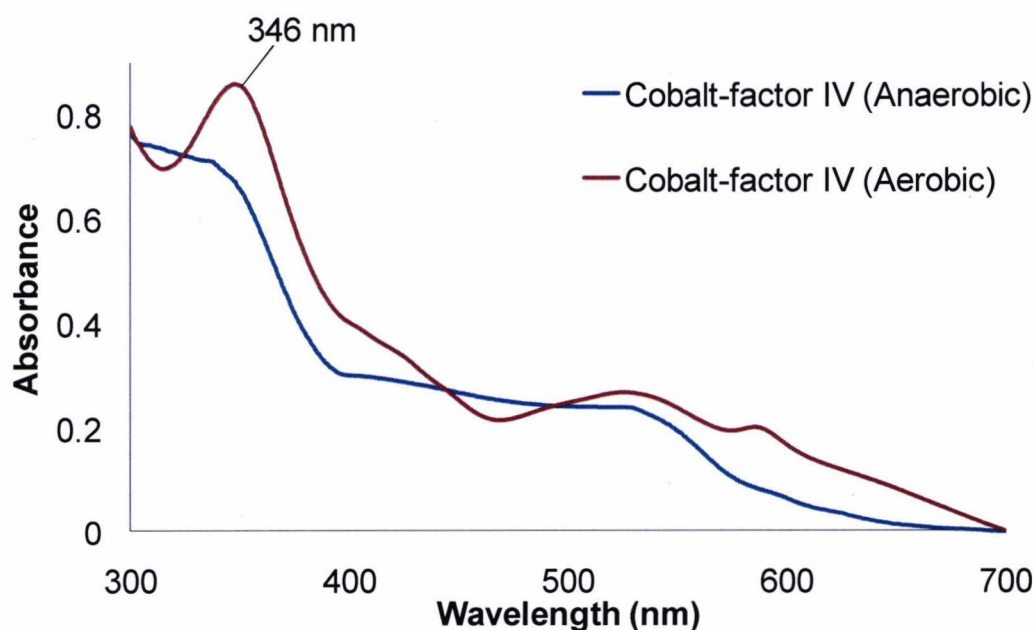


Figure 5.10. UV-Vis spectra of purified cobalt-factor IV. Cobalt-factor IV was purified from cobalt-factor III anaerobically as described and the UV-Vis spectrum was measured (blue line). The sample was then exposed to air and rescanned (red line).

5.2.8 EPR of CbiH₆₀ and cobalt-factor III incubations

To investigate the ring contraction reaction further, EPR was used to probe for changes in the environment of the cobalt ion and the [4Fe-4S] cluster in CbiH₆₀^{C-His} during the reaction. High concentrations of cobalt(II)-factor III and CbiH₆₀^{C-His} are required to obtain EPR signals. Cobalt(II)-factor III (600 μM) was prepared in a one-step incubation and used for EPR without any further purification. To purify CbiH₆₀^{C-His} at a high concentration, several litres of culture were grown, before the protein was purified anaerobically on a single His₆-bind column. To keep the protein stable and active, the purified protein (~230 μM) was immediately exchanged into Buffer H, before samples for EPR were set up. Only a limited quantity of CbiH₆₀^{C-His} was obtained, so two control incubations (enzyme only and enzyme + substrate) were omitted. However, other controls were prepared, which allowed for qualitative analysis of the reactions. The set-up for the reactions is shown in Table 5.1.

Table 5.1. Reactions for EPR analysis

Reaction	Sample description	Time	
		10 min	24 hr
1	Cobalt(II)-factor III	+	-
2	Cobalt(II)-factor III + SAM	+	-
3	Cobalt(II)-factor III + DTT	+	-
4	Cobalt(II)-factor III + SAM + DTT	+	-
5	CbiH ₆₀ + SAM	+	-
6	CbiH ₆₀ + DTT	+	-
7	CbiH ₆₀ + SAM + DTT	+	-
8	Cobalt(II)-factor III + CbiH ₆₀ + SAM	+	+
9	Cobalt(II)-factor III + CbiH ₆₀ + DTT	+	+
10	Cobalt(II)-factor III + CbiH ₆₀ + SAM + DTT	+	+

Incubations containing cobalt(II)-factor III (100 μM), CbiH₆₀^{C-His} (175 μM), SAM (5 mM) and DTT (10 mM) were prepared in Buffer H (see table 4.2). All samples were pre-incubated for 10 min at 37 °C and a 300 μl aliquot was withdrawn and transferred into an EPR tube, before sealing and freezing in liquid nitrogen. Samples containing CbiH₆₀ with cobalt-factor III and cofactors were also withdrawn after 24 hr at 37 °C.

After a 24 hr incubation only the reaction containing cobalt(II)-factor III, $\text{CbiH}_{60}^{\text{C-His}}$, SAM and DTT showed a change in the UV-Vis spectrum and visual colour. The UV-Vis spectra of the EPR samples for this reaction (10 min and 24 hr) are shown in Figure 5.11. Visually, it seems the reaction had only partially gone to completion, with a brown colour indicating a mixture of cobalt-factor III and cobalt-factor IV. The UV-Vis spectra estimated that an approximate 20-25% turnover had been achieved. This was indicated with a decrease in absorbance at 392 and 595 nm, with slight broad increases at 330 and 500 nm.

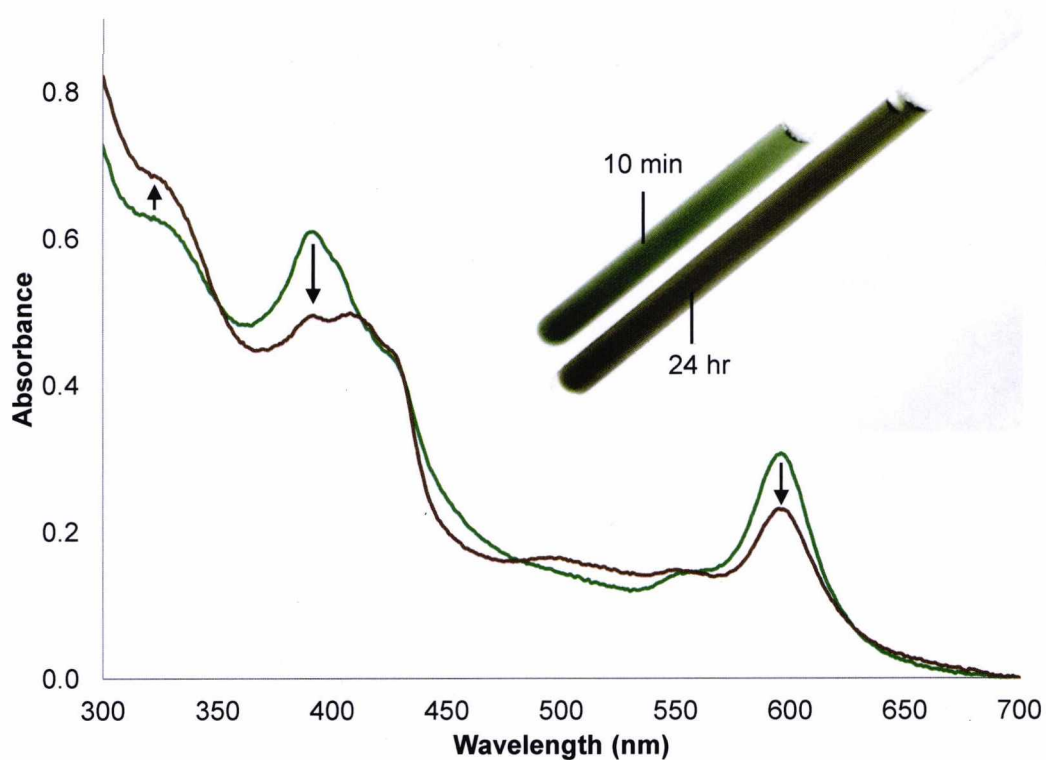


Figure 5.11. UV-Vis spectra and inset picture of EPR sample containing cobalt(II)-factor III, $\text{CbiH}_{60}^{\text{C-His}}$, SAM and DTT. UV-Vis spectra of sample pre-incubated at 37 °C for 10 min (green line) and 24 hr (brown line). Inset picture of EPR tubes. Note samples were analysed by EPR first, before the samples were defrosted under anaerobic conditions. The UV-Vis spectra and a picture were then taken.

The cobalt ion in cobalt(II)-factor III is paramagnetic with a single unpaired electron in the d_{z^2} orbital. This is a low-spin species with a cylindrical symmetry, the d_{z^2} orbital being perpendicular to the surrounding pyrrolic ring system leading to an axial EPR spectrum. The unpaired electron of cobalt(II) is sensitive to surrounding electron fields both perpendicular (\perp) and parallel (\parallel) to its orbital, leading to an EPR spectrum with two g values (g_{\perp} and g_{\parallel}). In addition the nuclear spin ($7/2$) of the ^{59}Co ion leads to hyperfine splitting (A) with two sets of hyperfine coupling. One set is on the g parallel (A_{\parallel}), providing information on axial ligation of the cobalt ion. The second set is the g perpendicular (A_{\perp}), which is attributed to the cobalt bonds provided by the four ^{14}N ligands from the porphyrin/corrin ring system.

For the control incubations (1-4), free cobalt(II)-factor III is in a low-spin base-on state with imidazole presumably bound as both upper and lower axial ligands. This leads to a g_{\perp} value of 2.28 with typical hyperfine splitting observed with previously observed spectra (Frank *et al.*, 2007). The addition of SAM leads to changes in the hyperfine splitting of A_{\perp} , but not A_{\parallel} . A similar but spectrally different effect is also observed with the addition of DTT to cobalt(II)-factor III. A summary of these reactions is shown in Figure 5.12. Changes in A_{\perp} suggest that both SAM and DTT can interact with cobalt(II)-factor III, although this interaction is not through axial ligation of the cobalt(II) ion. For $\text{CbiH}_{60}^{\text{C-His}}$ control incubations (5-7), the presence of SAM, DTT or both together, did not show any changes. Only a small amount of reduced $[\text{4Fe-4S}]^{2+}$ cluster was observed (data not shown), which was expected as the enzyme was purified in the oxidised state. In addition a small signal for a contaminating copper radical was observed with a $g = 2.00$, with other weak signals for high-spin rhombic iron at $g = 4.26$. This latter observation indicates that any $[\text{4Fe-4S}]$ cluster present is intact and potentially functional.

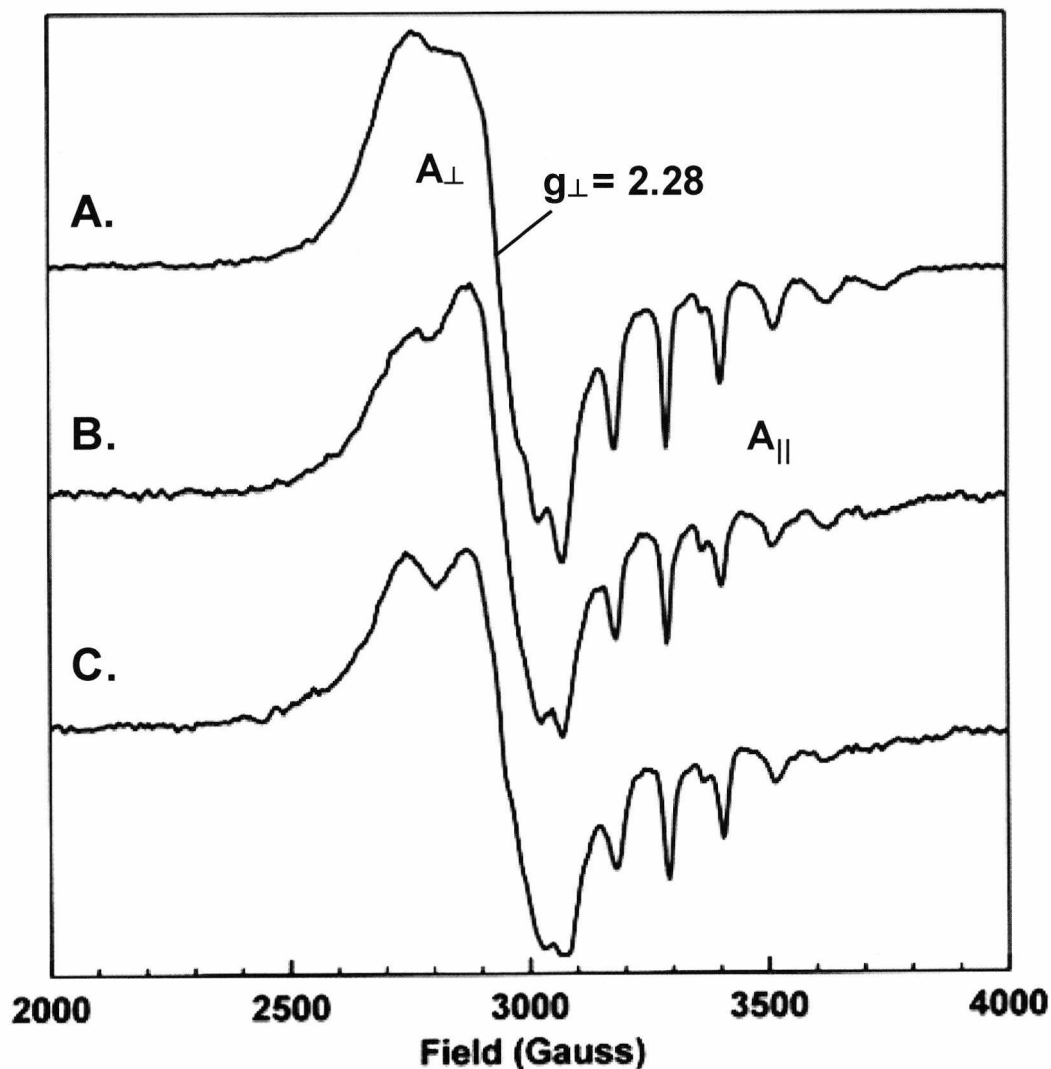


Figure 5.12. EPR trace of cobalt(II)-factor III (100 μM) incubated on its own (A), SAM (B) or DTT (C). The low-spin cobalt(II) produces an EPR signal with a $g_{\perp} = 2.28$, which indicates base-on cobalt(II). It is likely that contaminating imidazole is acting as an upper and lower axial ligand. Addition of SAM leads to changes in hyperfine splitting of A_{\perp} , whilst there are no changes in A_{\parallel} . This indicates that SAM interacts with cobalt-factor III, but not as a direct ligand to the cobalt(II) ion. Addition of DTT leads to similar changes as observed with SAM. This indicates that DTT also interacts with cobalt-factor III, but not the same as the SAM interaction.

As expected the control reaction containing cobalt(II)-factor III, $\text{CbiH}_{60}^{\text{C-His}}$ and SAM does not lead to the formation of cobalt-factor IV. By monitoring this reaction by EPR, the addition of $\text{CbiH}_{60}^{\text{C-His}}$ to cobalt(II)-factor III with SAM, leads to slow changes in both the g perpendicular and g parallel hyperfine splitting. Focusing on the parallel effects, this is dominated by three main features indicated in Figure 5.13. One feature with $g_{\parallel} = 2.020$ in the free cobalt(II)-factor III with SAM control, shows a small decrease to $g_{\parallel} = 2.018$ upon addition of $\text{CbiH}_{60}^{\text{C-His}}$. This shift is essentially complete after 24 hr, indicating a possible change in the ligation state of the cobalt ion. This could be interpreted as ternary complex formation, with binding of the cobalt ion to an amino acid ligand provided by $\text{CbiH}_{60}^{\text{C-His}}$. This has previously been observed with CbiL and cobalt(II)-factor II in the presence of SAM (Frank *et al.*, 2007). However, it is important not to exclude a change in orientation of the imidazole artefact that is presumably bound as an axial ligand to the cobalt ion. Moreover, it was noticed that there was no change in cobalt oxidation state upon the binding of $\text{CbiH}_{60}^{\text{C-His}}$ to cobalt(II)-factor III. This was unexpected, as previous UV-Vis analysis showed a shift in spectrum towards the spectrum for oxidised cobalt(III)-factor III. As there is no change in the oxidation state of the cobalt ion in the EPR spectrum, this suggests that this is a binding event and not oxidation of the cobalt ion.

For the control reaction with cobalt(II)-factor III, $\text{CbiH}_{60}^{\text{C-His}}$ and DTT, again this does not lead to the formation of cobalt-factor IV. Instead when compared to the equivalent reaction (cobalt-factor III, $\text{CbiH}_{60}^{\text{C-His}}$ and SAM), a distinct difference is observed after initial mixing of $\text{CbiH}_{60}^{\text{C-His}}$ with either of the cofactors SAM or DTT (data not shown). With the SAM present it is presumed parallel changes in the hyperfine splitting indicate ternary complex formation, with a ligand to the cobalt ion provided by $\text{CbiH}_{60}^{\text{C-His}}$. With DTT present instead of SAM, these changes do not occur during the initial mixing. However, after 24 hr, a shift in the same parallel features occurs. This suggests that the binding of cobalt(II)-factor III to $\text{CbiH}_{60}^{\text{C-His}}$ is ligand gated, with faster changes observed in the presence of SAM. Alternatively DTT could inhibit this binding.

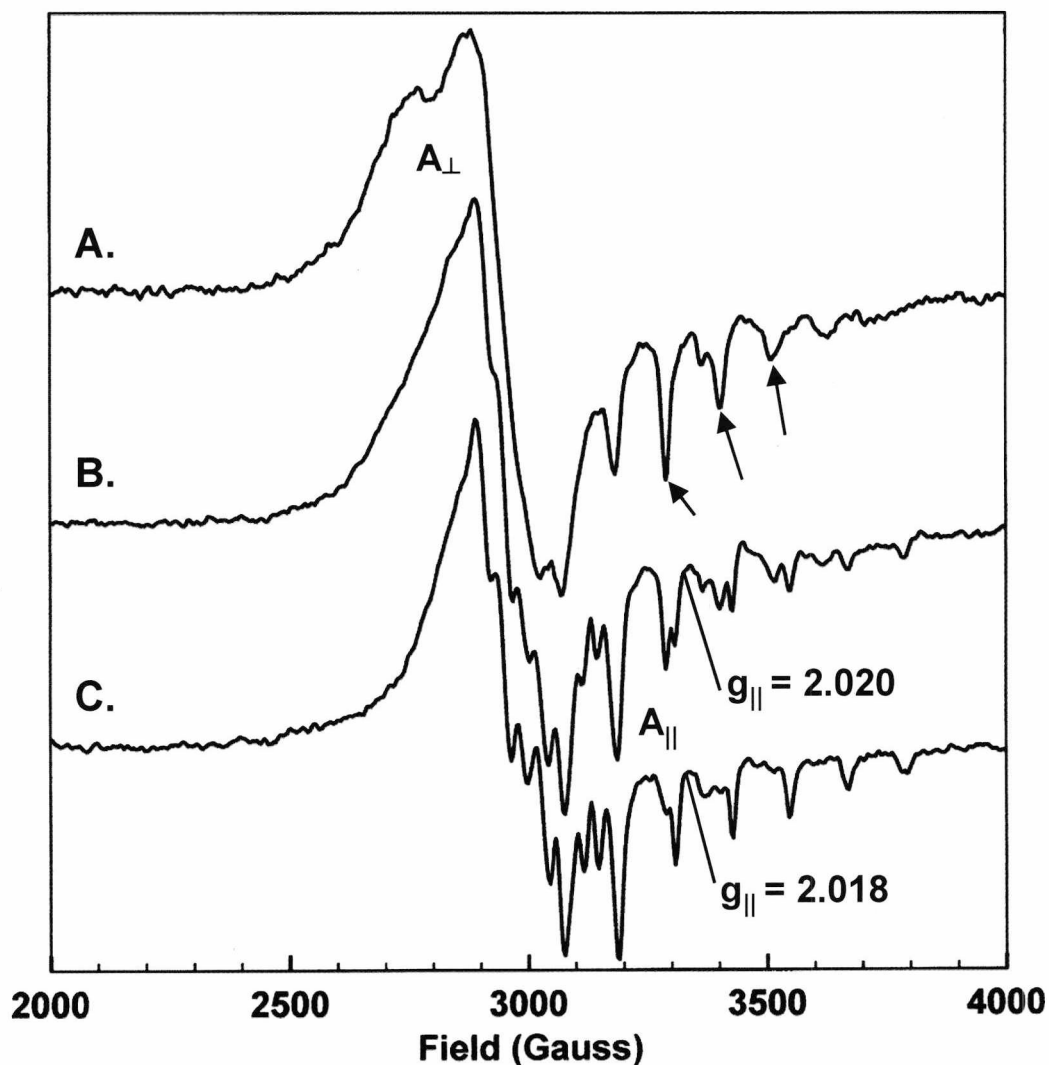


Figure 5.13. EPR time trace of cobalt(II)-factor III (100 μM) incubated with $\text{CbiH}_{60}^{\text{C-His}}$ (175 μM) in the presence of SAM (5 mM). Free cobalt(II)-factor II with SAM (A) is in a low-spin base-on cobalt(II) form, with contaminating imidazole bound as a lower axial ligand. Changes in both A_{\perp} and A_{\parallel} are observed with the addition of $\text{CbiH}_{60}^{\text{C-His}}$ (B). This could indicate a change in the orientation of the bound imidazole or show formation of a ternary complex. This process is essentially complete after 24 hr (C). Changes in the parallel hyperfine splitting are dominated by three components, indicated as black arrows in (A).

For the complete reaction that contains cobalt(II)-factor III, CbiH₆₀^{C-His}, SAM and DTT, a partial turnover with approximately 70-75% cobalt(II)-factor III and 20-25% cobalt-factor IV was observed. The initial mixing leads to small changes in the parallel features, as indicated in Figure 5.14. However, it seems that DTT slows this process, either through its interaction with cobalt(II)-factor III or the enzyme. After 24 hr, the partial turnover is observed as a visual colour change from green to brown. The EPR spectrum revealed no overall change in cobalt oxidation state, a feature that should be detected despite such a low yield of the reaction (Dr Steve Rigby, personal communication). In addition, the *g* perpendicular and *g* parallel features were almost identical to the same reaction lacking DTT. Essentially the decrease in $g_{||} = 2.020$ value to $g_{||} = 2.018$ is also present, suggesting the substrate or product is bound to the enzyme through cobalt ligation. Unfortunately, no signals were observed for the reduced [4Fe-4S]²⁺ cluster. This was surprising as it was anticipated that DTT could reduce the 4Fe-4S cluster. Thus, the function of the 4Fe-4S cluster in CbiH₆₀ still remains to be determined.

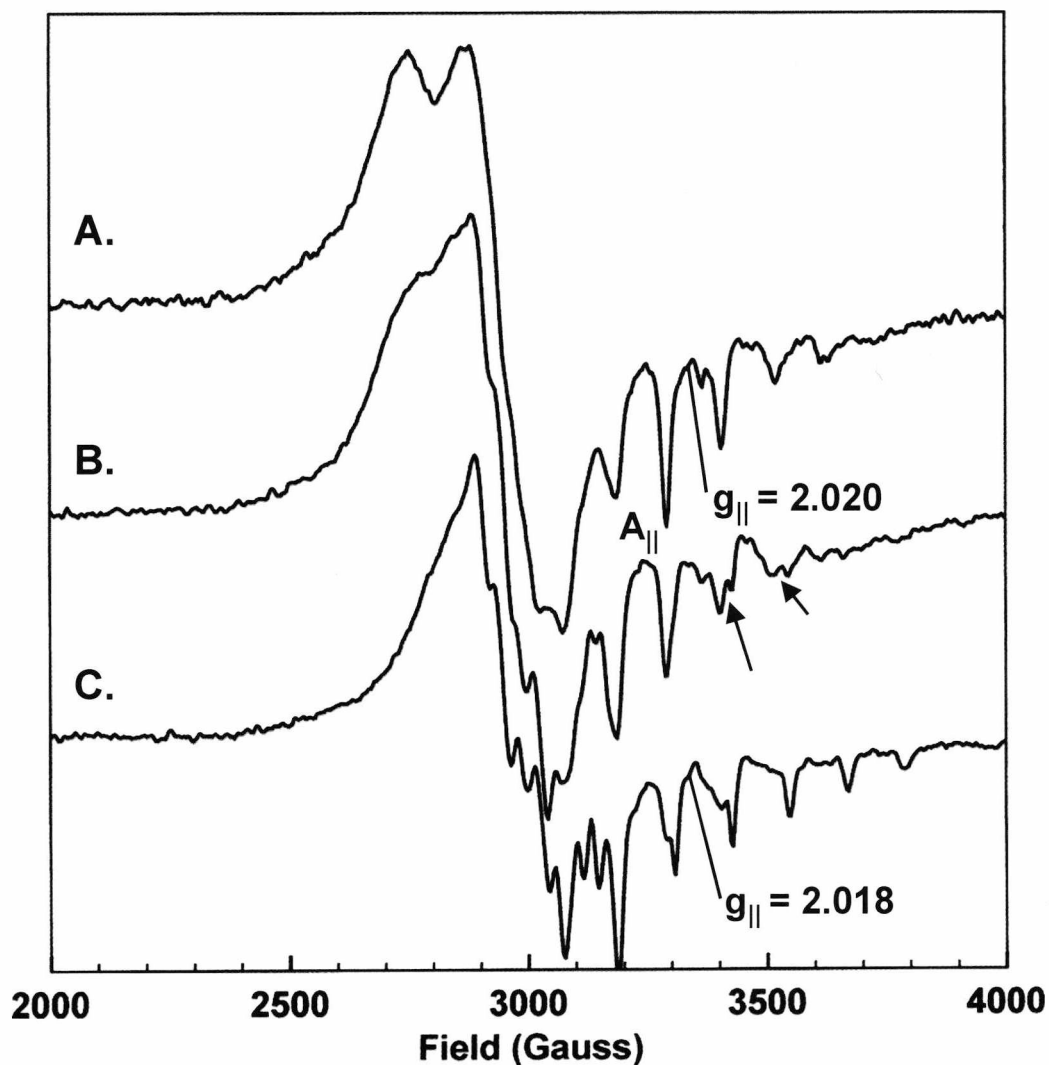


Figure 5.14. EPR time trace of cobalt(II)-factor III (100 μM) incubated with $\text{CbiH}_{60}^{\text{C-His}}$ (175 μM) in the presence of SAM (5 mM) and DTT (10 mM). Free cobalt(II)-factor II with SAM and DTT (A) is in a low-spin base-on cobalt(II) form, with contaminating imidazole bound as a lower axial ligand. With the addition of $\text{CbiH}_{60}^{\text{C-His}}$ (B) only slight changes are observed, indicating DTT inhibits the binding event. After 24 hr (C) the spectrum is essentially identical to Figure 5.13, with no change in cobalt oxidation state, whilst the cobalt ion is presumably bound to a ligand from the enzyme.

5.3 Discussion

The ring contraction step is a fascinating reaction that is unparalleled in Nature, yet its mechanism is yet to be fully resolved. The ring contraction step involves a C-17 methylation, removal of the methylene bridge between rings A-D and the formation of a lactone ring that requires the C-20 methyl group, although the order of these reactions is unknown. It seems the two pathways essentially differ in their dependence for oxygen. Whilst the anaerobic pathway is oxygen-independent, the aerobic pathway utilises molecular oxygen and incorporates a hydroxyl group at C-20 to form a tertiary alcohol (Debussche *et al.*, 1993; Schroeder *et al.*, 2009). This is catalysed by CobG with the reaction occurring subsequent to the ring contraction step catalysed by CobJ. The formation of a tertiary alcohol at C-20 has been shown to favour ring contraction using an organic chemistry based approach (Eschenmoser, 1988).

The limited knowledge of the anaerobic ring contraction step has created a “bottleneck” that limits further investigation into the pathway. To examine this further a novel anaerobic ring contraction enzyme CbiH₆₀ from *B. megaterium* DSM509 was investigated. After preliminary characterisation it was established CbiH₆₀ contains a reducible iron-sulphur centre, which was confirmed by EPR as a [4Fe-4S] type. Although not successfully demonstrated by redox potentiometry, it was suggested that the redox potential was high (~ -50 mV). Significantly CbiX^L also contains a [4Fe-4S] centre with a midpoint redox potential of $E^0 \approx -170 \pm 4$ mV (Leech *et al.*, 2003). A function for these redox groups has never been explained, although it is known that the aqua-Co(III) to Co(II) redox couple is around $E^0 \approx 170$ mV.

The activity of CbiH₆₀ was investigated using cobalt-factor III as a substrate, with the factor state chosen as it is known that this is the preferred substrate during the early stages of the anaerobic pathway (Frank *et al.*, 2007). The oxidised CbiH₆₀ [4Fe-4S] appeared to convert cobalt(II)-factor III to cobalt(III)-factor III, as shifts in the UV-Vis spectra were observed. Yet EPR experiments could not detect any change in cobalt oxidation state and suggested that the cobalt ion could be interacting with a ligand from CbiH₆₀. These two observations can be interpreted as

substrate binding, with EPR also suggesting that tetrapyrrole binding is ligand gated by SAM. As ligand binding to the cobalt is incomplete after initial mixing, the tetrapyrrole might initially bind in the active site through weak interactions with the propionate and acetate side chains. Then to secure the tetrapyrrole into the correct orientation the cobalt ion could act as anchor and bind to an unknown amino acid within the active site. Similar findings were also observed with the *M. thermoautotrophicus* CbiL (Frank *et al.*, 2007). Surprisingly the activation of CbiH₆₀^{C-His} requires the addition of the reducing agent DTT. Moreover, only the activity of the CbiH₆₀ could be demonstrated in its purified form, whilst other CbiH enzymes seemingly only show activity when present with an *E. coli* crude lysate. The C-terminus NiR/SiR domain is not essential to ring contraction (Raux *et al.*, 1998a). However, the activity *in vitro* can be studied further by modifying the conserved cysteine ligands for the [4Fe-4S] cluster. The C-terminus NiR/SiR domain may allow the reaction to proceed through an alternative route. This mechanism could adopt a more energetically favourable ES ternary complex and lower the activation energy. For other obligate anaerobic B₁₂ producing organisms these may utilise a non-specific oxidoreductase.

The enzymes belonging to the anaerobic pathway appear to be able to discretely recognise metal substrates and discriminate them from the metal-free intermediates of the aerobic pathway. This is despite remarkable sequence similarities between the enzymes shared between the two pathways. EPR of both CbiL and CbiH₆₀ suggests that the cobalt ion can form a bond with an enzyme ligand, with the formation of a five coordinate low-spin cobalt(II) ion being observed. However, for CbiH₆₀ it cannot be excluded that the contaminant imidazole may remain bound to the enzyme and thus a change in the parallel features of the cobalt ion could merely reflect altered orientation of the imidazole. The *M. thermoautotrophicus* CbiL substrate (cobalt-factor II) soaked crystals revealed only low occupancy and poor electron density for cobalt-factor II, with only a model to reveal the putative tetrapyrrole binding site (Frank *et al.*, 2007). Unfortunately, this model did not reveal any firm clues to a potential cobalt ligand, although it was suggested the structure may not be a true reflection of the active form of the enzyme. The *Clostridium tepidum* CbiL structure was also solved recently with SAH bound (Wadahama *et al.*, 2007). This publication used this structure and modelled the

docking of cobalt-factor II into its predicted active site. It was speculated that the conserved residues Met⁷¹ and Thr¹¹² may act as cobalt ligands, with Thr¹¹² in a more favourable position. These residues are not present in CobI sequences.

Interestingly, in close proximity to the *C. tepidum* Thr¹¹² is a conserved motif GDXXX(Y/W/F) found in all class III methyltransferases (Figure 5.15). The GD sequence is considered essential for SAM binding, but the aromatic group is close to the putative tetrapyrrole binding site, shown in Figure 5.16. Thr¹¹² is found two amino acids after this aromatic residue, but is only conserved for CbiL methyltransferases. For CbiH and CbiF the aromatic residue is strictly a tyrosine. However, this is also partially conserved in CbiL, CobI, CbiE and the N-terminus of CobL. Significantly, in CobJ this is replaced with a phenylalanine, whilst CobM contains a tryptophan in this position. It may act as a cobalt ligand, with the tetrapyrrole orientation perhaps different in other class III methyltransferases. Alternatively, these aromatic groups could be present for tetrapyrrole binding or participation in the methylation or ring contraction mechanism. In addition, the active site of CbiH is surrounded by four tyrosine residues, shown in Figure 5.16. These residues will be interesting to study further by mutagenesis.

It is considered that the ring contraction reaction is favoured by the prior formation of a tertiary alcohol that is observed in the aerobic pathway (Eschenmoser, 1988). However, the exact mechanism of ring contraction remains unknown. For the anaerobic pathway it was originally proposed that the cobalt is involved in an one or two electron redox reaction (Scott, 2003). However, results from this thesis suggest that the cobalt ion acts as a ligand to allow the enzyme to recognise the substrate. A threonine (Thr¹⁵⁹ *T. thermophilus*, Thr¹⁴⁴ *B. megaterium*) residue conserved only in the CbiH methyltransferases is found within the active site. Interestingly, this conserved residue is found only in CbiH and is associated with the predicted active site (Figure 5.16). In CobJ this is replaced with a conserved lysine, whilst in other class III methyltransferase it is not conserved. Mutagenesis could be employed to investigate the role of this residue. By altering these residues this could modify the activity and specificity of the enzyme. This is an important concept in not only understanding the mechanism of ring contraction but as an approach in synthetic biology to modify existing metabolic pathways to create novel compounds.

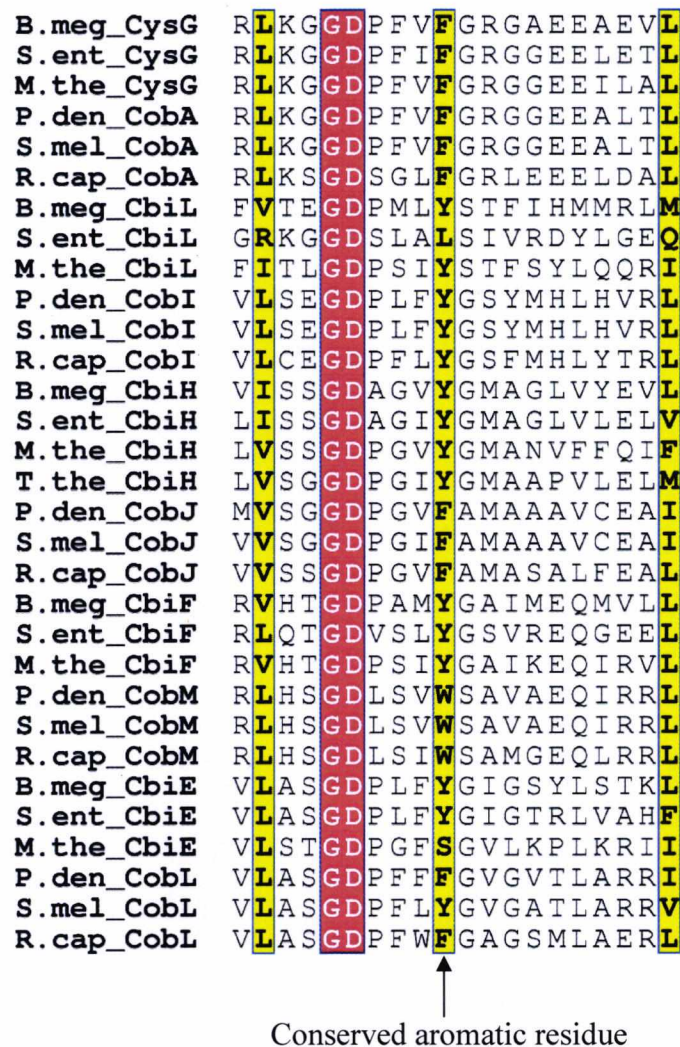
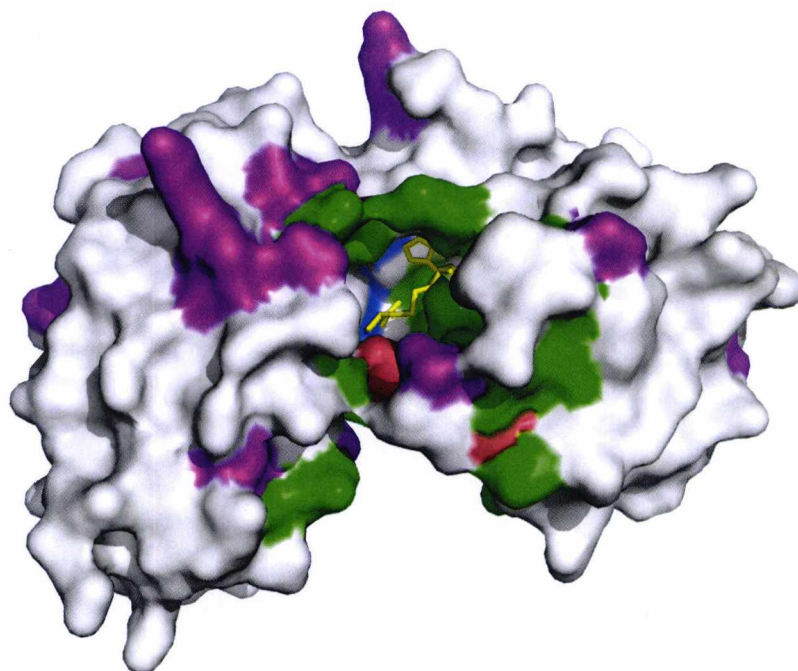


Figure 5.15. Alignment of a short sequence domain conserved in class III methyltransferase activity. A short mid-section sequence of CysG, CysG^A, CobA, CbiL, CobI, CbiH, CobJ, CbiF, CobM, CbiE and CobL (N-terminus) were aligned with ClustalW2 and graphically visualised on ESPrnt 2.2. Species abbreviations B.meg (*B. megaterium*), S.ent (*S. enterica*), M.the (*M. thermoautophylicus*), T.the (*T. thermophilus*), P.den (*P. denitrificans*), S.mel (*S. meliloti*), R.cap (*R. capsulatas*).

A.



B.

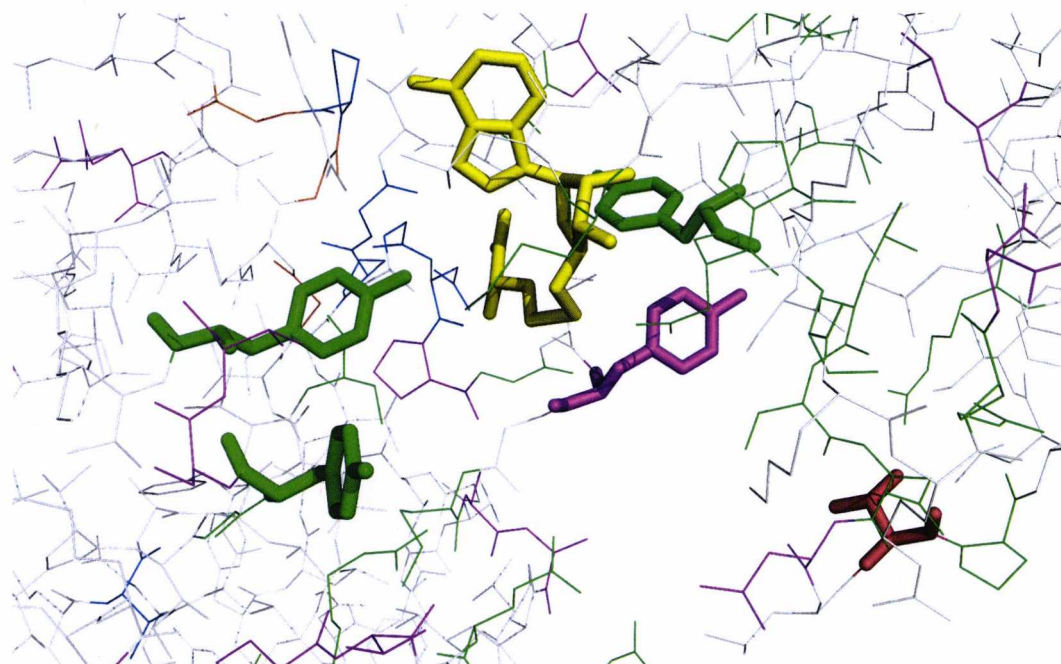


Figure 5.16. Regions of conservation between CbiH and CobJ using the structure of the *Thermus thermophilus* Hb8 CbiH (Shimizu, K and Kunishima, N., Unpublished data) as a model. A. Surface contour view of CbiH (PDB_2ZVB) with the putative tetrapyrrole binding site inside a cleft with SAH bound (yellow). Conserved residues are coloured accordingly. Conserved in all class III methyltransferases (blue), GXGXG motif with glycines highlighted (orange), conserved in only CbiH and CobJ (green), variant residues between CbiH and CobJ (magenta). **B.** Putative active site with SAH (yellow) bound. Surrounding tyrosine residues, shown in bold. In addition a single Thr¹⁵⁹ (red) residue conserved only in CbiH is shown in bold.

Chapter 6

The Biosynthesis of the Corrin Macrocycle in *B. megaterium*

6.1 Introduction

Two vitamin B₁₂ biosynthetic pathways exist, cryptically referred to as the aerobic (oxygen-dependent) and anaerobic (oxygen-independent), with the latter most likely evolving the earliest when the Earth's atmosphere was devoid of molecular oxygen (Raux, 1999). Starting at uroporphyrinogen III, the two pathways are broadly similar and involve eight methylations, cobalt insertion, ring contraction, formation and extrusion of a lactone ring, corrin ring reduction, methyl migration, decarboxylation, cobalt reduction, adenosylation and six amidations of the side chain carboxylate groups. The two pathways differ by two major events which involves the mechanism of ring contraction and the timing of cobalt insertion. Essentially, in the aerobic pathway cobalt is inserted late into a contracted corrin ring by an ATP-dependent cobaltochelataase, which is a multi-enzyme complex of CobN, CobS and CobT (Heldt *et al.*, 2005). In comparison the anaerobic pathway utilises a single enzyme (CbiX or CbiK) to insert cobalt at an early stage in the pathway into a pyrrolic ring system. Then to enable tighter coordination of the cobalt ion, the ring is contracted to the corrin ring by the enzyme CbiH. This enzyme also has a dual function as a methyltransferase and essentially performs a similar function as CobG and CobJ from the aerobic pathway (Santander *et al.*, 1997). The mechanism of ring contraction is unknown, but does not require molecular oxygen and forms a δ -lactone ring by using the acetate side chain at C-2 (please see Figure 1.3, for diagram of carbon assignment numbers). This differs in comparison to the ring contraction event in the aerobic pathway, which is an oxygen-dependent step used for the formation of a γ -lactone ring (Schroeder *et al.*, 2009).

The aerobic pathway is the better characterised of the two pathways, largely due to research with *P. denitrificans* (Cameron, 1989; Debussche *et al.*, 1993). In contrast, the anaerobic pathway has been characterised in detail only before the event of ring contraction. For *B. megaterium* precorrin-2 is first oxidised by SirC the NAD⁺ dependent dehydrogenase, by forming a double bond between C-14 and C-15. This produces the intermediate factor II which is the preferred substrate for insertion of cobalt by CbiX^L, producing cobalt-factor II (Leech *et al.*, 2003). However in other anaerobes that lack the SirC enzyme, these organisms generally possess a different

cobaltochelatase, CbiK, instead of CbiX^L. The next step involves a methylation at C-20 by CbiL, which is a prerequisite for ring contraction. This enzyme has been studied from *S. enterica* and *M. thermoautotrophicus*, and has been shown to prefer the substrate cobalt-factor II rather than cobalt-precorrin-2 (Frank *et al.*, 2007). Also there seems to be no preference for the oxidation state of the cobalt ion as either Co(III) or Co(II) (Frank *et al.*, 2007). Furthermore the enzyme cannot methylate precorrin-2, but will methylate cobalt-precorrin-2, cobalt-factor II and other metal-precorrin-2, metal-factor II substrates (Spencer *et al.*, 1998). This is also confirmed as a *S. enterica* *cbiL* mutant cannot be complemented by the *cobI* gene. This information suggests that although a number of enzymes share similarities between the two pathways, the anaerobic enzymes are highly specific for cobalt-containing substrates.

After ring contraction (Chapter 5), the next steps remained shrouded in mystery, with only a few intermediates revealed in trace amounts. This includes research supporting the role of CbiF to methylate at C-11, producing cobalt-precorrin-5A (Kajiwara *et al.*, 2006). In addition the CbiG was shown to remove the δ -lactone ring, releasing it as a C-2 unit acetaldehyde, producing cobalt-precorrin-5B. However, both cobalt-precorrin-5A and cobalt-precorrin-5B were only isolated in limited quantities. In addition, these were extracted by methyl esterification. To obtain these intermediates this approach utilised the minor quantities of cobalt-precorrin-4 (or cobalt-factor IV) produced by the *S. enterica* CbiH in crude lysates of *E. coli*. These lysates were coupled with *E. coli* lysates over-expressing CbiF, or CbiF and CbiG together. After incubation with cobalt, precorrin-3 and SAM, the tetrapyrroles were esterified before purification by TLC. These intermediates were then characterised by NMR. This revealed that the CbiF would methylate at C-11, whilst the δ -lactone ring was removed if CbiG was present. The full UV-Vis spectral data was not published, but cobalt-precorrin-5A and cobalt-precorrin-5B were described as brown and yellow in colour, respectively. In addition supplementary data provided absorbance maximums of the esterified compounds (Kajiwara *et al.*, 2006).

For the remaining steps, none of the intermediates between cobalt-precorrin-5B and cobyrinic acid have been isolated to date. In comparison to the aerobic pathway it

would appear that the C-1 methylation is the next proposed reaction to occur. No enzyme in the anaerobic pathway shares any identity with the analogous enzyme in the aerobic pathway, which is catalysed by CobF. CbiD is the predicted anaerobic C-1 methyltransferase (Roessner *et al.*, 2005). However, the sequence and structure (*A. fulgidus* PDB_1SR8) of CbiD, shares no similarity to any characterised enzyme. Although the intermediate cobalt-precorrin-6A has yet to be isolated, the CbiD was predicted to be the C-1 methyltransferase, using strains of *E. coli* expressing all the Cbi enzymes from *S. enterica* to produce cobyrinic acid. This work showed a deletion of *cbiD* resulted in production of a 1-desmethyl-cobyrinic acid *a,c* diamide intermediate lacking the C-1 methylation (Roessner *et al.*, 2005). Interestingly, the C-1 methylation would also be absent if the proteins CbiA or CbiP were missing. However even in the presence of CbiD, CbiA and CbiP, a significant amount of 1-desmethyl-cobyrinic acid *a,c*-diamide still accumulated. This evidence remains unclear, considering the amidases CbiA and CbiP are involved at a later step in the pathway. Evidence from these experiments supported the role for CbiD as the C-1 methyltransferase, but also suggested that its activity was dependent on the presence of both CbiA and CbiP. This might also suggest that the whole pathway interacts as a functional metabolon, but as of now, no protein-protein interactions have been detected. If this was to be true, then this might explain the difficulty in studying individual reactions.

This part of the project aims to further characterise the pathway and isolate intermediates that have never previously been observed. In Chapter 5, this identified a method to produce high quantities of cobalt-factor IV. This provides a reliable source of substrate to enable a further investigation into the anaerobic pathway. An outline of the steps investigated is shown in Figure 6.1. Following the success with CbiH₆₀, the remaining enzymes of the *B. megaterium* DSM509 anaerobic pathway from Chapter 4 were investigated. The purified enzymes were incubated with substrates in order to characterise their activity and selectivity. This characterisation involved a combination of UV-Vis spectroscopy, LC-MS and MS/MS.

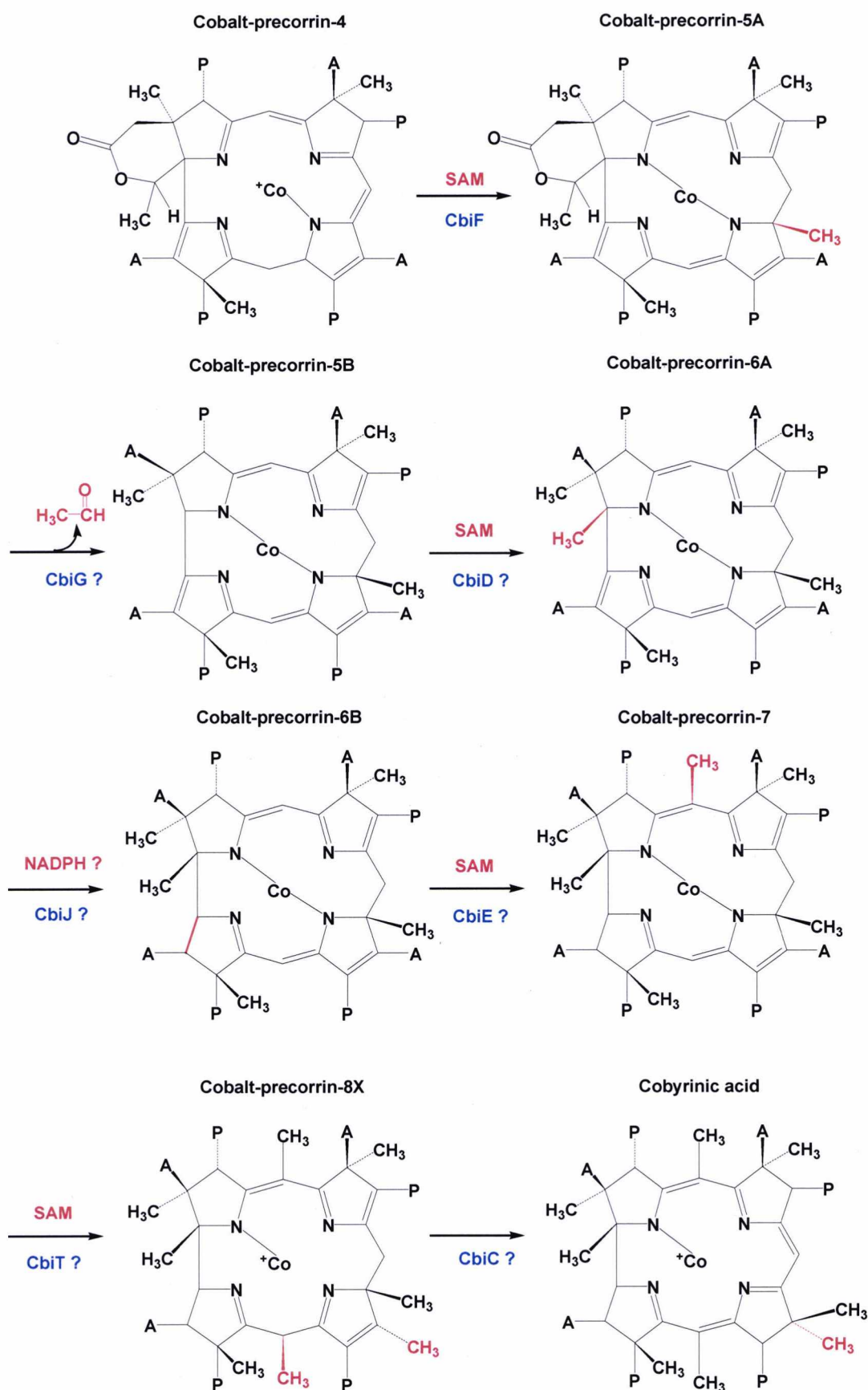


Figure 6.1. The missing steps of the anaerobic pathway. From cobalt-precorrin-6A to cobalt-precorrin-8X, none of the steps have been characterised, nor have the intermediates been isolated. Each step is highlighted in red.

6.2 Results

6.2.1 Isolation of cobalt-precorrin-5

This work aims to isolate and characterise true biosynthetic pathway intermediates. To begin with, different combinations of pure enzymes (CbiH₆₀, CbiF and CbiG) along with the substrate cobalt-factor III and cofactors SAM and DTT were co-incubated, summarised in Figure 6.2. This approach was used to verify the specificity and order of the reactions as outlined by Ian Scott's research (Kajiwara *et al.*, 2006). After incubation at 37 °C for 24 hrs, the reactions were analysed by UV-Vis spectroscopy before further analysis by LC-MS. Firstly, in the presence of CbiH₆₀ and cofactors (SAM and DTT), the starting cobalt-factor III substrate was observed to have changed to a brown colouration, with a UV-Vis spectrum consistent of a mixture of cobalt-factor III and cobalt-factor IV. In contrast if CbiF or CbiG were added separately or together to cobalt-factor III with cofactors, no colour change was observed. In fact CbiG was seen to precipitate leaving a cloudy green solution. However, if CbiF was added in combination with CbiH₆₀ and cofactors, a new yellow intermediate with novel UV-Vis spectral features was observed. Instead of a broad shoulder around 350 nm, this was replaced with a clear absorbance maximum at 356 nm, with double absorbance peaks around 470-500 nm (Figure 6.3). Although not identical, this was more similar to the spectra of cobalt-precorrin-5A and cobalt-precorrin-5B, previously described (please see Figure 6.3). The addition of CbiG to this new intermediate did not alter the visual appearance or UV-Vis spectrum. Therefore, the function and substrate specificity of CbiG could not be determined in this work. It should also be considered that these reactions were highly repeatable. To improve substrate turnover, the activity of the enzymes could be increased if *B. megaterium* DSM319 was induced for a shorter period (16 hrs, instead of 24 hrs).

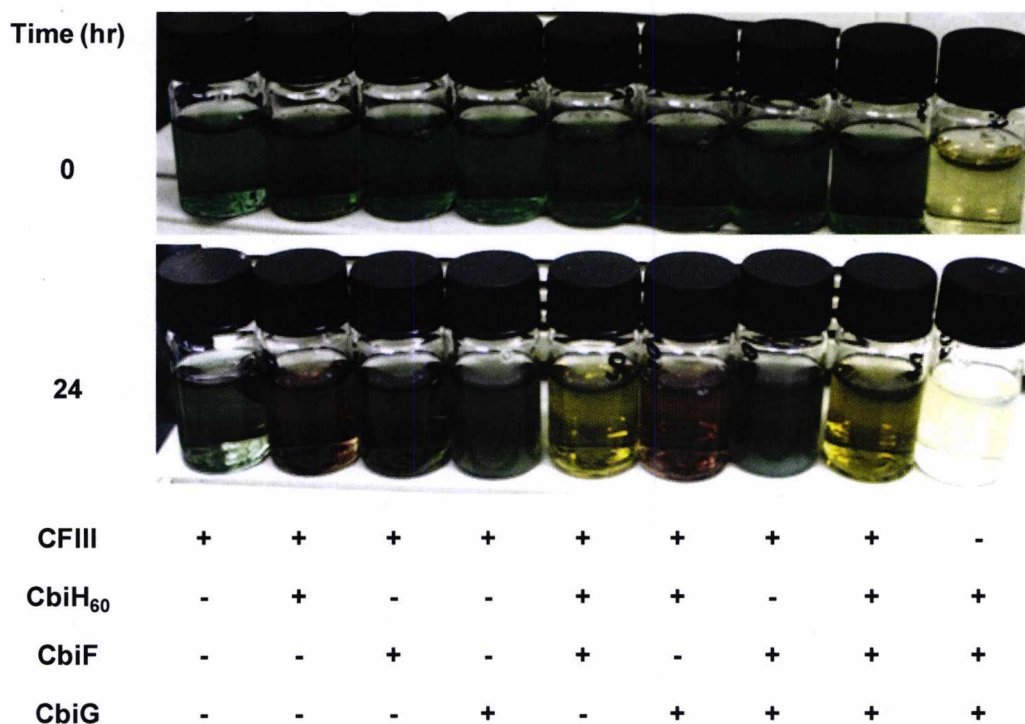


Figure 6.2. The visual appearance of reactions containing different combinations of CbiH₆₀, CbiF, CbiG with 30 μ M cobalt-factor III (CFIII), 1 mM SAM and 10 mM DTT. Top panel indicates appearance at beginning of incubation with bottom panel after 24 hrs at 37 °C. Table indicates presence of cobalt-factor III and enzymes. Note at beginning of reaction cobalt(III)-factor III is present as a bright green intermediate. Then after incubation with DTT this reduces the cobalt ion to cobalt(II)-factor III which is darker in colouration. The incubation containing CbiG on its own or in the presence of CbiF, shows a degree of precipitation. However if CbiG is incubated in the presence of CbiH₆₀ this is prevented.

This series of reactions demonstrated that CbiF produces an intermediate from cobalt-factor IV, presumably cobalt-precorrin-5A. In addition the UV-Vis spectra also indicated that some cobalt-factor II/III was still present. To obtain pure cobalt-precorrin-5 the majority of cobalt-factor II/III could be removed by heating at 80 °C for 30 min and the precipitate removed. The intermediate was subsequently bound to an anion-exchange column (see inset, Figure 6.3) revealing the presence of an orange pigment. After elution from the column the UV-Vis spectrum of the intermediate was taken, which revealed defined absorbance peaks at 356 (100), 466 (41) and 483 (39) nm (Figure 6.3). It appears that the original reaction producing cobalt-precorrin-5 appears yellow because it is mixture containing the green cobalt-factor II/III. Time limited further studies on the activity of CbiG. Therefore in all subsequent incubations, CbiH₆₀, CbiF and CbiG were incubated in combination

with cobalt-factor III to produce the intermediate cobalt-precorrin-5. For further details of the synthesis and handling of cobalt-precorrin-5, please see Appendix V.

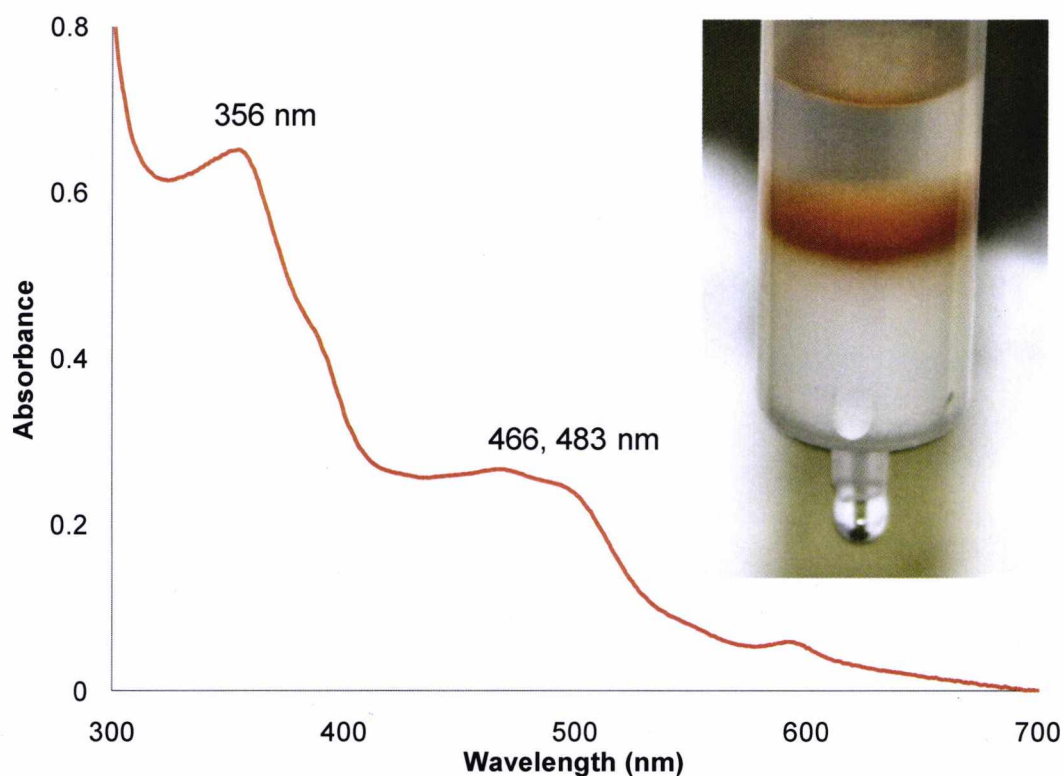


Figure 6.3. UV-Vis spectrum and visual appearance of new intermediate produced from cobalt-factor III, CbiH₆₀, CbiF, SAM and DTT. Inset picture showing orange intermediate concentrated using anion-exchange chromatography. The UV-Vis spectra of cobalt-precorrin-5A octamethylester [UV-Vis λ_{max} of 346 (100), 440 (58), 758 (8), 794 (8) nm] and cobalt-precorrin-5B octamethylester [UV-Vis λ_{max} of 312 (100), 379 (83), 483 (32), 595 (17), 885 (12) nm] were described previously (Kajiwara *et al.*, 2006).

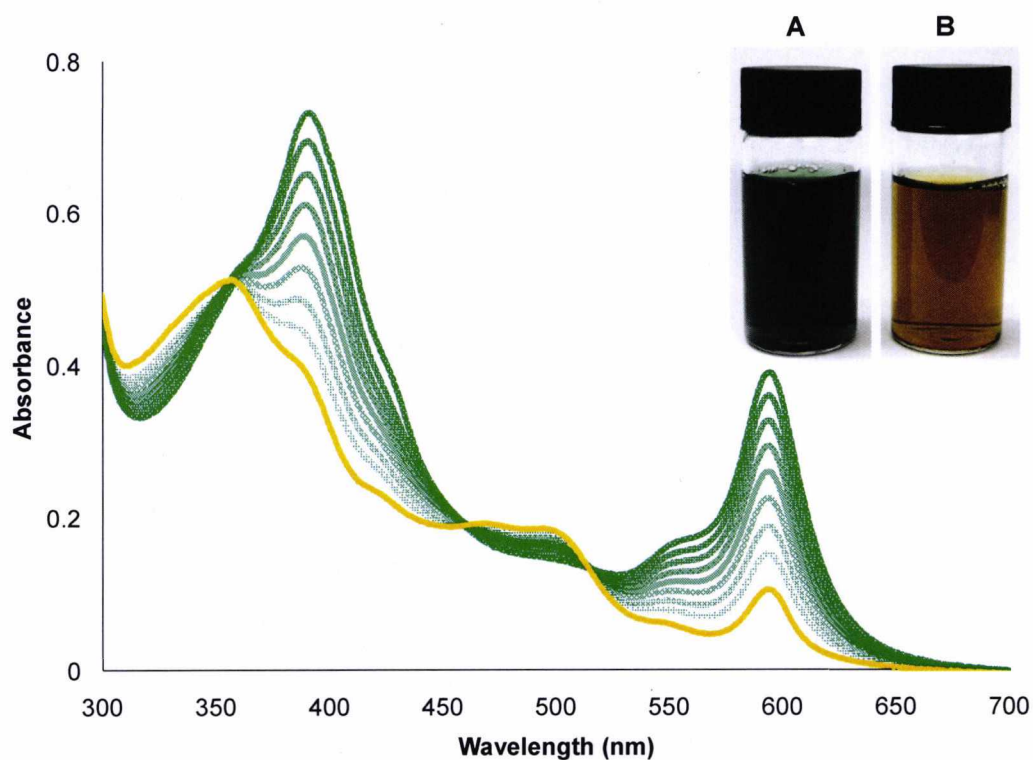


Figure 6.4. UV-Vis scanning spectra of 5 μM cobalt(II)-factor III incubated with CbiH₆₀, CbiF, CbiG, 1 mM SAM and 10 mM DTT. Cobalt-factor III (dark green line) is converted gradually (30 min scans) into this new intermediate (orange line). Inset picture shows 50 μM cobalt-factor III with 50 μM CbiH₆₀, 50 μM CbiF, 50 μM CbiG, 1 mM SAM and 10 mM DTT before incubation (A) and after 24 hrs incubation 37 °C (B). The presence of the peak at 595 nm indicates that a significant amount of cobalt-factor III remains unconverted.

It was assumed that the intermediate produced with CbiH₆₀ and CbiF was cobalt-precorrin-5A. The addition of CbiG would likely convert this into cobalt-precorrin-5B after opening of the δ -lactone ring. For routine LC-MS analysis, proteins were removed by adding acetic or hydrochloric acid (1% v/v), followed by centrifugation to clarify the samples. After an initial separation with an RP-18 column the eluted cobalt-precorrin-5 was found to be insoluble in water. However, this could be avoided if the intermediate was eluted in 40% (v/v) ethanol in water, instead of using 50% (v/v) ethanol in 1% (v/v) acetic acid. After this problem was corrected, samples were prepared for LC-MS analysis (Figure 6.5). Intermediates were separated using Method A, with 0.1% TFA and acetonitrile. However, cobalt-precorrin-5 was difficult to analyse by LC-MS, since it appeared as a series of breakdown products. In addition the UV-Vis spectrum shifted to produce two prominent peaks at 320 and 404 nm, along with peaks at m/z 934.24 and 920.24.

These masses were analysed by MS/MS which revealed that the more dominant ion at m/z 920.24 was derived from the m/z 934.24 species. As LC-MS did not yield conclusive results, only further confirmation by NMR will achieve this. However this was not within the scope of this thesis, so for now it was more intriguing to investigate the rest of the pathway.

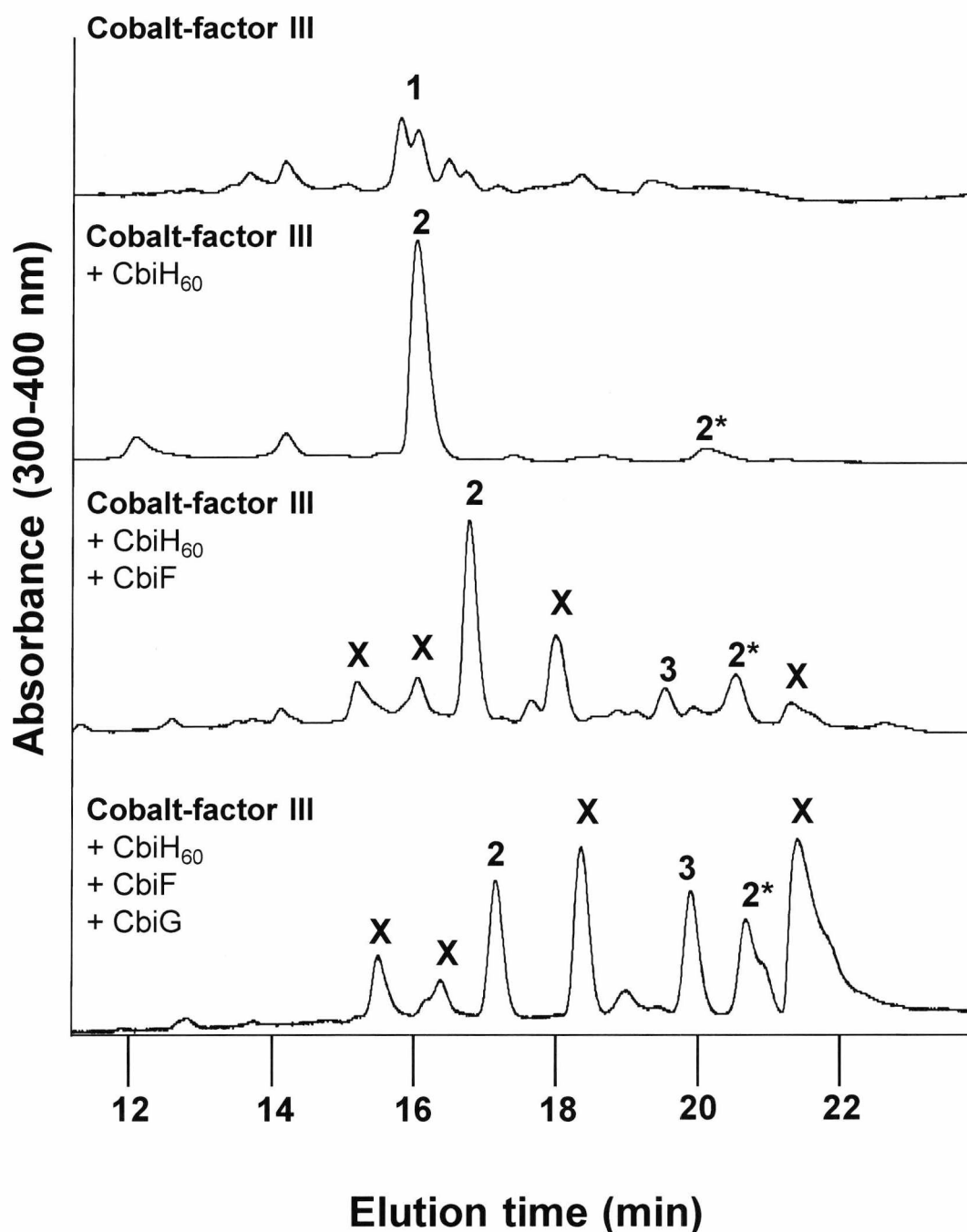


Figure 6.5. LC-MS UV-Vis chromatogram of CbiH₆₀, CbiF, CbiG incubations with cobalt-factor III, SAM and DTT. Separation of tetrapyrroles was achieved using Method A, with 0.1% TFA and acetonitrile. A combination of UV-Vis, MS and MS/MS was used to identify each compound. Labels are as follows, **(1)** cobalt-factor III (m/z 931.19), **(2)** cobalt-factor IV (m/z 948.23), **(2*)** oxidised cobalt-factor IV (m/z 946.21), **(3)** oxidised cobalt-precorrin-5, with ion at m/z 934.24 and the more abundant peak at m/z 920.24, **(X)** unknown peaks of m/z 966.20 and 950.20 (15.5-16.5 min) and m/z less than 900 (18.5 and 21.5 min).

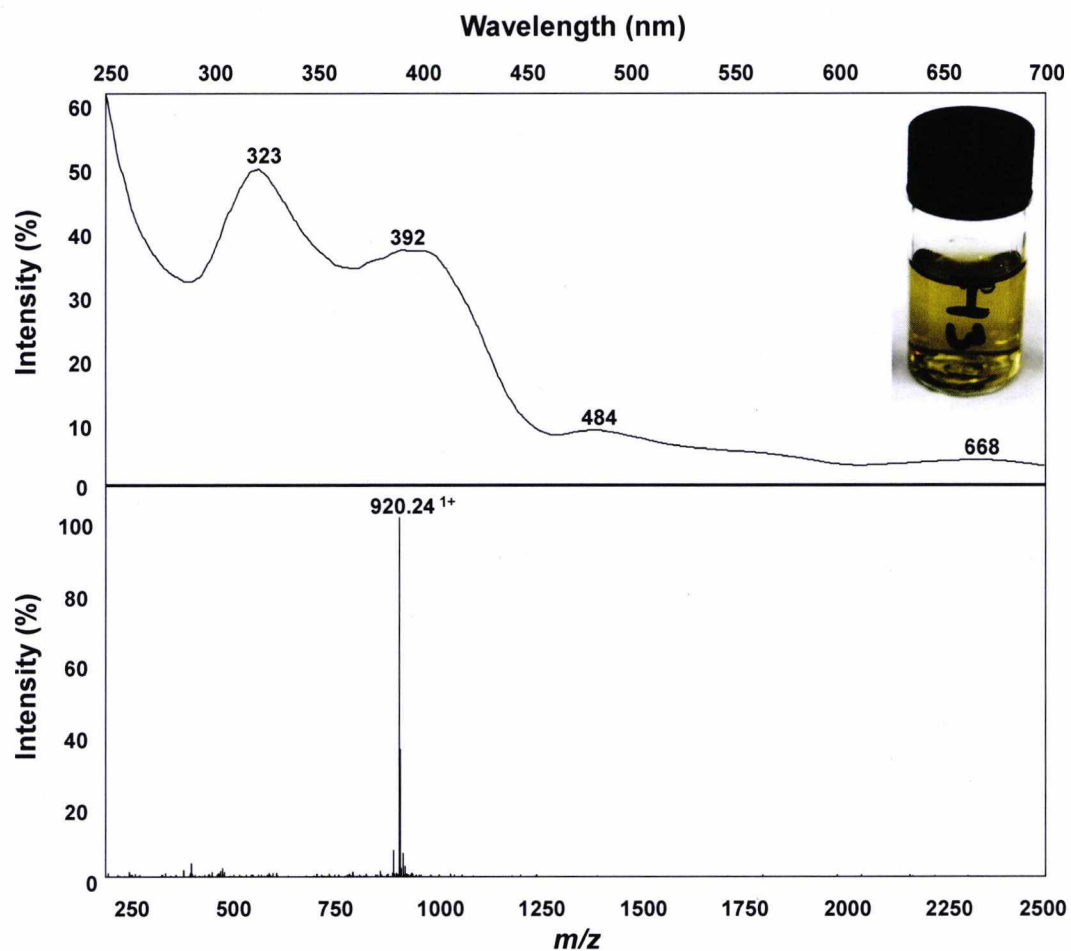


Figure 6.6. UV-Vis and MS of oxidised cobalt-precorrin-5 eluting at 19.7 min. UV-Vis with inset picture of intermediate in anaerobic conditions and MS spectrum.

6.2.3 Isolation of cobalt-precorrin-6A

Previous data suggested that CbiD may require the presence of other enzymes (CbiA and CbiP) for activity and it seemed unlikely that the *B. megaterium* DSM509 CbiD would catalyse any methylation reaction with cobalt-precorrin-5B. However, it was noted that the UV-Vis spectrum of cobalt-precorrin-5B shifted dramatically in the presence of CbiD and SAM. The broad absorbance at 460-510 nm was displaced by a new peak at approximately 420 nm. In addition the main absorbance peak for cobalt-precorrin-5B at 356 nm showed a shift to 325 nm. This spectral change was also shown to be dependent on the presence of SAM and CbiD. For further details of the synthesis and handling of cobalt-precorrin-6A, please see Appendix V.

The UV-Vis spectral changes that were observed were also concurrent with a subtle colour change from orange to a bronze coloured solution (Figure 6.7). This new intermediate and the control incubations (without SAM or enzyme) were analysed by LC-MS. Using Method A, the same peaks previously obtained with analysis of cobalt-precorrin-5B (m/z 934.24 and 920.24) were obtained for the substrate only, substrate plus enzyme and substrate plus SAM controls. However, remarkably for the incubation containing cobalt-precorrin-5B, CbiD and SAM a new peak was obtained which eluted at 17.5 min. This new peak had an m/z 951.29, with a predicted molecular formula of $(C_{44}H_{50}CoN_4O_{16}^+)$. Based on the proposed structure of cobalt-precorrin-6A (952.29 Da), it seems likely that this compound is indeed the next intermediate in the pathway. Interestingly, the LC UV-Vis spectrum of cobalt-precorrin-6A is not dissimilar to the spectra obtained under anaerobic conditions (Figure 6.8). This is also reflected by the mass obtained, suggesting that this is the first intermediate that is relatively stable even under aerobic conditions.

To transform the putative cobalt-precorrin-6A to the next known intermediate, cobyric acid (a characterised intermediate with experimentally confirmed extinction coefficients), a reduction of the corrin ring, methylation at C-5 and C-15, decarboxylation at the *e* acetate side chain and the unusual 1,5-sigmatropic methyl group rearrangement from C-11 to C-12 are required.

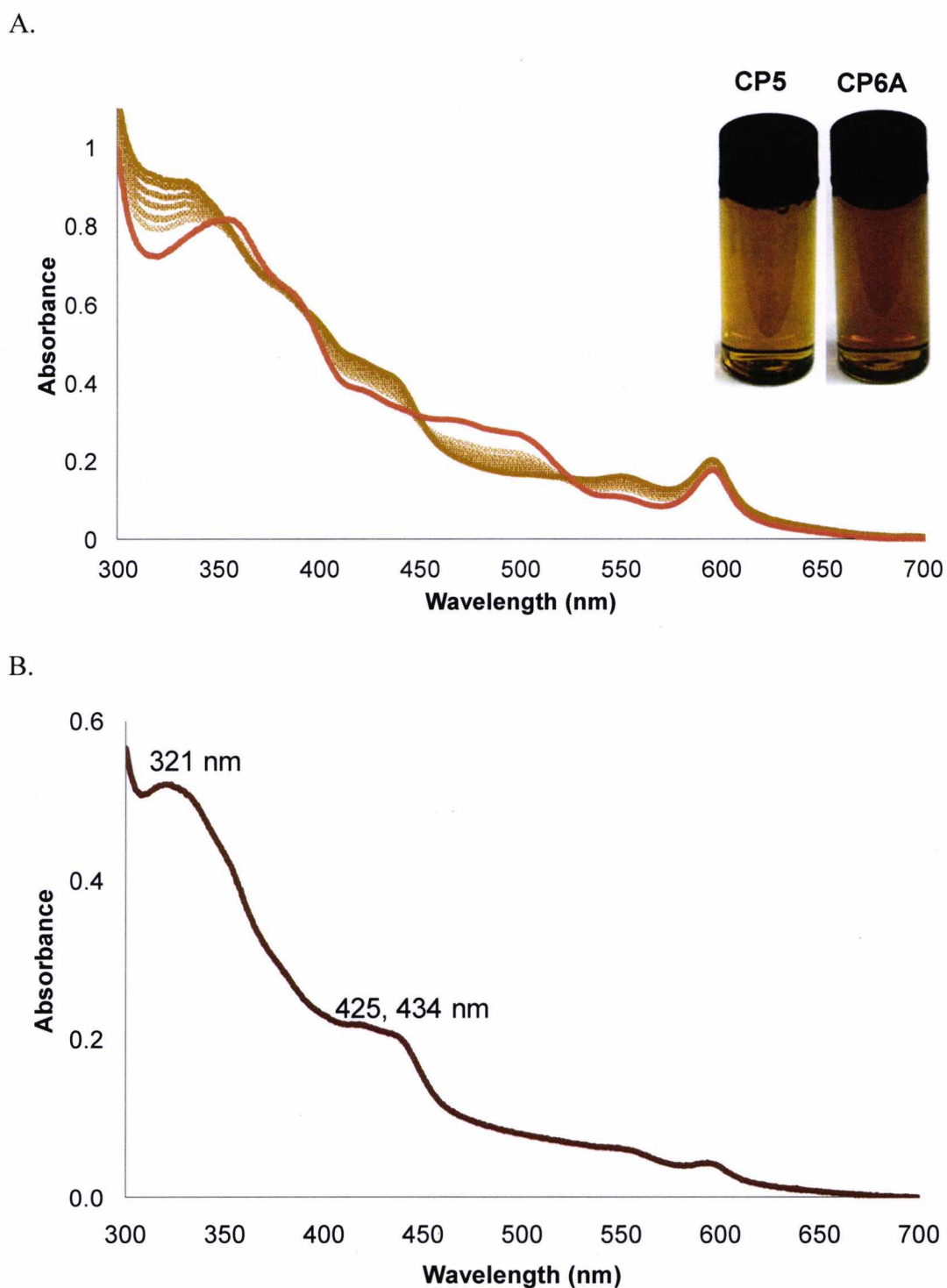


Figure 6.7. UV-Vis spectra of cobalt-precorrin-5 incubated with CbiD and SAM. **A.** Reaction at 37 °C with scans every 10 min. Inset picture of cobalt-precorrin-5 (CP5) and cobalt-precorrin-6A (CP6A). **B.** Cobalt-precorrin-6A after purification by reverse-phase chromatography, confirmed by LC-MS.

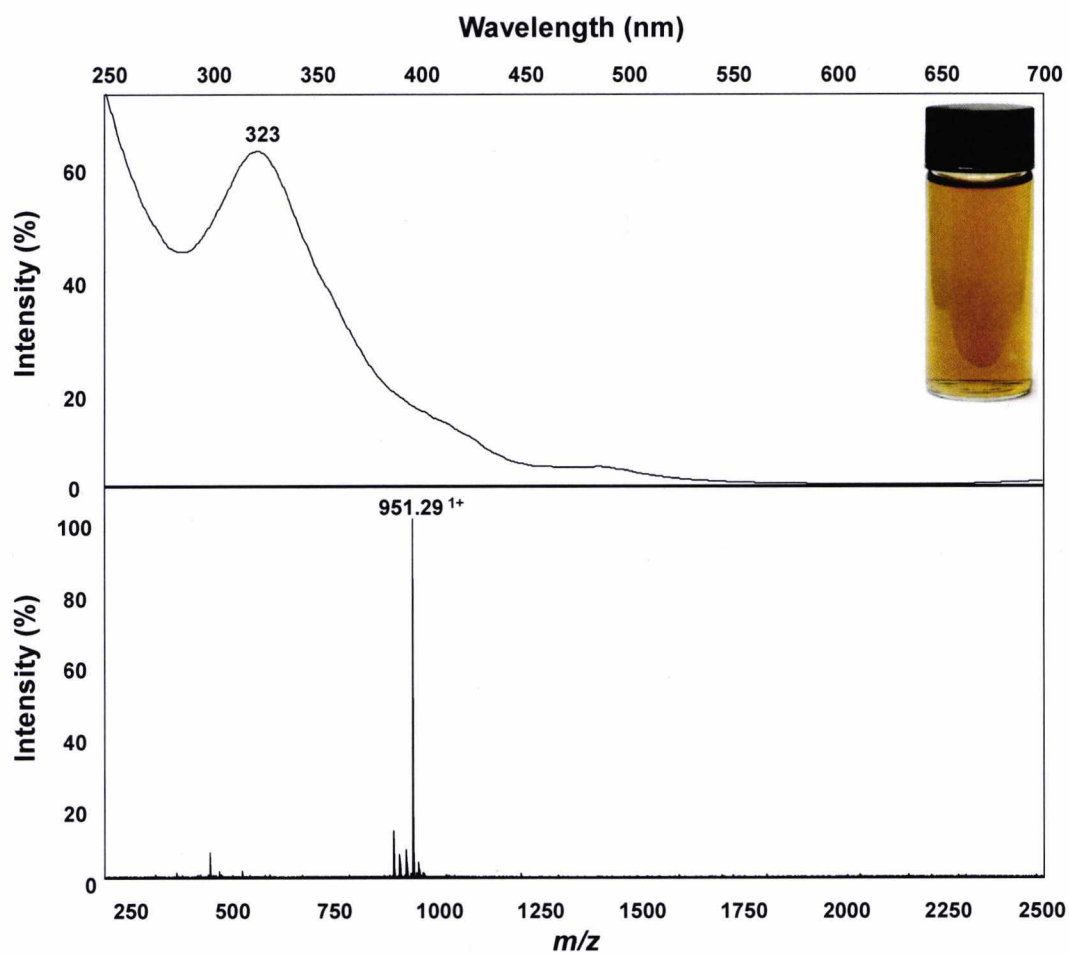


Figure 6.8. UV-Vis and MS of cobalt-precorrin-6A eluting at 17.4 min. UV-Vis with inset picture of intermediate in anaerobic conditions and MS spectrum.

6.2.4 The characterisation of the corrin ring reductase CbiJ

The orthologue of CobK in the anaerobic pathway is CbiJ. Using this past knowledge an assumption was made that CbiJ was the most likely enzyme to catalyse the reduction of cobalt-precorrin-6A, with NADPH as the likely reducing agent. So with newly synthesised cobalt-precorrin-6A the activity of CbiJ was investigated.

Initially the intermediate cobalt-precorrin-6A was incubated with and without CbiJ (estimated as close to equimolar concentrations) and the UV-Vis spectra was monitored for any shifts that might indicate enzyme binding, but nothing was observed. However, the addition of CbiJ (10 μ M) along with NADH (25 μ M) revealed a slight but repeatable spectral shift. The absorbance peak at 322 nm shifted to 318 nm (Figure 6.9). In addition the peak at 418 nm increased its absorbance slightly. Visually it was noticed that the bronze coloured cobalt-precorrin-6A would appear as a brighter yellow pigment. These changes were not observed if this reaction was repeated in the presence of NADPH (25 μ M) and CbiJ. This indicates that the reaction is dependent on the presence of CbiJ and NADH. The rate of reaction was comparatively fast. Lowering the concentration of enzyme reduced the rate of reaction. An absolute rate cannot be given as the concentration of cobalt-precorrin-6A was not known.

To confirm this intermediate as dihydro-cobalt-precorrin-6 or cobalt-precorrin-6B, further confirmation by LC-MS was required. However, there appeared to be difficulties with preparing this intermediate for LC-MS analysis. After acidification with 1% acetic acid (v/v) and purification by reverse phase chromatography, the UV-Vis spectrum that was observed in these reactions reverted back to the original spectrum of cobalt-precorrin-6A (Figure 6.9). This could also be observed visually as the addition of acid to this intermediate led to a bleaching effect with a bronze coloured solution leftover. This could be avoided if the mixture was simply heated (80 °C, 30 minutes) and purified by anion-exchange chromatography. Thus, cobalt-precorrin-6B may be sensitive to oxidation making it difficult to detect under aerobic conditions. The exact nature of cobalt-precorrin-6B may therefore have to await NMR determination.

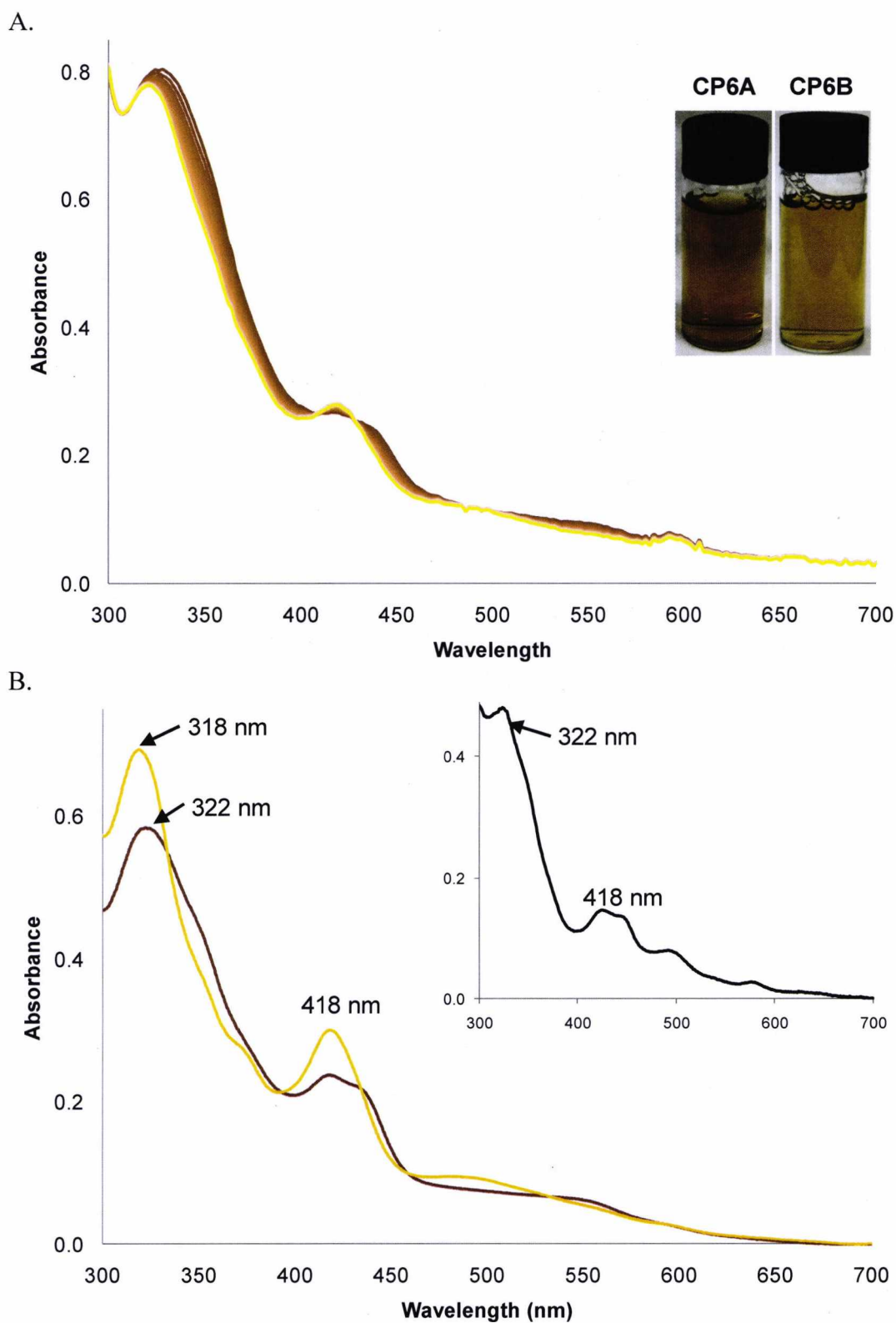


Figure 6.9. UV-Vis of cobalt-precorrin-6A incubated with CbiJ and NADH. A. UV-Vis scans every 2 sec, with inset picture of concentrated intermediates. **B.** UV-Vis spectra after anion-exchange chromatography, cobalt-precorrin-6A (dark brown line) and cobalt-precorrin-6B (yellow line). Inset graph shows cobalt-precorrin-6B after acidification and reverse phase chromatography. Acidification causes this intermediate to revert back to cobalt-precorrin-6A, which was confirmed by LC-MS.

6.2.5 Biosynthesis of cobyrinic acid from cobalt-precorrin-6A

Along with the two electron reduction of the C-C double bond between C-18 and C-19, a further two methylations at C-5 and C-15, decarboxylation of the ring C acetate side chain and methyl group migration from C-11 to C-12 are required, for the transformation of cobalt-precorrin-6A into cobyrinic acid. In order to look at the individual reactions in the transformation of cobalt-precorrin-6B into precorrin-8X the CbiE and CbiT enzymes from *M. thermoautotrophicus* were studied and compared to the incubations with the *B. megaterium* CbiET fusion protein.

6.2.6 Isolation of cobalt-precorrin-7

This first set of incubations began with incubating the presumed cobalt-precorrin-5A with *B. megaterium* DSM509 CbiG, *B. megaterium* DSM509 CbiD, *B. megaterium* DSM509 CbiJ, *M. thermoautotrophicus* CbiE and *M. thermoautotrophicus* CbiT. All enzymes were individually purified and then added at approximately equal concentrations, with respect to the substrate. A set of reactions were prepared based on the predicted pathway, with each individual enzyme added as a set of progressive reactions as shown in Figure 6.10. All incubations contained the cofactors NADH (50 μ M) and SAM (1 mM), with the reactions prepared in Buffer H to maintain stability of all enzymes. Incubations were kept for 24 hours at 37 °C before samples were prepared for LC-MS analysis using Method A. In summary for reactions 1-6, significant quantities of contaminant intermediates were found including cobalt-factor IV (16 min) and cobalt-factor II (25-32 min). For R1-4, known intermediates were detected as described in Sections 6.2.1 to 6.2.4. In reaction 5 this contained the starting substrate cobalt-precorrin-5A, CbiG, CbiD, CbiJ, CbiE and cofactors. The addition of CbiE led to a visual loss of colour, whilst LC-MS revealed that a new intermediate (Figure 6.11) eluted at 16.5 min with a peak at m/z 965.34. Furthermore a derivative of this intermediate eluted slightly later at 17 min with a peak at m/z 967.35. It is possible that this intermediate is cobalt-precorrin-7. For reaction 6, the further addition of CbiT to this mixture led to a novel peak eluting earlier at 12-13 min with peaks at m/z 920.33 and 922.34. It is possible that this intermediate is a new decarboxylated form of cobalt-precorrin-7 and could be cobalt-precorrin-7X, with X referring to decarboxylated.

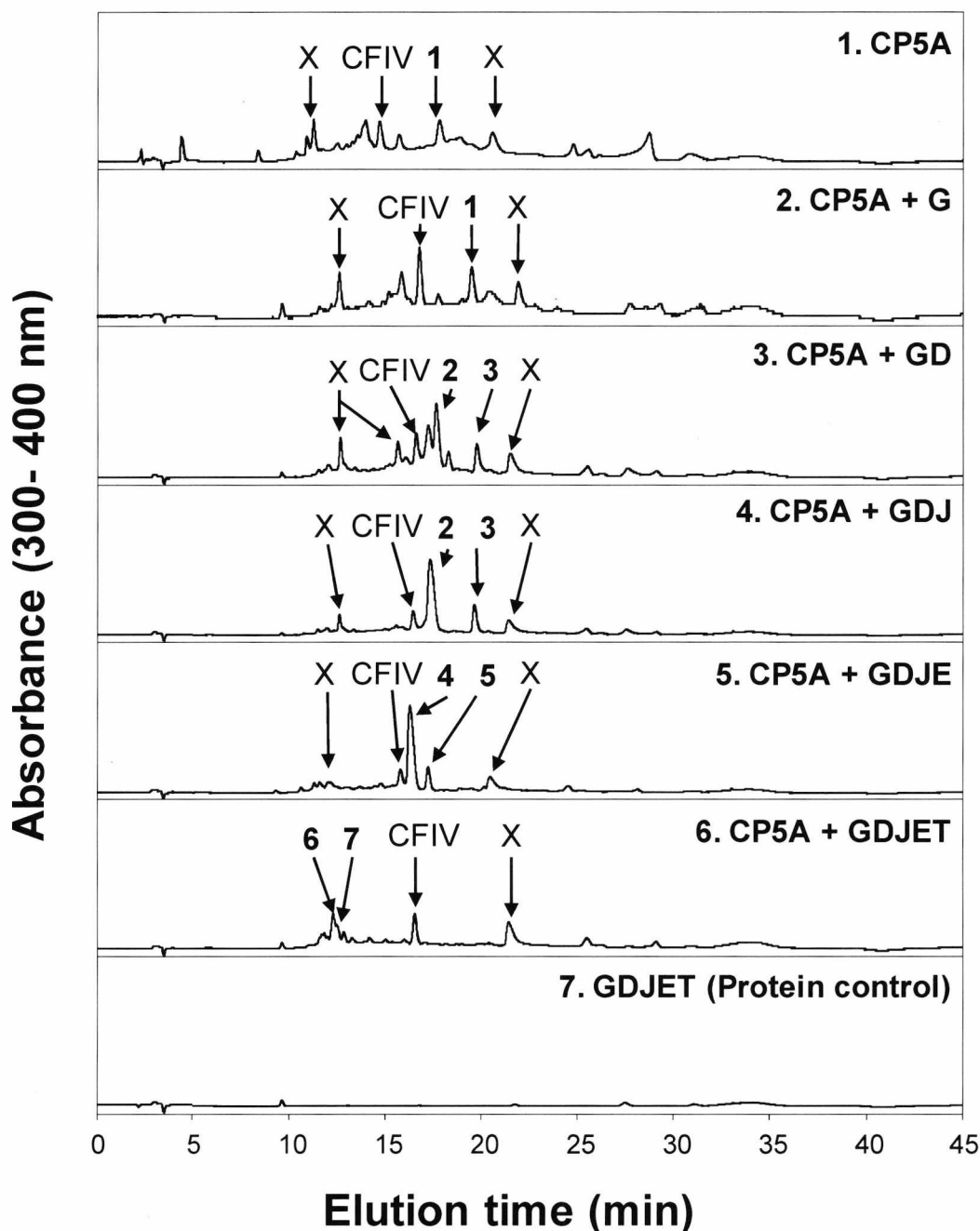
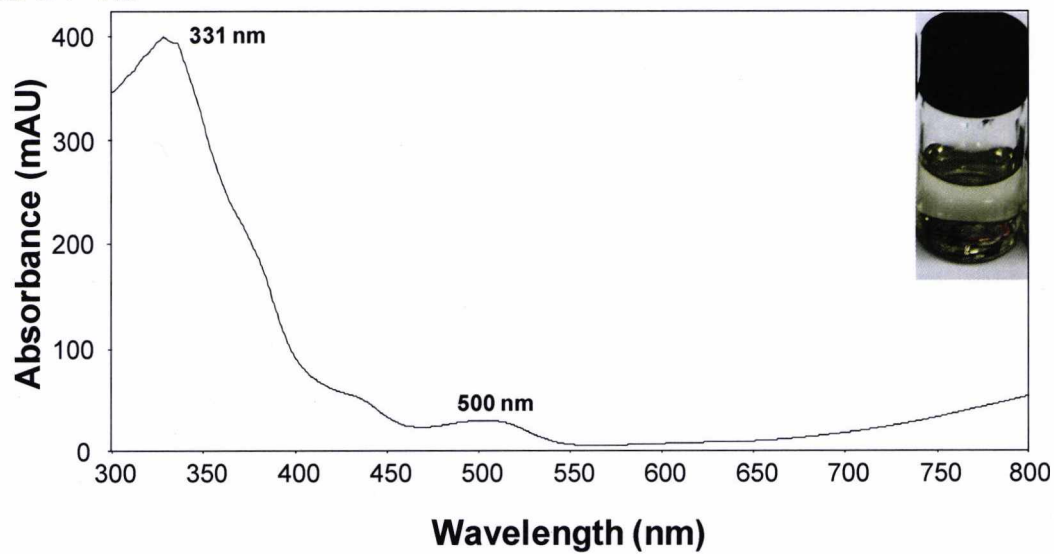


Figure 6.10. UV-Vis chromatogram of cobalt-precorrin-5A incubated with CbiG, -D, -J, -E, -T and cofactors (1 mM SAM, 50 μ M NADH), then separated and analysed by LC-MS. Reactions indicated in upper right corner of each run, Cbi proteins abbreviated as single letters. Cobalt-factor IV (CFIV) and unknown species (X) indicated by arrows. Tetrapyrrole intermediates indicated with arrows and numbered as follows, (1) oxidised cobalt-precorrin-5 (m/z 934.24, 920.24), (2) cobalt-precorrin-6A (m/z 951.31), (3) dihydro-cobalt-precorrin-6A (m/z 953.33) and the putative intermediates (4) cobalt-precorrin-7 (m/z 965.34), (5) dihydro-cobalt-precorrin-7 (m/z 967.35), (6) cobalt-precorrin-7X (920.33) and (7) dihydro-cobalt-precorrin-7X. The UV-Vis and MS spectrum of putative cobalt-precorrin-7 is shown next.

A. UV-Vis



B. MS

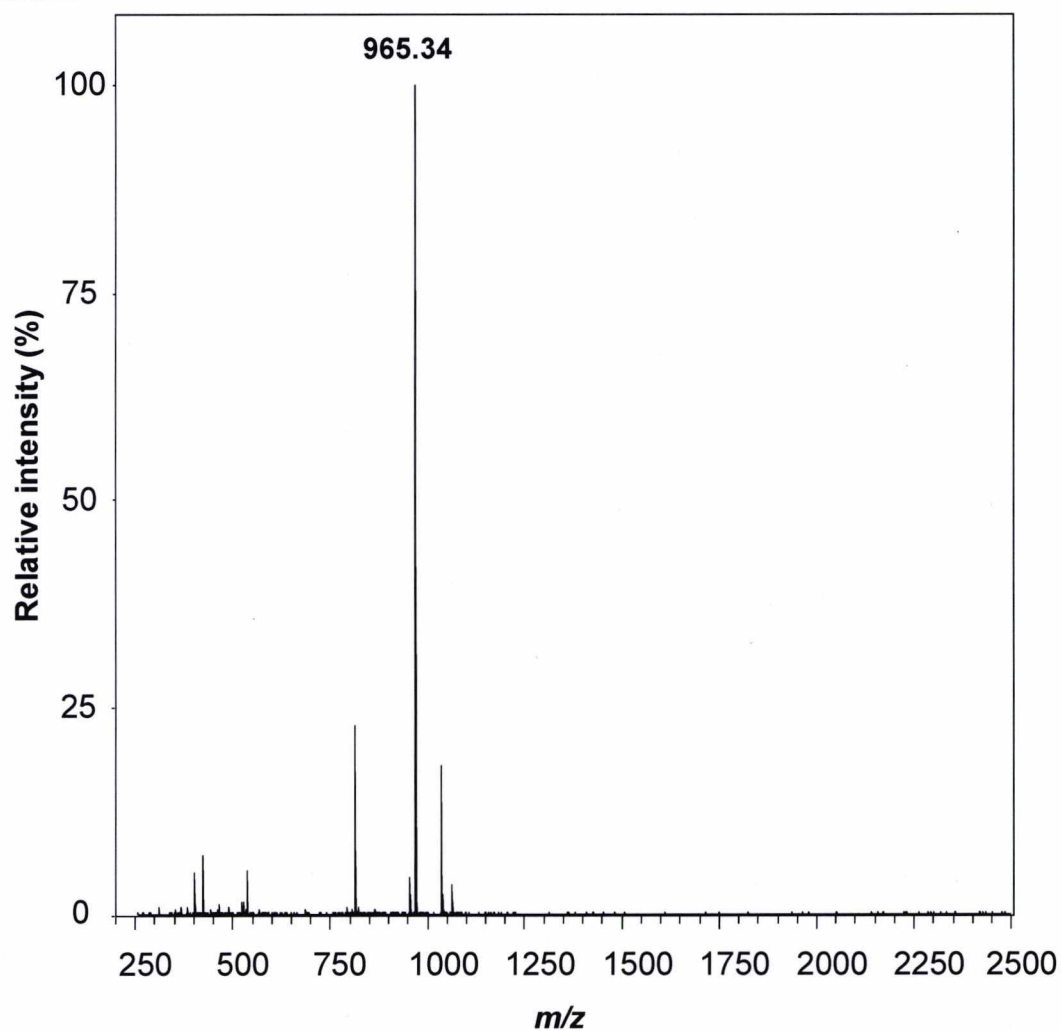


Figure 6.11. UV-Vis and MS of putative cobalt-precorrin-7 eluting at 16.5 min. A. UV-Vis of intermediate in anaerobic conditions. B. MS spectrum, single charged species at m/z 965.34. Inset picture represents a mixture containing cobalt-precorrin-7.

6.2.7 Isolation of cobalt-precorrin-8X and cobyrinic acid

The *M. thermoautotrophicus* CbiT did not seem to catalyse the final methylation event to form cobalt-precorrin-8X, possibly due to the N-terminal His₆-tag inactivating the enzyme. Therefore, the overall reaction of CbiET was studied using the *B. megaterium* DSM509 enzyme. In addition, it was important to determine if the expected intermediate could be converted into cobyrinic acid. To do this the CbiC from *B. megaterium* DSM509 was also investigated for activity. The enzymes CbiJ, CbiET and CbiC were overproduced in *B. megaterium* DSM319 and purified as before. Reactions were set up, as shown in Table 6.1.

Table 6.1: Reaction list to study the activity of CbiJ, CbiET and CbiC

Reaction	CP6A*	CbiJ	CbiET	CbiC
R1	-	-	-	-
R2	+	-	-	-
R3	+	+	-	-
R4	+	-	+	-
R5	+	-	-	+
R6	+	+	+	-
R7	+	+	-	+
R8	+	-	+	+
R9	+	+	+	+
R10	-	+	+	+

*Cobalt-precorrin-6A (CP6A) used at an unknown concentration (approximately 10-50 μ M)

All incubations contained SAM (5 mM), NADH (250 μ M) and were left at 37 °C for overnight incubation before purification, see Figure 6.12 for visual appearance. To eliminate cobalt-factor II/III, after incubation the samples were first heated at 80 °C for 30 min before acidification. Precipitate was removed, followed by purification with an RP-18 column. Samples were eluted in 40% ethanol (v/v), then subsequently freeze-dried and resuspended in 1% (v/v) acetic acid. After the purification the UV-Vis spectrum was recorded (Figure 6.12). The spectrum for cobalt-precorrin-6A with CbiJ (R3) was similar to the substrate control, cobalt-precorrin-6A, on its own (R2). However, with the addition of CbiET with or without CbiJ (R4, R6) the main absorbance peak for cobalt-precorrin-6A at 324 nm seemed to narrow and shift to 326 nm. In addition the broad absorbance between 350-500 nm diminished. This resulted in a visual colour loss, just as was seen with

the previous experiment with the *M. thermoautotrophicus* CbiE and CbiT. When CbiC was added to cobalt-precorrin-6A, on its own or with CbiJ (R5, R7), this seemed to produce a new spectrum. Although the main absorbance peak was at 324 nm, a new smaller peak appeared at 447 nm. However if cobalt-precorrin-6A was added to CbiET in combination with CbiC (R8, R9) a completely new spectrum was formed. There was no requirement for CbiJ in this reaction as new peaks formed at 318 (100) and 468 (32) nm. This spectrum is reminiscent of the previously characterised intermediate cob(II)yrinic acid (Nussbaumer *et al.*, 1981). Coupled with this, if this intermediate was exposed to air the spectrum slowly shifted (~1 hour) to a typical spectrum of cob(III)yrinic acid (Figure 6.13). However, there did not seem to be a requirement for CbiJ in these set of reactions.

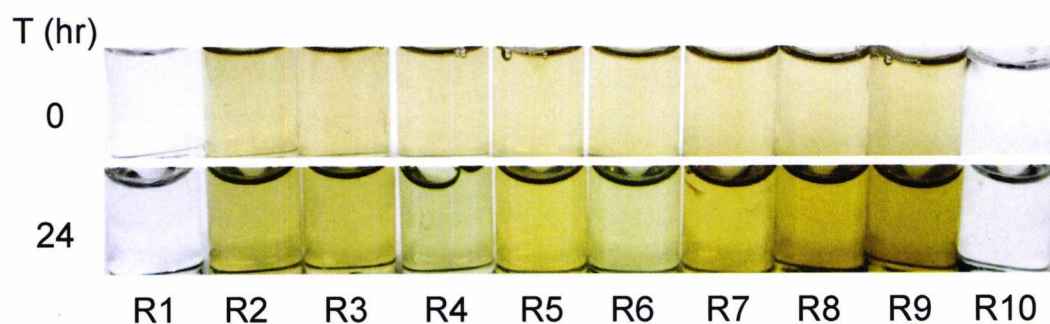
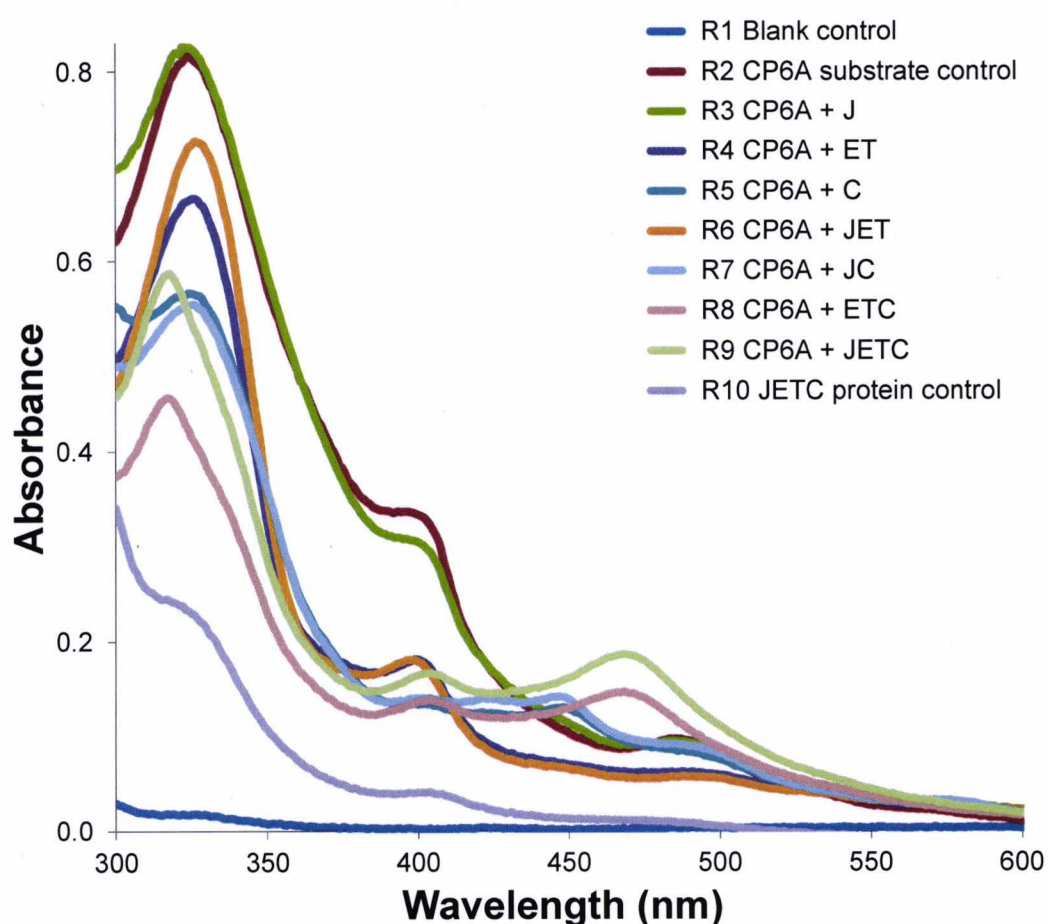
A. Visual appearance before and after incubation**B. UV-Vis spectra of incubations after RP-18 purification**

Figure 6.12. The Visual appearance and UV-Vis spectra of reactions containing cobalt-precorrin-6A, CbiJ, CbiET and CbiC with SAM. Reactions were incubated in Buffer H with cobalt-precorrin-6A (10-50 μM), CbiJ (2 μM), CbiET (2 μM), CbiC (2 μM), SAM (5 mM) and NADH (250 μM). **A.** Visual appearance before and after incubation. **B.** UV-Vis spectra. Samples were heated before acidification and centrifugation. The supernatant was purified using an RP-18 column, before freeze-drying and resuspension in 1% (v/v) acetic acid.

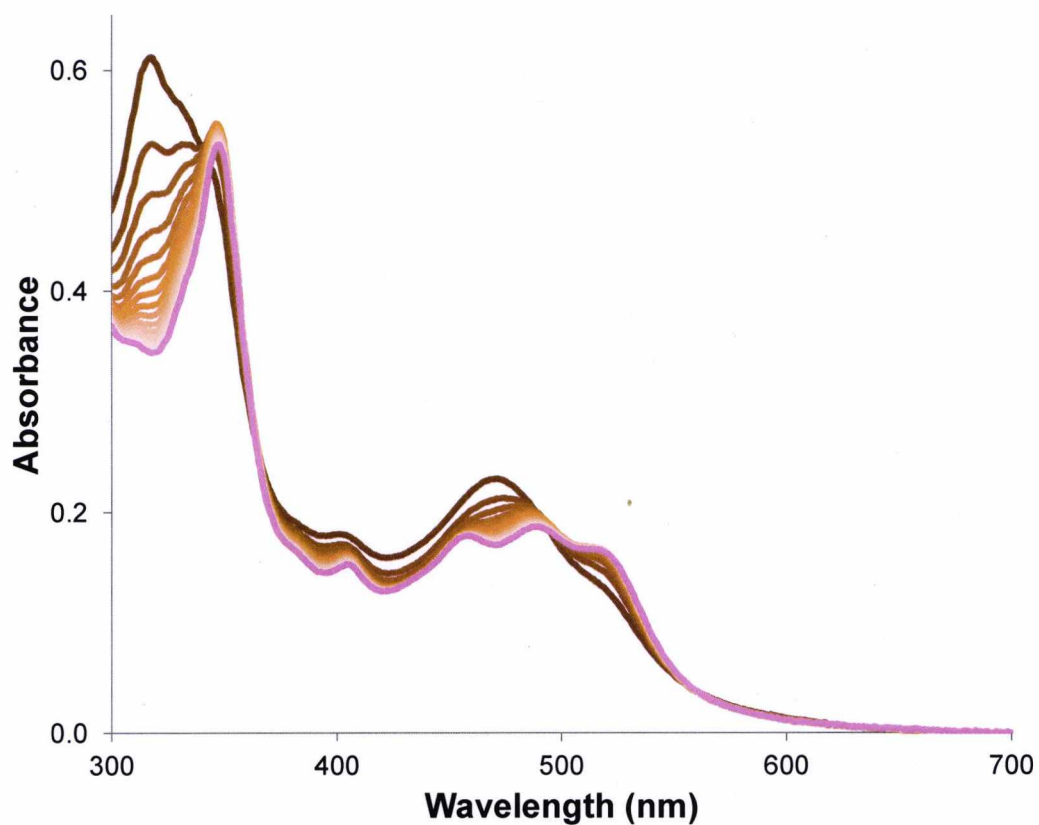


Figure 6.13. UV-Vis spectra showing the oxidation of cob(II)yrinic acid to cob(III)yrinic acid. Purified cob(II)yrinic acid from R9 was taken from anaerobic conditions and exposed to air, with UV-Vis scans taken every 10 min.

LC-MS was performed using Method A with 0.1% (v/v) TFA and acetonitrile as solvents, as summarised in Figure 6.14. No tetrapyrroles were detected in the protein and cofactor only control. The substrate control (R2) showed a couple of absorbance peaks mainly constituting of cobalt-precorrin-6A (m/z 951.28), which eluted between 17-18 min. As expected the addition of CbiJ (R3) did little to alter this, as was the same with CbiC added (R5). However CbiET (R4) produced quite a remarkable result, whereby a new series of peaks eluted earlier between 10-16 min. The peak eluting at 15 min had a single symmetrical absorbance peak close to 323-325 nm with a clean MS spectrum including a single charged ion with a peak at m/z 938.36, along with a more dominant double charged ion with an m/z 469.68. In addition to this, another peak eluting earlier at 13 min with a main absorbance peak at 305 nm had a single charge m/z 936.35, along with a double charged ion at m/z 468.67.

If CbiJ was incubated with CbiET and cobalt-precorrin-6A (R6) then the same masses were obtained, although the intermediates appeared to elute slightly earlier. The incubation containing CbiJ and CbiC (R7) only revealed the unconverted substrate cobalt-precorrin-6A. The theoretical mass of cobalt-precorrin-8X is the same as the mass for cobyrinic acid, which is 938.36 Da. Once CbiC was incubated with CbiET (R8), new intermediates formed that eluted later at 24-27 min, with the same masses obtained with the same incubation lacking CbiC (R4, R6). Although mostly present as a single charged species at m/z 938.34, this confirmed along with the UV-Vis spectra that this enzyme mixture had converted cobalt-precorrin-6A to cobyrinic acid. This suggested that the intermediate that formed in the absence of CbiC was cobalt-precorrin-8X. In addition to cobyrinic acid, another peak eluting slightly later at 26 min had an unusual UV-Vis spectrum. Several absorbance peaks at 270 (100), 345 (67), 390 (50), 480 (46) nm suggested a novel intermediate had formed. This new compound could be obtained regardless of whether CbiJ was added to the incubation. In addition the MS spectrum revealed the compound had a single charged peak at m/z 954.33 and a double charged peak at m/z 477.66. The true identity of this compound remains unknown but it is possibly a derailed intermediate. DAD UV-Vis and MS spectra are shown for each compound obtained.

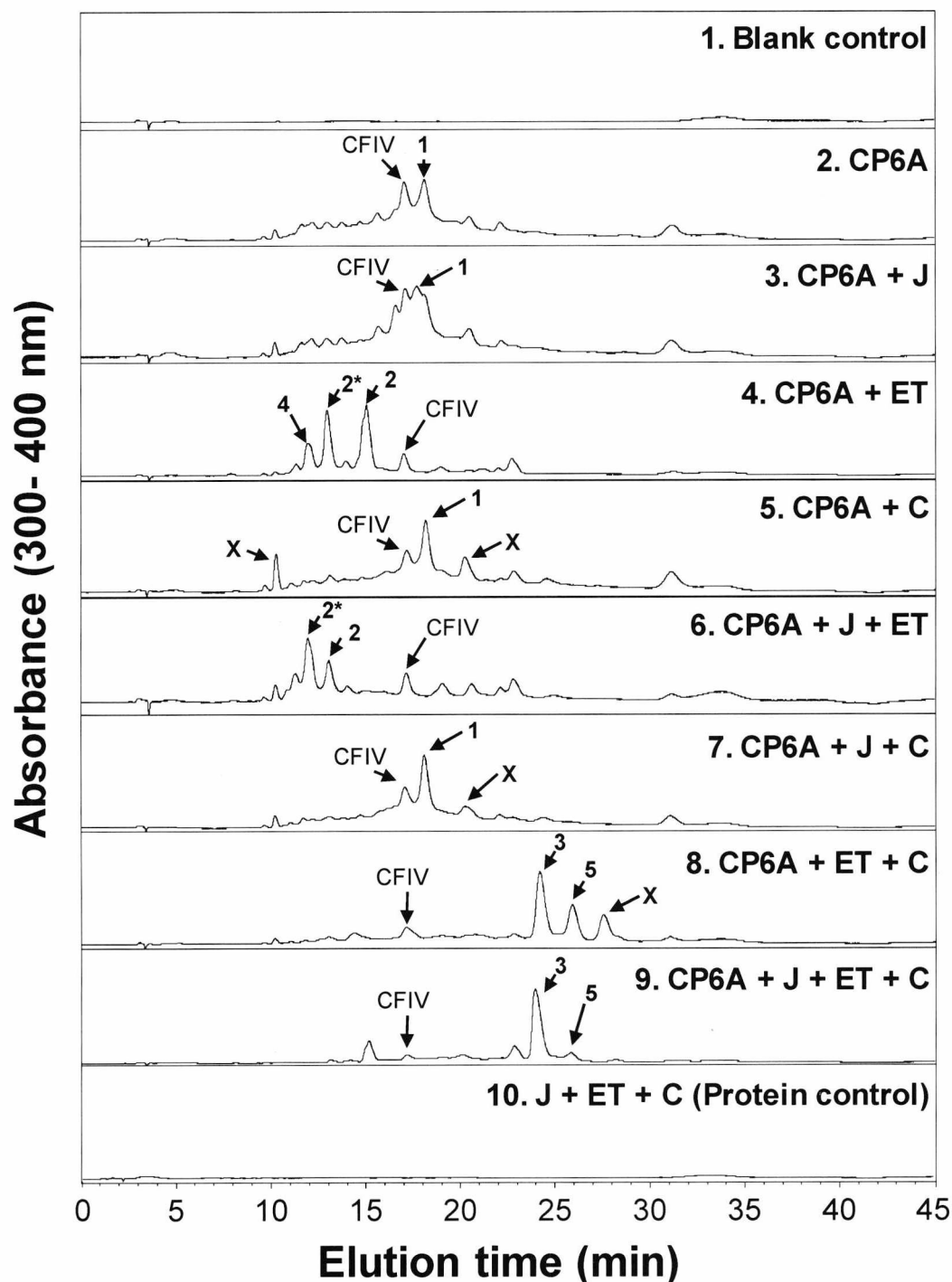


Figure 6.14. UV-Vis chromatogram of cobalt-precorrin-6A (CP6A) incubated with CbiJ, CbiET, CbiC and cofactors (5 mM SAM, 250 μ M NADH), then separated and analysed by LC-MS. Reactions indicated in upper right corner of each run, Cbi proteins abbreviated as single letters. Cobalt-factor IV (CFIV) and (X) peaks with $m/z < 900$ indicated by arrows. Tetrapyrrole intermediates indicated with arrows and numbered as follows, (1) cobalt-precorrin-6A (m/z 951.28), (2) cobalt-precorrin-8X (m/z 938.36), (2*) oxidised cobalt-precorrin-8X (m/z 936.34). (3) cob(II)yrinic acid (m/z 938.34), (4) unknown species (m/z 952.34) and (5) unknown species (m/z 954.33). (2, 2*) cobalt-precorrin-8X, (3) cob(II)yrinic acid and (4,5) are shown next with UV-Vis and MS spectrums provided. Compounds labelled with X had an m/z less than 900.

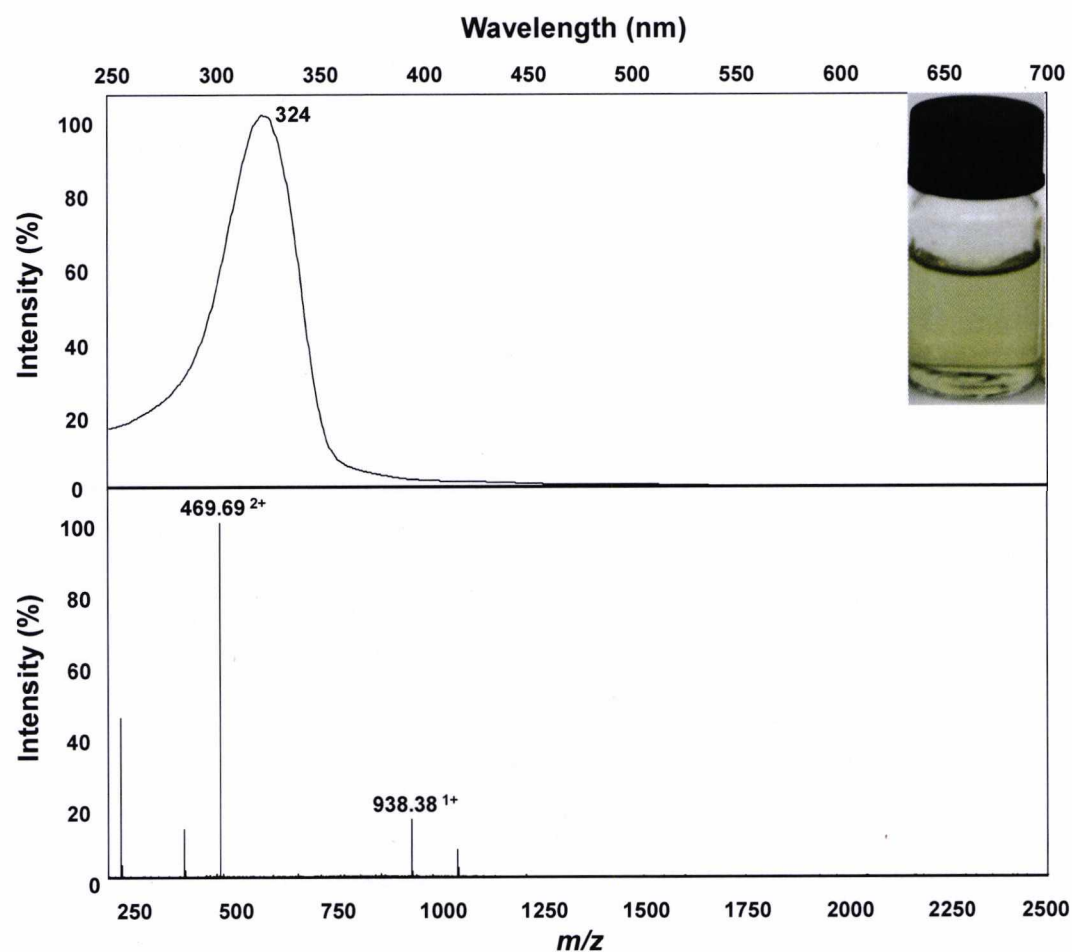


Figure 6.15. UV-Vis and MS of (2) cobalt-precorrin-8X eluting at 15 min. UV-Vis of intermediate in anaerobic conditions and MS spectrum, with a single charged species m/z 938.38 and double charged species m/z 469.69. The inset picture is of a mixture containing cobalt-precorrin-8X and is not representative of a single intermediate.

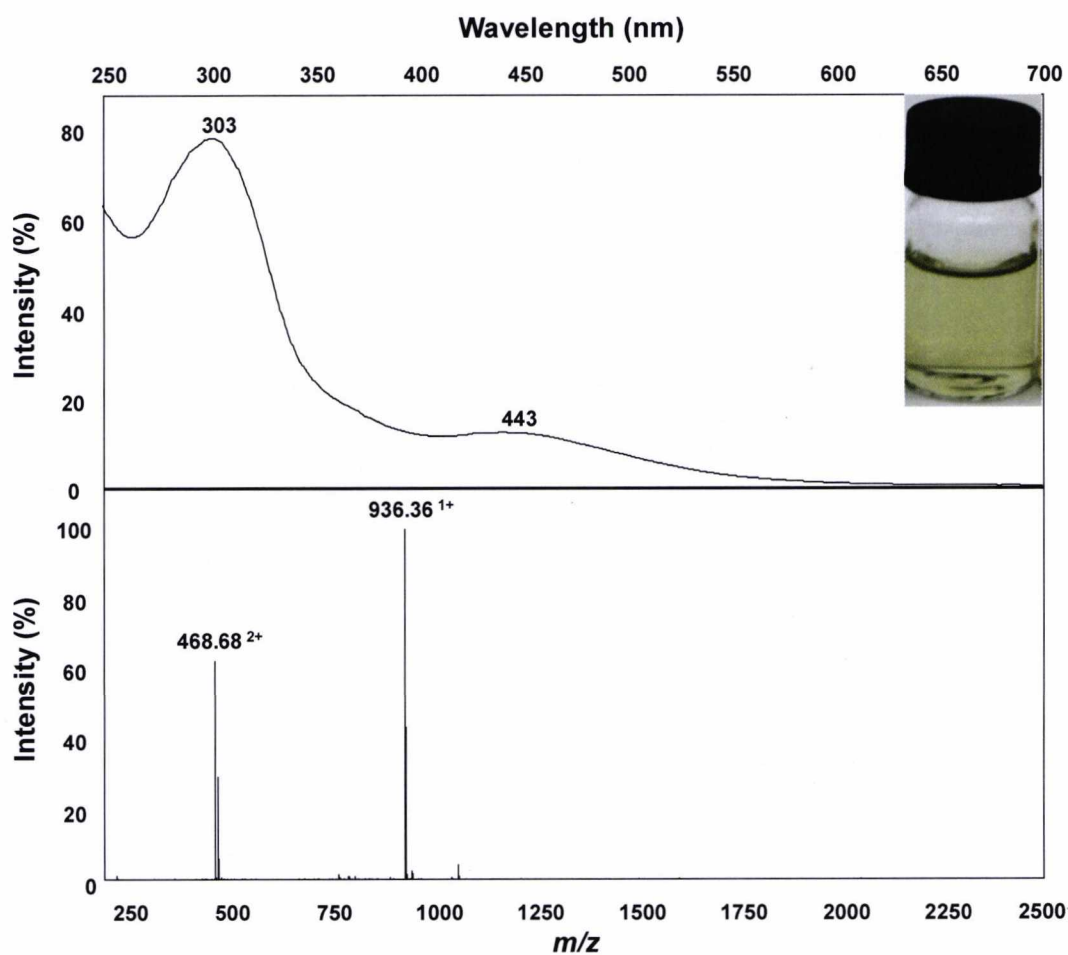


Figure 6.16. UV-Vis and MS of (2*) oxidised cobalt-precorrin-8X eluting at 13 min. UV-Vis of intermediate in anaerobic conditions and MS spectrum, with a single charged species m/z 936.36, double charged m/z 468.68. The inset picture is of a mixture containing oxidised cobalt-precorrin-8X and is not representative of a single intermediate.

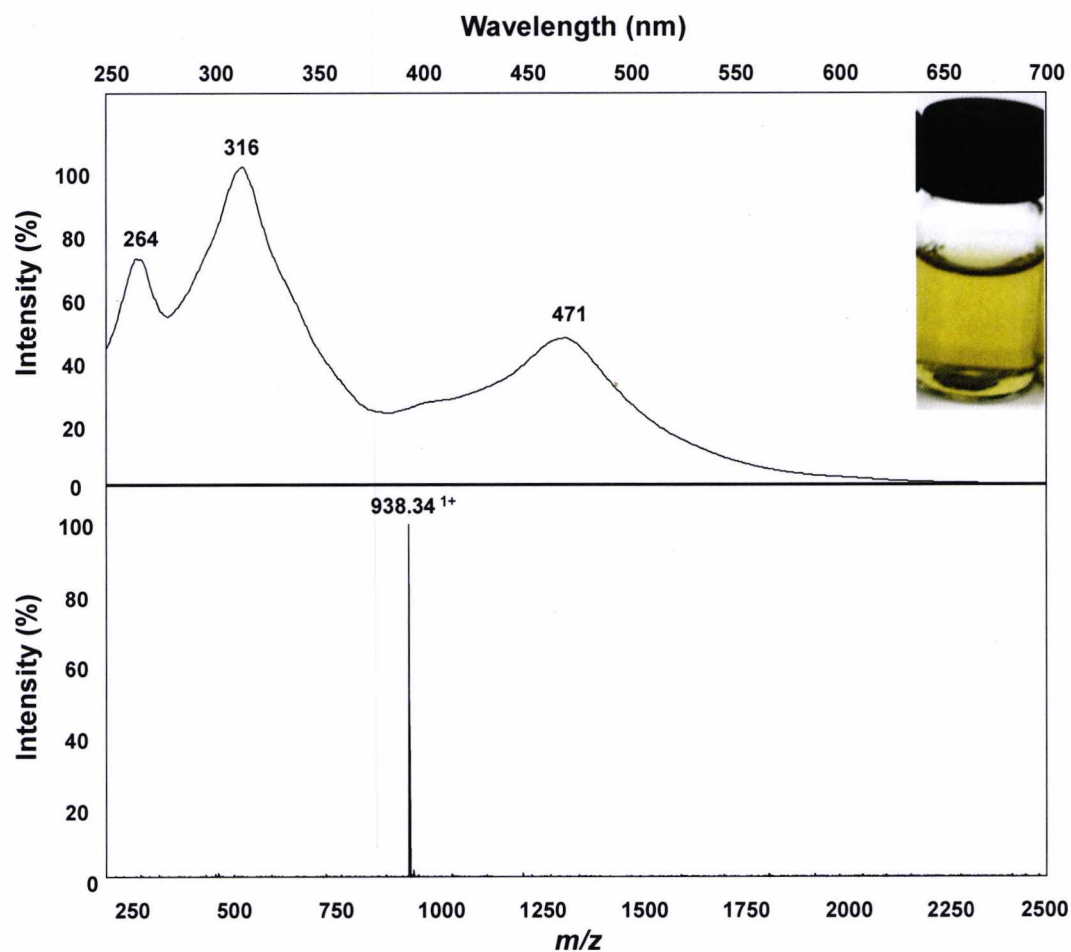


Figure 6.17. UV-Vis and MS of (3) cob(II)yrinic acid eluting at 24 min. UV-Vis of intermediate in anaerobic conditions and MS spectrum, with a single charged species m/z 938.34. The inset picture is of a mixture containing cob(II)yrinic acid and is not representative of a single intermediate.

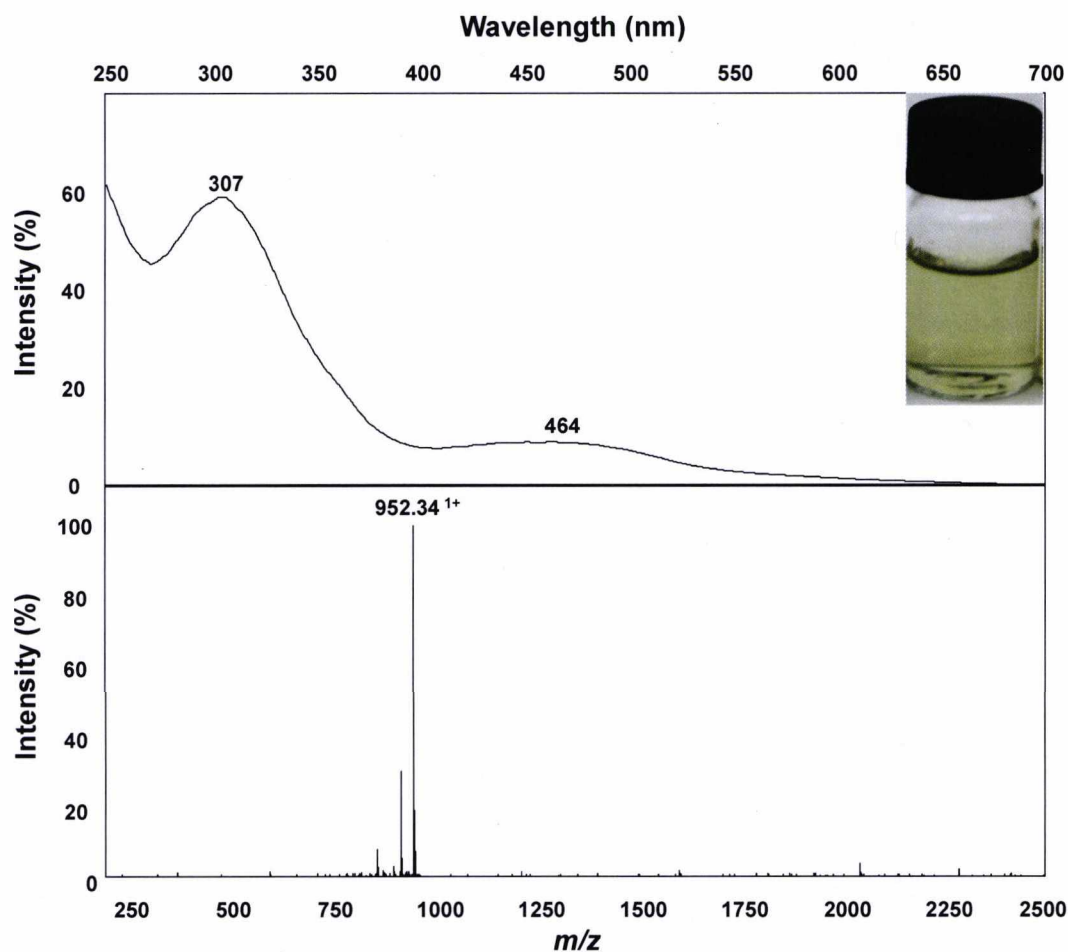


Figure 6.18. UV-Vis and MS of (4) unknown species produced by CbiET and cobalt-precorrin-6A, eluting at 11.9 min. UV-Vis of intermediate in anaerobic conditions and MS spectrum, with a single charged species m/z 952.34. The inset picture is of a mixture containing the unknown intermediate and is not representative of a single intermediate. The identity of this compound is unknown, but it could be an adduct of cobalt-precorrin-8X or may contain an additional methyl group

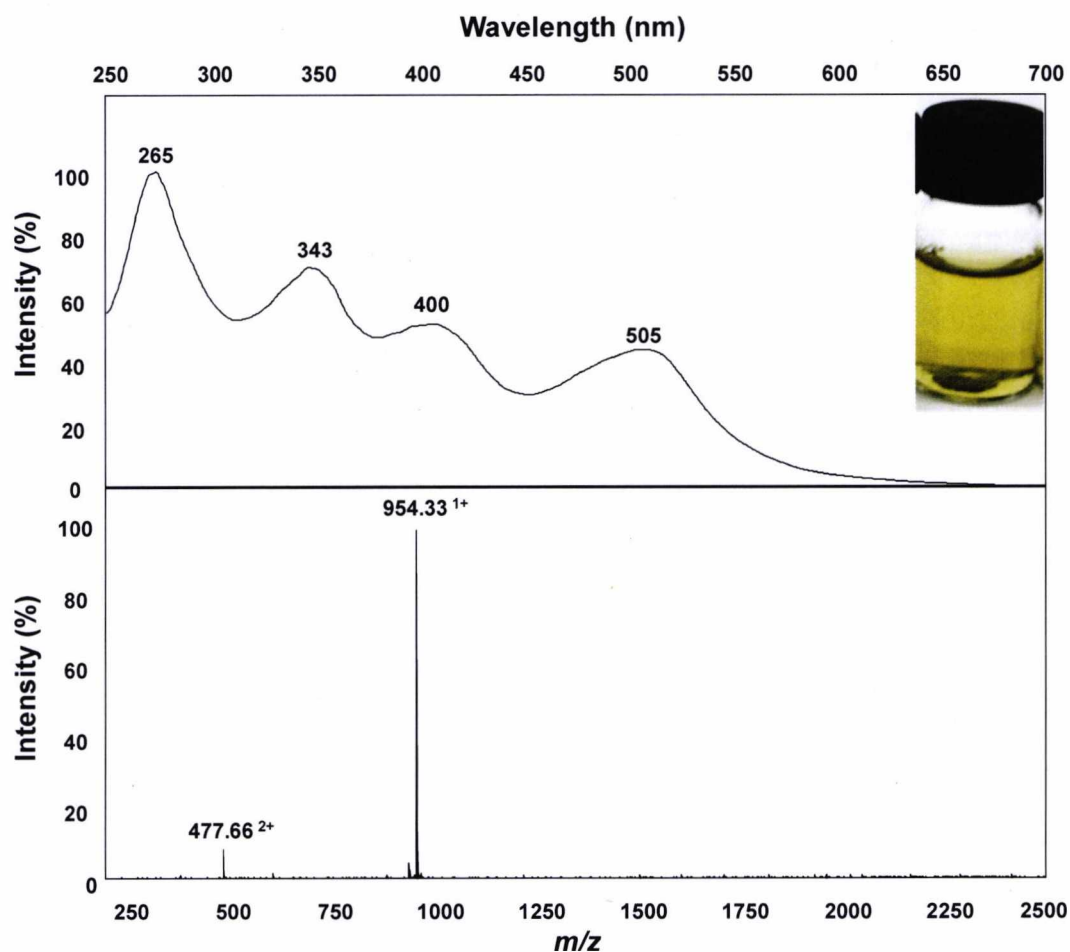


Figure 6.19. UV-Vis and MS of (5) unknown species produced by CbiJ, CbiET, CbiC and cobalt-precorrin-6A, eluting at 26 min. UV-Vis of intermediate in anaerobic conditions and MS spectrum, with a single charged species m/z 954.33, double charged m/z 477.66. The inset picture is of a mixture containing the unknown intermediate and is not representative of a single intermediate. The identity of this compound is unknown.

6.3 Discussion

The major dilemma with the anaerobic biosynthesis of vitamin B₁₂ is that although some of the intermediates have been isolated, many of the pathway enzymes and intermediates have remained uncharacterised. Indeed, those that have been isolated have only been produced in minute quantities and not in their free-acid form. Thus, little is understood on how the process of ring contraction occurs and why the cobalt ion is inserted at such an early step. This work details the first full elucidation of the anaerobic biosynthesis of the corrin macrocycle of vitamin B₁₂ using the *B. megaterium* DSM509 enzymes. The strain DSM509 is a past industrial producer of vitamin B₁₂ and its complete anaerobic vitamin B₁₂ biosynthetic cluster was isolated and characterised in *E. coli* (Raux *et al.*, 1998b). By using the recent advances in recombinant protein production and purification in *B. megaterium* DSM319 (Biedendieck *et al.*, 2007b; Malten *et al.*, 2006; Vary *et al.*, 2007), the relevant B₁₂ biosynthetic genes were homologously expressed in the host to obtain high yields of soluble and active enzymes, as shown in Chapter 4.

The original design of this thesis was to overproduce the B₁₂ biosynthetic enzymes homologously in *B. megaterium* DSM319 to overcome solubility problems that occurred during heterologous production in *E. coli*. The successful production of a range of enzymes demonstrates *B. megaterium* as a good alternative tool in molecular biotechnology. Apart from the odd exceptions such as CbiL and CbiW, the majority of the enzymes could successfully be produced to high yields thus providing an easily obtained source of active enzymes. Once the key barrier of anaerobic ring contraction was unlocked (Chapter 5) and high quantities of the ring contracted product cobalt-factor IV were obtained, these key elements were combined in order to characterise the rest of the pathway in detail. Beginning with CbiF and CbiG, this work details the first UV-Vis spectral details of cobalt-precorrin-5 and shows its first isolation in its free-acid form. One major assumption is made and that is that the intermediate is presumed to have been reduced back to the precorrin level. This is mainly because the factor intermediates have absorbance close to 595 nm and are more intense in colour than their precorrin counterparts. In addition, in reference to previous published data, the UV-Vis spectra observed with this work are close to the previous described cobalt-precorrin-5B. Unfortunately

time limited an individual examination of this reductive process and a more thorough investigation will hopefully identify what enzyme is responsible.

It was interesting to find that the CbiF catalysed methylation (presumably at C-11) would not occur unless CbiH₆₀ was present. This could suggest that CbiF is substrate inhibited or alternatively a reductive process (factor → precorrin) is essential before CbiF can methylate. Therefore if cobalt-factor IV is converted directly to cobalt-precorrin-5A this would require a two electron reduction of the corrin macrocycle along with the C-11 methylation. Previous research had suggested that cobalt-precorrin-5A/B can be biosynthesised from either precorrin-3 or factor III (Kajiwara *et al.*, 2006). However, it remains uncertain how the factor intermediate is reduced back to the precorrin level. It is unlikely a simple methyltransferase could catalyse this, which may suggest a true role for the [4Fe-4S] cluster in CbiH₆₀. But how could this process occur in organisms such as *S. enterica* or *P. shermanii*, which do not contain this iron-sulphur containing CbiH? This particular question remains unsolved in both pathways and will be a key focus in follow up experiments.

The activity of CbiG could not be confirmed. The enzyme was found to be unstable and precipitated if incubated without CbiH₆₀. Although this was not conclusive evidence for the activity of CbiG, it should be noted that there are several gene fusions between CbiH, CbiF and CbiG in other organisms possessing the anaerobic pathway. This may explain why the reactions catalysed by these enzymes have been a major limitation in studying the rest of the pathway. Although cobalt-factor IV itself is indeed quite stable aerobically, the next intermediate isolated after the CbiF methylation is not. Many organisms have fusion proteins, with CbiET, CbiHF, CbiHG, CbiDJ, CbiXC, CbiJET, CbiXW, CbiGW, CbiH-SirC and CbiHC combinations found in other organisms (data not shown). As the intermediates are relatively unstable, this form of enzyme fusion may aid in metabolic channelling to increase flux through this particularly oxygen-sensitive stage of the pathway. Cobalt-precorrin-5A was previously described as having a different spectrum to cobalt-precorrin-5B with a maximum peak absorbance around 320 nm (Kajiwara *et al.*, 2006). This was not observed in the current work described in this thesis. Further to this, the conversion of cobalt-precorrin-6A to cobyrinic acid does not require the

further addition of CbiJ. These observations could be explained by the presence of some CbiG and CbiJ. Both CbiG and CbiJ could be adventitiously purified from the *B. megaterium* DSM319 background proteome, either through a protein-protein interaction or non-specifically. To negate this possibility the experiments could be repeated by purifying the enzymes from cultures grown with a high level of exogenous vitamin B₁₂, to suppress expression of the genomic B₁₂ biosynthetic genes. This approach has been demonstrated successfully with the CbiX^L, as this is a common contaminant in all His₆-bind protein purification from *B. megaterium* DSM319. Furthermore, the B₁₂ intermediates found co-purified with CbiF, CbiJ and CbiD could also be greatly reduced with this method.

The role of CbiD is one of the most intriguing findings of this chapter. The elucidation of the *A. fulgidus* CbiD structure (PDB_1SR8) revealed that the protein does not contain a Rossman fold, which is considered an essential structural feature for SAM binding. The *B. megaterium* CbiD shares 36% identity and 52% similarity with the *A. fulgidus* CbiD. However, sequence and structural alignments suggest that there is no similar enzyme known in Nature. The structure does, however, contain a deep pocket which could accommodate cobalt-precorrin-5B, which was modelled with AutoDock Vina1 (Figure 6.20). Although CbiD does contain several glycine rich regions these are found on the outer surface of the enzyme and it seems unlikely that they play a role in SAM binding. In this work it has been shown *B. megaterium* CbiD is active as the missing methyltransferase since it transforms the orange coloured cobalt-precorrin-5B into the bronze coloured cobalt-precorrin-6A, but only in the presence of SAM. The final structure of cobalt-precorrin-6A requires confirmation by NMR spectroscopy. The fact that CbiD is able to function in isolation suggests that there is no need for a multi-enzyme complex.

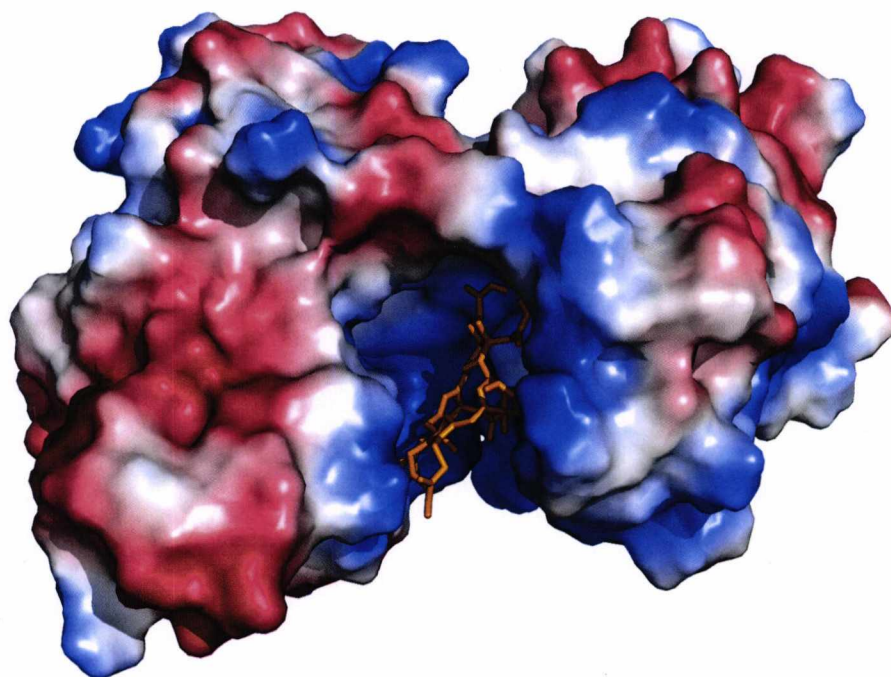



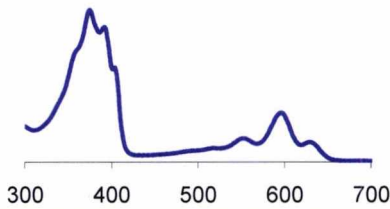

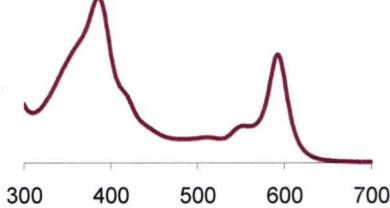

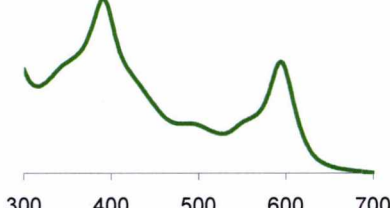

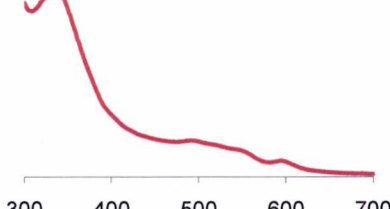

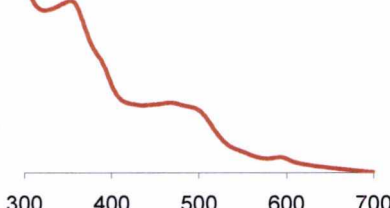
Figure 6.20. Protein-ligand docking model of cobalt-precorrin-5B and *A. fulgidus* CbiD (PDB_1SR8). AutoDock Vina 1 (Trott & Olson, 2010) was used to model the binding of cobalt-precorrin-5B (Orange) with CbiD (Mapping potential to molecular surface). The red color represents negative potential, the blue color expresses positive potential, and the white color expresses zero potential.

The reactions catalysed by CbiET and CbiC were only briefly studied. Essentially this work shows the completion of the anaerobic pathway with the first reported step-by-step enzymatic synthesis of cobyrinic acid, starting from the substrate ALA. CbiET catalyses the C-5, C-15 methylations and the decarboxylation of the *e* acetate side chain. This yields another new intermediate cobalt-precorrin-8X, which CbiC converts into the more spectrally distinguishable cobyrinic acid. Furthermore, it seems likely that the CbiE can methylate cobalt-precorrin-6A/B to produce cobalt-precorrin-7. Previous research showed the CbiT from *S. enterica* could methylate the C-15 position and decarboxylate cobalt-factor III in an out-of-order reaction (Roessner & Scott, 2006). Now further research will be needed to confirm that cobalt-precorrin-7 is a true intermediate and that this can be converted into cobalt-precorrin-8X and hence into cobyrinic acid. A summary of the entire anaerobic pathway is provided in Table 6.2.


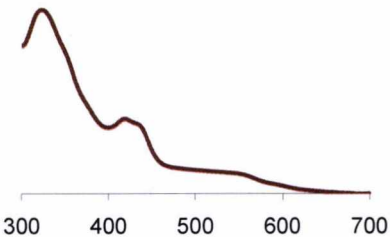

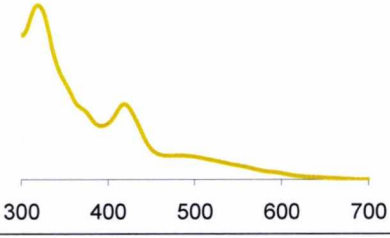

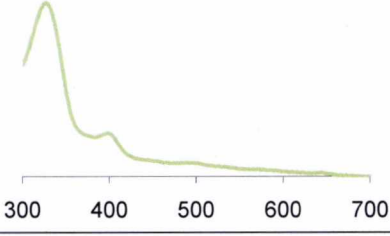

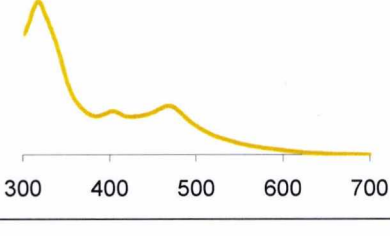

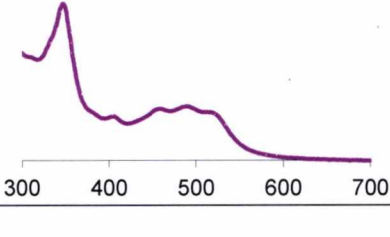
It is interesting to speculate how the B₁₂ biosynthetic enzymes achieve their specificity – is it down to substrate recognition or is it some form of substrate channelling within a larger multi-enzyme complex? The lack of specificity gives the

opportunity of using the enzymes in a combinatorial approach to make a range of new compounds based on the corrin template. Understanding how this happens is an important pre-requisite within the area of synthetic biology if we are to use biology to assist in the production of new chemicals and drugs.

Table 6.2. Summary of the anaerobic pathway intermediates from factor II to cobyrinic acid

Intermediate	Visual	UV-Vis Spectrum	<i>m/z</i>
Factor II			863.30
Co-factor II			919.21 917.20
Co-factor III			931.21 929.20
Co-factor IV or Co-precorrin 4 ?			948.23 946.22
Co-precorrin-5			934.24 920.24

Continued

Intermediate	Visual	UV-Vis Spectrum	<i>m/z</i>
Co-precorrin-6A			951.29
Co-precorrin-6B			ND
Co-precorrin-8X			938.36
Cob(II)yrinic acid			938.36
Cob(III)yrinic acid			938.34

Chapter 7

General Discussion

The dawn of biochemistry began in 1893 with Eduard Buchner showing the juice of yeast cells could ferment sugars. Whilst modern biochemistry has unveiled the majority of central metabolic pathways, the anaerobic biosynthesis of the most complex of small molecules, vitamin B₁₂, has proved an elusive challenge in biochemistry. Indeed, the biosynthesis of vitamin B₁₂ is symbolically referred to as a biochemical Mount Everest (Battersby, 2000), although in evolutionary terms it is perhaps more pertinent to consider it as Mount Improbable. In fact, to tackle the question of how such a complex metabolic pathway may have arisen one has to understand the chemical logic underpinning each step. This thesis confronts the challenge of the anaerobic biosynthesis of vitamin B₁₂ and successfully delineates the first step-by-step biosynthesis of the corrin macrocycle. Key to this breakthrough has been the unlocking of the anaerobic ring contraction process.

An electron to spark the ignition

In essence ring contraction is the most distinctive reaction in cobalamin biosynthesis and is essential to impound the cobalt ion within its organic enclosure. Attempts have previously been made to study the anaerobic ring contraction reaction, but in reality only small amounts of the product, cobalt-factor IV or cobalt-precorrin-4, have ever been isolated. These approaches have involved the use of crude lysates with enhanced levels of CbiH in *E. coli*. It seems that the CbiH₆₀ catalysed ring contraction is dependent not only on SAM, but also an electron source. It was previously hypothesised that the reduction of the cobalt ion could play a role during ring contraction (Scott, 2003). However, this seems to be unlikely, with EPR studies of CbiH₆₀ indicating that the anaerobic ring contraction step does not involve any change in redox state of the cobalt ion. Instead, it seems more important for these enzymes to recognise the cobalt ion and bind through an unknown ligand to anchor the tetrapyrrole substrate within the active site. The reaction also requires an electron donor (DTT, NaDT), although this is an unexplained finding. *B. megaterium* is capable of producing vitamin B₁₂ under aerobic or anaerobic conditions, whilst organisms such as *S. typhimurium* are only capable of producing vitamin B₁₂ under anaerobic conditions (Raux *et al.*, 1998b). It is possible that *B. megaterium* may possess auxiliary enzymes to protect the oxygen labile steps of the pathway, thus providing a route during aerobic growth (Raux *et al.*, 1998b). Nonetheless, in the context of this thesis, it appears that the

anaerobic ring contraction mechanism may be promoted under reductive conditions and is blocked by raised oxygen concentrations during aerobic growth of organisms such as *S. enterica*. Thus, the C-terminal extension of *B. megaterium* may provide an environment for the ring contraction reaction to proceed under aerobic conditions. Previous research has shown that CbiH₆₀'s C-terminal extension is not required in *E. coli* for vitamin B₁₂ biosynthesis (Raux *et al.*, 1998a). However, these *in vivo* studies may have overlooked the role of this C-terminal extension. It is tempting to speculate that CbiW may act as CbiH₆₀'s redox partner. However it should be considered that both CbiH₆₀ and CbiW are insoluble when produced in *E. coli*, which may affect these *in vivo* findings.

A panoply of colours

The anaerobic B₁₂ pathway proceeds via colourless (uroporphyrinogen III), tangerine orange (precorrin-2), deep purple (factor II), crimson red or sky blue (cobalt-factor II can appear as both), chlorophyll green (cobalt-factor III), strawberry red (cobalt-factor IV), tango orange (cobalt-factor V or cobalt-precorrin-5), roman bronze (cobalt-precorrin-6A), straw yellow (cobalt-precorrin-6B), almost water colourless (cobalt-precorrin-8X) and flame yellow [cob(II)yrinic acid], or oxidation to carnation pink [cob(III)yrinic acid]. However, this panoply of colour belies a Pandora's Box of problems linked with the lability of the pathway intermediates. Thus the beauty of the individual components of the biosynthesis is matched by the malevolence of molecular oxygen.

The research described in this thesis successfully demonstrates the first elucidation of the anaerobic biosynthesis of the corrin macrocycle of vitamin B₁₂. Yet a number of questions still need to be addressed including the ring contraction mechanism, the isolation of cobalt-precorrin-6B, the role of CbiG and the process by which cobalt-factor III is reduced. Although this research could not demonstrate a mechanism for anaerobic ring contraction, an efficient method to make large quantities of cobalt-factor IV was found. This was the major limiting barrier in all previous research into this pathway as yields of only around 5% had been obtained (PhD Thesis: Frank, 2007). The next intermediate that was acquired in this research appears to be cobalt-precorrin-5. It seems the C-11 methylation catalysed by CbiF creates a highly unstable intermediate, which seems to become stable only after the

rearrangement of the C-11 methyl group to the C-12 position. The next logical step will be to employ NMR to investigate when the δ -lactone ring is opened.

One of the major findings of the current research has been the confirmation of a novel methyltransferase CbiD, representing a novel class of methyltransferase. With further structural information on the mode of SAM and cobalt-precorrin-5 binding, this will reveal more detail on the exact mechanism. In addition this research also proposes that NADH is the preferred reducing agent for the CbiJ catalysed C-19 reduction. This was surprising as many enzymes utilising this cofactor can use both NAD(P)H forms. This is in contrast to the aerobic pathway that utilises NADPH (Blanche *et al.*, 1992c). There is no typical sequence motif for NAD(P)H binding in either CobK or CbiJ. Structural information should therefore reveal how binding of this cofactor is afforded. Although cobalt-precorrin-6B was found to be acid sensitive, further attempts to characterise this intermediate will be sought by utilising NMR. For CbiET it was strange to observe a loss of colour during the production of cobalt-precorrin-8X. Unfortunately there was not enough time to allow for a more detailed examination of this reaction, but with completion of the pathway up to cobyrinic acid this allows each intermediate to be studied and an individual extinction coefficient can be calculated. This will permit a kinetic analysis of the majority of the enzymatic reactions.

Despite any significant evidence, it has long been suspected that the B₁₂ pathway enzymes may interact in the cell to enable a more rapid and efficient processing of these unstable intermediates. This could be through a multi-enzyme complex (Roessner *et al.*, 2005) or alternatively through a more loosely assembled metabolon. Such a scenario would allow the product of one enzyme to be channelled to the next enzyme in the pathway. However, the research outlined provides little evidence to support either of these theories. After cobalt-precorrin-5 is formed, there seems to be no requirement for other enzymes to be present when studying the activities of CbiD, CbiJ, CbiET and CbiC *in vitro*. Indeed all of the intermediates in these steps can be kept for several days/months in an anaerobic state. Although much is still to be learnt concerning the overall systems biology of a prokaryotic cell, from these studies it seems that the intermediates in the vitamin B₁₂ pathway are more stable than previously thought.

Enzymes as a trap to extract metabolic pathway intermediates

New evidence as to how pathway intermediates can be extracted from the cellular metabolic pool using enzymes has been provided. Not only has this been successfully demonstrated in the aerobic pathway (Evelyn Raux, unpublished data), but this is currently being developed as an approach to examine novel uncharacterised metabolic pathways. Now that the anaerobic pathway has successfully been elucidated, it should be possible to reconstruct the pathway in non-B₁₂ producing bacteria such as *E. coli* or *B. subtilis*. An organism such as *B. subtilis* offers its advantages as an easy model organism to genetically modify and as a host for recombinant protein production. In addition it is better characterised than *B. megaterium* and is capable of high-cell density growth under anaerobic conditions. Furthermore, as the *B. megaterium* CbiH₆₀ is insoluble when produced in *E. coli*, an efficient enzyme is required to make cobalt-factor IV *in vivo*. Indeed, engineered strains of *E. coli* have only been able to produce very limited vitamin B₁₂ quantities *de novo*, with the ring contraction step the major block in the pathway (PhD Thesis: Frank, 2007). However, other problems regarding cobalt transport and chelation into factor II have also shown to be troublesome in *B. megaterium*.

The evolution of B₁₂ biosynthesis

It has long been speculated as to how Nature's biosynthetic Mount Everest (Battersby, 2000) evolved. The anaerobic route to vitamin B₁₂ is believed to have appeared millions of years before the atmosphere first became abundant with molecular oxygen $\sim 1.7 \times 10^9$ years ago. It is challenging to support either the retrograde or patchwork theories of evolution in the origin of vitamin B₁₂ biosynthesis (Holliday *et al.*, 2007; Raux, 1999; Warren *et al.*, 2002). However, it is important to understand which of the tetrapyrroles was first to evolve and why? Originally it was believed that vitamin B₁₂ was the most ancient, due to its link with ribonucleotide reduction (Dickman, 1977). Nonetheless, the significance of this reaction is questionable because it can be catalysed by three different enzymes, all requiring a metallocofactor (Nordlund & Reichard, 2006). In addition haem and chlorophyll biosynthesis could have appeared before vitamin B₁₂, with haem required for early terminal respiratory oxidases (Schafer *et al.*, 1996). Photosynthesis is believed to have evolved 3×10^9 years ago and sparked the rise in atmospheric oxygen levels (Tice & Lowe, 2006). It is speculated that sirohaem is

the most ancient of the tetrapyrroles. Not only is it the easiest to make biochemically, but it is also structurally the simplest (Warren & Scott, 1990). Based on this theory the next part of this discussion shall now carefully link some of the evolutionary clues to the findings of this research and provide a possible evolutionary theory for vitamin B₁₂ biosynthesis.

The most significant clue is the sirohaem pathway. This pathway is considered the oldest tetrapyrrole pathway, based on its biosynthetic simplicity and its requirement for NiR/SiR enzymes. The sirohaem pathway could have provided a foundation stage from which the vitamin B₁₂ pathway could evolve. The initial steps (C-2, C-7 methylation, ring oxidation, chelation) are shared between the sirohaem, anaerobic vitamin B₁₂ pathways (Leech, 2002) and the recently discovered alternative haem pathway (Storbeck *et al.*, 2010). This is also represented at the genetic level, although the iron chelation step is catalysed non-enzymatically (Professor Martin Warren, personal communication). It is important to highlight that precorrin-2 does not chelate metals very well whereas, in contrast, factor II chelates cobalt, iron, magnesium, nickel and copper with greater efficiency, as the ring system is more planar (PhD Thesis: Leech, 2004). Whether a precorrin-2 dehydrogenase was originally required is unknown, but it should be considered that this reaction can occur spontaneously even under anaerobic conditions. It is likely that metal chelation remained non-enzymatic before the chelatase enzymes (SirB, CbiX and CbiK) first appeared. As iron is more predominant in the environment than cobalt, it is likely this would determine iron-factor II (sirohaem) as the major product, with cobalt-factor II a minor side product. Eventually through gene duplication or fusion a NiR/SiR enzyme would evolve that could recognise iron-factor II (sirohaem), with its higher concentration favouring its selection over cobalt-factor II. However cobalt chelation would still have been an advantage to the cell as cobalt is toxic, even at low concentrations (Fantino *et al.*, 2010; Thorgersen & Downs, 2007).

The next steps are a bit trickier to predict. The most noticeable clue to the evolutionary origin of vitamin B₁₂ lies with the similarity shared between the class III methyltransferases, found in both pathways. It has previously been proposed that an ancestral methyltransferase gave birth to more regioselective methyltransferases, through the process of gene duplication and specialisation (Holliday *et al.*, 2007).

This could indicate that originally a single non-specific methyltransferase was responsible for all eight methylations. This theory supports the patchwork mechanism, whereby originally a primitive enzyme that lacked substrate specificity was replaced by a more efficient enzyme through gene duplication. These original starter enzymes would have existed in ancient cells that had simple genomes, estimated to be a minimum of 562 kb (Itaya, 1995). In addition these enzymes would produce unwanted side products (e.g factor S3), thus reducing yield and efficiency (Jensen, 1976). Therefore, mutations and gene duplications that aided the development of new enzymes that increased efficiency would provide a selective advantage.

The methylation events can be compared to the reactions that catalyse the amidations of six of the surrounding carboxylic acid side chains (CobB/CbiA and CobQ/CbiP, in their respective pathways). Although there is a preferred order of amidations, this is not always strictly followed (Fresquet *et al.*, 2007), suggesting that these reactions were once random. Interestingly the amidation reactions occur at a much later stage in the pathway when the intermediates are stable. However the methylation events occur in the more redox sensitive stages of both pathways. It is likely that the methylations originally developed to provide the corrin ring with stability from oxidation and allow it to form an appropriate conjugation state (Eschenmoser & Wintner, 1977). This requirement for stability could have driven the evolution of specific methyltransferases to increase efficiency and quickly process these labile intermediates through the pathway. Moreover ring contraction would be essential to tightly impound the cobalt within the macrocycle and provide extra stability. The findings from this research indicate this reaction favours a strong reductive process. This could have occurred spontaneously during the early stages of pathway development, as the early microbial world would have been strongly reducing in the absence of oxygen. The presence of the cobalt ion is likely to act only as a receptor for the enzymes to recognise its presence.

The next stage of pathway development would require a driving force for the positive selection of an ancient ancestor of vitamin B₁₂. This structure must have been simpler, with the lower ligand DMB and the nucleotide perhaps added at a later stage. It is impossible to predict what enzyme evolved to give the cells a

selective advantage. However, it is likely that some form of rearrangement reaction may have been useful, with a reaction that benefited carbon or nitrogen utilisation initially providing this driving force. If this was the original framework for a functional B₁₂ like pathway, then patchwork evolution can be applied to systematically improve efficiency and energy conservation. Over time, gene duplications and acquisition of novel enzyme functions (horizontal transfer, gene fusion events) would replace unfavourable non-enzymatic events. It is understandable that CbiT is a unique methyltransferase in the pathway, as this also catalyses a decarboxylation event. However, why did the pathway acquire a unique methyltransferase in CbiD? If the C-1 methylation is examined closer, it occurs at an unfavourable position on the macrocycle, presumably after the δ -lactone ring is removed by CbiG. The development or acquisition of a methyltransferase to catalyse this reaction may have been essential, as the C-1 methylation seems to provide extra stability, as shown in this work.

Next attention turns to the development of the aerobic pathway, presumably during the rise of oxygen in the atmosphere. This thesis draws a large assumption that aerobic pathway developed from the anaerobic pathway. It has been argued that the two pathways evolved separately, as enzymes such as CobI cannot replace the activity of CbiL in their respective pathways (Frank *et al.*, 2005; Spencer *et al.*, 1998). This was predicted as the enzymes can discriminate whether cobalt is inside the macrocycle. However, it is likely the rise of oxygen levels would have been the main driving force for the development of the aerobic pathway from its anaerobic counterpart. A transitory period may have occurred as the anaerobic pathway struggled to operate in organisms that were adapting to grow under rising oxygen levels. The link in evolution may be explained by the C-terminal extension of CbiH₆₀. Just from its similarity with the NiR/SiR family which is shared with CobG, this could be a connection in the evolution and development of the aerobic pathway from the anaerobic route. It could be proposed that the presence of this redox group might not be essential but instead allows the anaerobic ring contraction reaction to occur via an alternative mechanism or provide a driving force when the redox potential is raised in aerobic conditions. However further increased levels of oxygen may have driven a cataclysmic split in the pathways because of the oxygen sensitive nature of cobalt-organic intermediates. A new mechanism evolved using

molecular oxygen to replace the original anaerobic pathway mechanism. Over time the aerobic enzymes would lose their ability to recognise the cobalt ion. In addition enzymes that preferred non-cobalt intermediates were favoured as these would produce less side products. Other processes that were replaced or modified include the non-requirement for precorrin-2 oxidation and late cobalt insertion. This can be explained as factor intermediates readily chelate metals and to avoid this process it was essential this intermediate was not oxidised. This may have driven the assembly of the multifunctional CysG to separate sirohaem and vitamin B₁₂ biosynthesis further. In addition, further evidence for the later development of the aerobic pathway is provided by the late cobalt insertion step. The multi-enzyme complex of CobNST catalyses this reaction and shares remarkable resemblance to the magnesium chelatase, found in chlorophyll biosynthesis (Lundqvist *et al.*, 2009). It is highly likely that the genes encoding this were acquired through horizontal transfer and then modified to insert cobalt into a stable ring contracted intermediate, a reaction that is unfavourable and requires energy. Enzyme complexes are modular assemblies and are indicative of more recently acquired functions. This provides further evidence that the aerobic pathway evolved after the anaerobic route to vitamin B₁₂. A summary of this outline is provided in Figure 7.1.

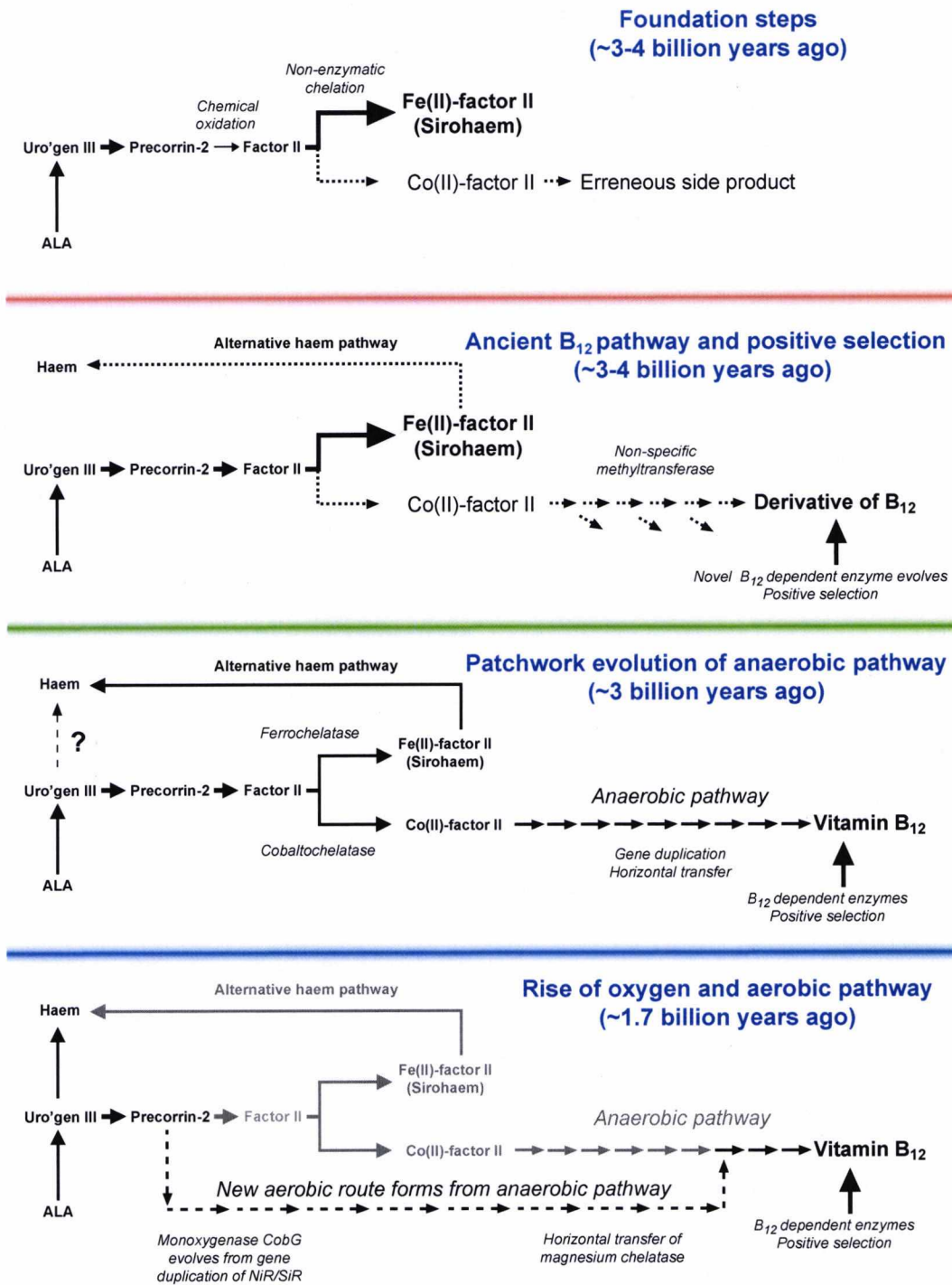


Figure 7.1. Theoretical model for tetrapyrrole pathway evolution from an ancestral semi-biosynthetic sirohaem pathway. For simplicity purposes chlorophyll, coenzyme F₄₃₀ and haem d₁ pathways are omitted.

Concluding remarks

The original ancestral anaerobic pathway may have required only a dozen enzymes along with some precursors. This has gradually has been replaced by over thirty enzymes through a mixture of gene duplication, gene fusion, horizontal transfer and modular assembly of enzymes to acquire new function. It should be considered that evolution is continuous, with our perception only a transitory conception of time. Therefore, in the next few millions years, Nature may yet find a new optimum route to the biosynthesis of vitamin B₁₂. However it is a question of whether simplicity or complexity is the answer to the future biosynthesis of vitamin B₁₂.

- Abe, R., Kudou, M., Tanaka, Y., Arakawa, T. & Tsumoto, K. (2009).** Immobilized metal affinity chromatography in the presence of arginine. *Biochem Biophys Res Commun* **381**, 306-310.
- Astner, I., Schulze, J. O., van den Heuvel, J., Jahn, D., Schubert, W. D. & Heinz, D. W. (2005).** Crystal structure of 5-aminolevulinate synthase, the first enzyme of heme biosynthesis, and its link to XLSA in humans. *EMBO J* **24**, 3166-3177.
- Balachandran, S., Vishwakarma, R. A., Monaghan, S. M., Prella, A., Stamford, P. J., Leeper, F. J. & Battersby, A. R. (1994).** *Journal of the Chemical Society, Perkins Transactions* **1**, 487.
- Banerjee, R. & Ragsdale, S. W. (2003).** The many faces of vitamin B₁₂: catalysis by cobalamin-dependent enzymes. *Annu Rev Biochem* **72**, 209-247.
- Battersby, A. R., Bushell, M. J., Jones, C., Lewis, N. G. & Pfenninger, A. (1981).** Biosynthesis of vitamin B₁₂: identity of fragment extruded during ring contraction to the corrin macrocycle. *Proc Natl Acad Sci U S A* **78**, 13-15.
- Battersby, A. R. (2000).** Tetrapyrroles: the pigments of life. *Nat Prod Rep* **17**, 507-526.
- Bernhard, K., Schrepf, H. & Goebel, W. (1978).** Bacteriocin and antibiotic resistance plasmids in *Bacillus cereus* and *Bacillus subtilis*. *J Bacteriol* **133**, 897-903.
- Biedendieck, R., Beine, R., Gamer, M., Jordan, E., Buchholz, K., Seibel, J., Dijkhuizen, L., Malten, M. & Jahn, D. (2007a).** Export, purification, and activities of affinity tagged *Lactobacillus reuteri* levansucrase produced by *Bacillus megaterium*. *Appl Microbiol Biotechnol* **74**, 1062-1073.
- Biedendieck, R., Yang, Y., Deckwer, W. D., Malten, M. & Jahn, D. (2007b).** Plasmid system for the intracellular production and purification of affinity-tagged proteins in *Bacillus megaterium*. *Biotechnol Bioeng* **96**, 525-537.
- Biedendieck, R., Malten, M., Barg, H., Bunk, B., Martens, J. H., Deery, E., Leech, H., Warren, M. J. & Jahn, D. (2010).** Metabolic engineering of cobalamin (vitamin B₁₂) production in *Bacillus megaterium*. *Microb Biotechnol* **3**, 24-37.
- Blanche, F., Debussche, L., Thibaut, D., Crouzet, J. & Cameron, B. (1989).** Purification and characterization of S-adenosyl-L-methionine: uroporphyrinogen III methyltransferase from *Pseudomonas denitrificans*. *J Bacteriol* **171**, 4222-4231.
- Blanche, F., Thibaut, D., Frechet, D. & other authors (1990).** Hydrogenobyric acid: isolation, biosynthesis, and function. *Angewandte Chemie International Edition* **29**, 884-886.
- Blanche, F., Couder, M., Debussche, L., Thibaut, D., Cameron, B. & Crouzet, J. (1991a).** Biosynthesis of vitamin B₁₂: stepwise amidation of carboxyl groups *b*, *d*,

e, and *g* of cobyrinic acid *a,c*-diamide is catalyzed by one enzyme in *Pseudomonas denitrificans*. *J Bacteriol* **173**, 6046-6051.

Blanche, F., Debussche, L., Famechon, A., Thibaut, D., Cameron, B. & Crouzet, J. (1991b). A bifunctional protein from *Pseudomonas denitrificans* carries cobinamide kinase and cobinamide phosphate guanylyltransferase activities. *J Bacteriol* **173**, 6052-6057.

Blanche, F., Famechon, A., Thibaut, D., Debussche, L., Cameron, B. & Crouzet, J. (1992a). Biosynthesis of vitamin B₁₂ in *Pseudomonas denitrificans*: the biosynthetic sequence from precorrin-6Y to precorrin-8X is catalyzed by the *cobL* gene product. *J Bacteriol* **174**, 1050-1052.

Blanche, F., Maton, L., Debussche, L. & Thibaut, D. (1992b). Purification and characterization of cob(II)yrinic acid *a,c*-diamide reductase from *Pseudomonas denitrificans*. *J Bacteriol* **174**, 7452-7454.

Blanche, F., Thibaut, D., Famechon, A., Debussche, L., Cameron, B. & Crouzet, J. (1992c). Precorrin-6x reductase from *Pseudomonas denitrificans*: purification and characterization of the enzyme and identification of the structural gene. *J Bacteriol* **174**, 1036-1042.

Borovok, I., Gorovitz, B., Schreiber, R., Aharonowitz, Y. & Cohen, G. (2006). Coenzyme B₁₂ controls transcription of the *Streptomyces* class Ia ribonucleotide reductase *nrdABS* operon via a riboswitch mechanism. *J Bacteriol* **188**, 2512-2520.

Brindley, A. A., Raux, E., Leech, H. K., Schubert, H. L. & Warren, M. J. (2003). A story of chelatase evolution: identification and characterization of a small 13-15-kDa "ancestral" cobaltochelatase (CbiX^S) in the archaea. *J Biol Chem* **278**, 22388-22395.

Butler, C. S., Charnock, J. M., Bennett, B. & other authors (1999). Models for molybdenum coordination during the catalytic cycle of periplasmic nitrate reductase from *Paracoccus denitrificans* derived from EPR and EXAFS spectroscopy. *Biochemistry* **38**, 9000-9012.

Cameron, B., Blanche, F., Rouyez, M. C. & other authors (1991). Genetic analysis, nucleotide sequence, and products of two *Pseudomonas denitrificans* cob genes encoding nicotinate-nucleotide: dimethylbenzimidazole phosphoribosyltransferase and cobalamin (5'-phosphate) synthase. *J Bacteriol* **173**, 6066-6073.

Cameron, B., Briggs, K., Pridmore, S., Brefort, G. & Crouzet, J (1989). Cloning and analysis of genes involved in coenzyme B₁₂ biosynthesis in *Pseudomonas denitrificans*. *Journal of Bacteriology* **171**, 547-557.

Christie, G. E., Farnham, P. J. & Platt, T. (1981). Synthetic sites for transcription termination and a functional comparison with tryptophan operon termination sites in vitro. *Proc Natl Acad Sci U S A* **78**, 4180-4184.

- Condon, C. (2007).** Maturation and degradation of RNA in bacteria. *Curr Opin Microbiol* **10**, 271-278.
- Crane, B. R. & Getzoff, E. D. (1996).** The relationship between structure and function for the sulfite reductases. *Curr Opin Struct Biol* **6**, 744-756.
- Crouzet, J., Cameron, B., Cauchois, L., Rigault, S., Rouyez, M. C., Blanche, F., Thibaut, D. & Debussche, L. (1990a).** Genetic and sequence analysis of an 8.7-kilobase *Pseudomonas denitrificans* fragment carrying eight genes involved in transformation of precorrin-2 to cobyrinic acid. *J Bacteriol* **172**, 5980-5990.
- Crouzet, J., Cauchois, L., Blanche, F., Debussche, L., Thibaut, D., Rouyez, M. C., Rigault, S., Mayaux, J. F. & Cameron, B. (1990b).** Nucleotide sequence of a *Pseudomonas denitrificans* 5.4-kilobase DNA fragment containing five cob genes and identification of structural genes encoding S-adenosyl-L-methionine: uroporphyrinogen III methyltransferase and cobyrinic acid *a,c*-diamide synthase. *J Bacteriol* **172**, 5968-5979.
- Crouzet, J., Levy-Schil, S., Cameron, B., Cauchois, L., Rigault, S., Rouyez, M. C., Blanche, F., Debussche, L. & Thibaut, D. (1991).** Nucleotide sequence and genetic analysis of a 13.1-kilobase-pair *Pseudomonas denitrificans* DNA fragment containing five *cob* genes and identification of structural genes encoding Cob(I)alamin adenosyltransferase, cobyrinic acid synthase, and bifunctional cobinamide kinase-cobinamide phosphate guanylyltransferase. *J Bacteriol* **173**, 6074-6087.
- Dann, C. E., 3rd, Wakeman, C. A., Sieling, C. L., Baker, S. C., Irnov, I. & Winkler, W. C. (2007).** Structure and mechanism of a metal-sensing regulatory RNA. *Cell* **130**, 878-892.
- Debussche, L., Couder, M., Thibaut, D., Cameron, B., Crouzet, J. & Blanche, F. (1991).** Purification and partial characterization of cob(I)alamin adenosyltransferase from *Pseudomonas denitrificans*. *J Bacteriol* **173**, 6300-6302.
- Debussche, L., Couder, M., Thibaut, D., Cameron, B., Crouzet, J. & Blanche, F. (1992).** Assay, purification, and characterization of cobaltochelataase, a unique complex enzyme catalyzing cobalt insertion in hydrogenobyrinic acid *a,c*-diamide during coenzyme B₁₂ biosynthesis in *Pseudomonas denitrificans*. *J Bacteriol* **174**, 7445-7451.
- Debussche, L., Thibaut, D., Cameron, B., Crouzet, J. & Blanche, F. (1993).** Biosynthesis of the corrin macrocycle of coenzyme B₁₂ in *Pseudomonas denitrificans*. *J Bacteriol* **175**, 7430-7440.
- Demain, A. L. & Vaishnav, P. (2009).** Production of recombinant proteins by microbes and higher organisms. *Biotechnol Adv* **27**, 297-306.
- Dickman, S. R. (1977).** Ribonucleotide reduction and the possible role of cobalamin in evolution. *J Mol Evol* **10**, 251-260.

- Drummond, J. T., Huang, S., Blumenthal, R. M. & Matthews, R. G. (1993).** Assignment of enzymatic function to specific protein regions of cobalamin-dependent methionine synthase from *Escherichia coli*. *Biochemistry* **32**, 9290-9295.
- Einsle, O., Messerschmidt, A., Huber, R., Kroneck, P. M. & Neese, F. (2002).** Mechanism of the six-electron reduction of nitrite to ammonia by cytochrome c nitrite reductase. *J Am Chem Soc* **124**, 11737-11745.
- Eppinger, M., Bunk, B., Johns, M. A. & other authors (2011).** Genome sequences of the biotechnologically important *B. megaterium* strains QM B1551 and DSM319. *J Bacteriol*.
- Escalante-Semerena, J. C., Suh, S. J. & Roth, J. R. (1990).** *cobA* function is required for both de novo cobalamin biosynthesis and assimilation of exogenous corrinoids in *Salmonella typhimurium*. *J Bacteriol* **172**, 273-280.
- Escalante-Semerena, J. C., Johnson, M. G. & Roth, J. R. (1992).** The CobII and CobIII regions of the cobalamin (vitamin B₁₂) biosynthetic operon of *Salmonella typhimurium*. *J Bacteriol* **174**, 24-29.
- Escalante-Semerena, J. C. (2007).** Conversion of cobinamide into adenosylcobamide in bacteria and archaea. *J Bacteriol* **189**, 4555-4560.
- Eschenmoser, A. & Wintner, C. E. (1977).** Natural product synthesis and vitamin B₁₂. *Science* **196**, 1410-1420.
- Eschenmoser, A. (1988).** Vitamin B₁₂ experiments concerning the origin of its molecular structure. *Angewandte Chemie International Edition England* **27**, 5-29.
- Fantino, J. R., Py, B., Fontecave, M. & Barras, F. (2010).** A genetic analysis of the response of *Escherichia coli* to cobalt stress. *Environ Microbiol* **12**, 2846-2857.
- Frank, S., Brindley, A. A., Deery, E., Heathcote, P., Lawrence, A. D., Leech, H. K., Pickersgill, R. W. & Warren, M. J. (2005).** Anaerobic synthesis of vitamin B₁₂: characterization of the early steps in the pathway. *Biochem Soc Trans* **33**, 811-814.
- Frank, S., Deery, E., Brindley, A. A. & other authors (2007).** Elucidation of substrate specificity in the cobalamin (vitamin B₁₂) biosynthetic methyltransferases. Structure and function of the C20 methyltransferase (CbiL) from *Methanothermobacter thermautotrophicus*. *J Biol Chem* **282**, 23957-23969.
- Fresquet, V., Williams, L. & Raushel, F. M. (2007).** Partial randomization of the four sequential amidation reactions catalyzed by cobyrinic acid synthetase with a single point mutation. *Biochemistry* **46**, 13983-13993.
- Fujii, K. & Huennekens, F. M. (1974).** Activation of methionine synthetase by a reduced triphosphopyridine nucleotide-dependent flavoprotein system. *J Biol Chem* **249**, 6745-6753.

- Gamer, M., Frode, D., Biedendieck, R., Stammen, S. & Jahn, D. (2009).** A T7 RNA polymerase-dependent gene expression system for *Bacillus megaterium*. *Appl Microbiol Biotechnol* **82**, 1195-1203.
- Gartner, D., Geissendorfer, M. & Hillen, W. (1988).** Expression of the *Bacillus subtilis* xyl operon is repressed at the level of transcription and is induced by xylose. *J Bacteriol* **170**, 3102-3109.
- Heinemann, I. U., Jahn, M. & Jahn, D. (2008).** The biochemistry of heme biosynthesis. *Arch Biochem Biophys*.
- Heldt, D., Lawrence, A. D., Lindenmeyer, M., Deery, E., Heathcote, P., Rigby, S. E. & Warren, M. J. (2005).** Aerobic synthesis of vitamin B₁₂: ring contraction and cobalt chelation. *Biochem Soc Trans* **33**, 815-819.
- Hodgkin, D. G., Pickworth, J., Robertson, J. H., Trueblood, K. N., Prosen, R. J. & White, J. G. (1955).** The crystal structure of the hexacarboxylic acid derived from B₁₂ and the molecular structure of the vitamin. *Nature* **176**, 325-328.
- Holliday, G. L., Thornton, J. M., Marquet, A., Smith, A. G., Rebeille, F., Mendel, R., Schubert, H. L., Lawrence, A. D. & Warren, M. J. (2007).** Evolution of enzymes and pathways for the biosynthesis of cofactors. *Nat Prod Rep* **24**, 972-987.
- Itaya, M. (1995).** An estimation of minimal genome size required for life. *FEBS Lett* **362**, 257-260.
- Jensen, R. A. (1976).** Enzyme recruitment in evolution of new function. *Annu Rev Microbiol* **30**, 409-425.
- Johansson, P. & Hederstedt, L. (1999).** Organization of genes for tetrapyrrole biosynthesis in Gram-positive bacteria. *Microbiology* **145** (Pt 3), 529-538.
- Kajiwara, Y., Santander, P. J., Roessner, C. A., Perez, L. M. & Scott, A. I. (2006).** Genetically engineered synthesis and structural characterization of cobalt-precorrin 5A and -5B, two new intermediates on the anaerobic pathway to vitamin B₁₂: definition of the roles of the CbiF and CbiG enzymes. *J Am Chem Soc* **128**, 9971-9978.
- Kazanov, M. D., Vitreschak, A. G. & Gelfand, M. S. (2007).** Abundance and functional diversity of riboswitches in microbial communities. *BMC Genomics* **8**, 347.
- Kim, I. K., Yim, Y. I., Kim, Y. M., Lee, J. W., Yim, H. S. & Kang, S. O. (2003).** CbiX-homologous protein (CbiX^{hp}), a metal-binding protein, from *Streptomyces seoulensis* is involved in expression of nickel-containing superoxide dismutase. *FEMS Microbiol Lett* **228**, 21-26.
- Komeda, H., Kobayashi, M. & Shimizu, S. (1997).** A novel transporter involved in cobalt uptake. *Proc Natl Acad Sci USA* **94**, 36-41.

- Kunkel, A., Marten, J-H., Jahn, D., Barg, H., and Warren, M.J. (2004).** Method for producing vitamin B₁₂. *United States Patent 2004/0235120 A1*.
- Kunnimalaiyaan, M. & Vary, P. S. (2005).** Molecular characterization of plasmid pBM300 from *Bacillus megaterium* QM B1551. *Appl Environ Microbiol* **71**, 3068-3076.
- Lawrence, A. D., Deery, E., McLean, K. J., Munro, A. W., Pickersgill, R. W., Rigby, S. E. & Warren, M. J. (2008).** Identification, characterization, and structure/function analysis of a corrin reductase involved in adenosylcobalamin biosynthesis. *J Biol Chem* **283**, 10813-10821.
- Leech, H., Raux-Deery, E., Heathcote, P., and Warren, M. J. (2002).** Production of cobalamin and sirohaem in *Bacillus megaterium*: an investigation into the role of the branchpoint chelataases sirohydrochlorin ferrochelataase (SirB) and sirohydrochlorin cobalt chelataase (CbiX). *Biochemical Society Transactions* **30**, 610-613.
- Leech, H. K., Raux, E., McLean, K. J. & other authors (2003).** Characterization of the cobaltochelataase CbiX^L: evidence for a 4Fe-4S center housed within an MXCXC motif. *J Biol Chem* **278**, 41900-41907.
- Levican, G., Katz, A., de Armas, M., Nunez, H. & Orellana, O. (2007).** Regulation of a glutamyl-tRNA synthetase by the heme status. *Proc Natl Acad Sci USA* **104**, 3135-3140.
- Luer, C., Schauer, S., Virus, S., Schubert, W. D., Heinz, D. W., Moser, J. & Jahn, D. (2007).** Glutamate recognition and hydride transfer by *Escherichia coli* glutamyl-tRNA reductase. *FEBS J* **274**, 4609-4614.
- Lundqvist, J., Elmlund, D., Heldt, D., Deery, E., Soderberg, C. A., Hansson, M., Warren, M. & Al-Karadaghi, S. (2009).** The AAA(+) motor complex of subunits CobS and CobT of cobaltochelataase visualized by single particle electron microscopy. *J Struct Biol* **167**, 227-234.
- Maggio-Hall, L. A. & Escalante-Semerena, J. C. (1999).** *In vitro* synthesis of the nucleotide loop of cobalamin by *Salmonella typhimurium* enzymes. *Proc Natl Acad Sci USA* **96**, 11798-11803.
- Malten, M., Biedendieck, R., Gamer, M., Drews, A. C., Stammen, S., Buchholz, K., Dijkhuizen, L. & Jahn, D. (2006).** A *Bacillus megaterium* plasmid system for the production, export, and one-step purification of affinity-tagged heterologous levansucrase from growth medium. *Appl Environ Microbiol* **72**, 1677-1679.
- Mandal, M. & Breaker, R. R. (2004).** Gene regulation by riboswitches. *Nat Rev Mol Cell Biol* **5**, 451-463.
- Martens, J. H., Barg, H., Warren, M. J. & Jahn, D. (2002).** Microbial production of vitamin B₁₂. *Appl Microbiol Biotechnol* **58**, 275-285.

- Martens, J. H., Barg, H., Warren, M. J. & Jahn, D (2002).** Microbial production of vitamin B₁₂. *Applied Microbiology and Biotechnology* **58**, 275-285.
- Minot, G. R. & Murphy, W. P. (1926).** *JAMA, The Journal of the American Medical Association* **87**, 470.
- Mombelli, L., Nussbaumer, C., Weber, H., Muller, G. & Arigoni, D. (1981).** Biosynthesis of vitamin B₁₂: nature of the volatile fragment generated during formation of the corrin ring system. *Proc Natl Acad Sci U S A* **78**, 11-12.
- Monedero, V., Poncet, S., Mijakovic, I., Fieulaine, S., Dossonnet, V., Martin-Verstraete, I., Nessler, S. & Deutscher, J. (2001).** Mutations lowering the phosphatase activity of HPr kinase/phosphatase switch off carbon metabolism. *EMBO J* **20**, 3928-3937.
- Moser, J., Schubert, W. D., Beier, V., Bringemeier, I., Jahn, D. & Heinz, D. W. (2001).** V-shaped structure of glutamyl-tRNA reductase, the first enzyme of tRNA-dependent tetrapyrrole biosynthesis. *EMBO J* **20**, 6583-6590.
- Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L. & Breaker, R. R. (2002).** Genetic control by a metabolite binding mRNA. *Chem Biol* **9**, 1043.
- Navarro, C., Wu, L. F. & Mandrand-Berthelot, M. A. (1993).** The *nik* operon of *Escherichia coli* encodes a periplasmic binding-protein-dependent transport system for nickel. *Mol Microbiol* **9**, 1181-1191.
- Neuberger, A. & Scott, J. J. (1953).** Aminolaevulinic acid and porphyrin biosynthesis. *Nature* **172**, 1093-1094.
- Nordlund, P. & Reichard, P. (2006).** Ribonucleotide reductases. *Annu Rev Biochem* **75**, 681-706.
- Nussbaumer, C., Imfeld, M., Worner, G., Muller, G. & Arigoni, D. (1981).** Biosynthesis of vitamin B₁₂: mode of incorporation of factor III into cobyrinic acid. *Proc Natl Acad Sci U S A* **78**, 9-10.
- Ollagnier-De Choudens, S., Sanakis, Y., Hewitson, K. S., Roach, P., Baldwin, J. E., Munck, E. & Fontecave, M. (2000).** Iron-sulfur center of biotin synthase and lipoate synthase. *Biochemistry* **39**, 4165-4173.
- Olteanu, H. & Banerjee, R. (2001).** Human methionine synthase reductase, a soluble P-450 reductase-like dual flavoprotein, is sufficient for NADPH-dependent methionine synthase activation. *J Biol Chem* **276**, 35558-35563.
- Ozaki, S., I., Roessner, C. A., Stolowich, N. J., Atshaves, B. P., Hertle, R., Muller, G. & Scott, A. I. (1993).** Multienzyme synthesis and structure of factor S3. *Journal of the American Chemical Society* **115**, 7935-7938.

- Perkins, D. N., Pappin, D. J., Creasy, D. M. & Cottrell, J. S. (1999).** Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551-3567.
- Pfaltz, A., Kobelt, A., Huster, R. & Thauer, R. K. (1987).** Biosynthesis of coenzyme F₄₃₀ in methanogenic bacteria. Identification of 15,17(3)-seco-F₄₃₀-17(3)-acid as an intermediate. *Eur J Biochem* **170**, 459-467.
- PhD Thesis: Barg, H. (2003).** Genetic optimisation of *Bacillus megaterium* for vitamin B₁₂ production.
- PhD Thesis: Frank, S. (2007).** Biosynthesis of cobalamin (Vitamin B₁₂): Characterisation of the anaerobic pathway in *Methanobacter thermoautotrophicus*.
- PhD Thesis: Leech, H. K. (2004).** The biosynthesis of sirohydrochlorin and its transformation into sirohaem and vitamin B₁₂.
- PhD Thesis: Malten, M. (2005).** Protein production and secretion in *Bacillus megaterium*.
- Raux, E., Lanois, A., Levillayer, F., Warren, M. J., Brody, E., Rambach, A. & Thermes, C. (1996).** Salmonella typhimurium cobalamin (vitamin B₁₂) biosynthetic genes: functional studies in *S. typhimurium* and *Escherichia coli*. *J Bacteriol* **178**, 753-767.
- Raux, E., Lanois, A., Rambach, A., Warren, M. J. & Thermes, C. (1998a).** Cobalamin (vitamin B₁₂) biosynthesis: functional characterization of the *Bacillus megaterium* *cbi* genes required to convert uroporphyrinogen III into cobyrinic acid *a,c*-diamide. *Biochem J* **335** (Pt 1), 167-173.
- Raux, E., Lanois, A., Warren, M. J., Rambach, A. & Thermes, C. (1998b).** Cobalamin (vitamin B₁₂) biosynthesis: identification and characterization of a *Bacillus megaterium* *cobI* operon. *Biochem J* **335** (Pt 1), 159-166.
- Raux, E., Schubert, H. L. & Warren, M. J. (2000).** Biosynthesis of cobalamin (vitamin B₁₂): a bacterial conundrum. *Cell Mol Life Sci* **57**, 1880-1893.
- Raux, E., Leech, H. K., Beck, R. & other authors (2003).** Identification and functional analysis of enzymes required for precorrin-2 dehydrogenation and metal ion insertion in the biosynthesis of sirohaem and cobalamin in *Bacillus megaterium*. *Biochem J* **370**, 505-516.
- Raux, E., Schubert, H. L., Roper, J. M., Wilson, K. S., and Warren, M. J. (1999).** Vitamin B₁₂: Insights into Biosynthesis's Mount Improbable. *Bioorganic Chemistry* **27**, 100-118.
- Ravnum, S. & Andersson, D. I. (1997).** Vitamin B₁₂ repression of the *btuB* gene in *Salmonella typhimurium* is mediated via a translational control which requires leader and coding sequences. *Mol Microbiol* **23**, 35-42.

- Rickes, E. L., Brink, N. G., Koniuszy, F. R., Wood, T. R. & Folkers, K. (1948). *Science* **107**, 396.
- Riemer, J., Hoepken, H. H., Czerwinska, H., Robinson, S. R. & Dringen, R. (2004). Colorimetric ferrozine-based assay for the quantitation of iron in cultured cells. *Anal Biochem* **331**, 370-375.
- Robin, C., Blanche, F., Cauchois, L., Cameron, B., Couder, M. & Crouzet, J. (1991). Primary structure, expression in *Escherichia coli*, and properties of S-adenosyl-L-methionine:uroporphyrinogen III methyltransferase from *Bacillus megaterium*. *J Bacteriol* **173**, 4893-4896.
- Roessner, C. A., Spencer, J. B., Stolowich, N. J. & other authors (1994). Genetically engineered synthesis of precorrin-6B and the complete corrinoid, hydrogenobyric acid, an advanced precursor of vitamin B₁₂. *Chem Biol* **1**, 119-124.
- Roessner, C. A., Williams, H. J. & Scott, A. I. (2005). Genetically engineered production of 1-desmethylcobyrinic acid, 1-desmethylcobyrinic acid *a,c*-diamide, and cobyrinic acid *a,c*-diamide in *Escherichia coli* implies a role for CbiD in C-1 methylation in the anaerobic pathway to cobalamin. *J Biol Chem* **280**, 16748-16753.
- Roessner, C. A. & Scott, A. I. (2006). Fine-tuning our knowledge of the anaerobic route to cobalamin (vitamin B₁₂). *J Bacteriol* **188**, 7331-7334.
- Rosso, M. L. & Vary, P. S. (2005). Distribution of *Bacillus megaterium* QM B1551 plasmids among other *B. megaterium* strains and *Bacillus* species. *Plasmid* **53**, 205-217.
- 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.
- Rygas, T. & Hillen, W. (1991). Inducible high-level expression of heterologous genes in *Bacillus megaterium* using the regulatory elements of the xylose-utilization operon. *Appl Microbiol Biotechnol* **35**, 594-599.
- Rygas, T., Scheler, A., Allmansberger, R. & Hillen, W. (1991). Molecular cloning, structure, promoters and regulatory elements for transcription of the *Bacillus megaterium* encoded regulon for xylose utilization. *Arch Microbiol* **155**, 535-542.
- Rygas, T. & Hillen, W. (1992). Catabolite repression of the xyl operon in *Bacillus megaterium*. *J Bacteriol* **174**, 3049-3055.
- Sambrook, J. & Frisch, E. F. (1989). Molecular cloning: A laboratory manual, 2nd edition. *Cold Spring Harbour Laboratory*.

- Santander, P. J., Roessner, C. A., Stolowich, N. J., Holderman, M. T. & Scott, A. I. (1997).** How corrinoids are synthesized without oxygen: nature's first pathway to vitamin B₁₂. *Chem Biol* **4**, 659-666.
- Santillan, M. & Mackey, M. C. (2005).** Dynamic behaviour of the B₁₂ riboswitch. *Phys Biol* **2**, 29-35.
- Schafer, G., Purschke, W. & Schmidt, C. L. (1996).** On the origin of respiration: electron transport proteins from archaea to man. *FEMS Microbiol Rev* **18**, 173-188.
- Schilling, O., Langbein, I., Muller, M., Schmalisch, M. H. & Stulke, J. (2004).** A protein-dependent riboswitch controlling *ptsGHI* operon expression in *Bacillus subtilis*: RNA structure rather than sequence provides interaction specificity. *Nucleic Acids Res* **32**, 2853-2864.
- Schroeder, S., Lawrence, A. D., Biedendieck, R. & other authors (2009).** Demonstration that CobG, the monooxygenase associated with the ring contraction process of the aerobic cobalamin (vitamin B₁₂) biosynthetic pathway, contains an Fe-S center and a mononuclear non-heme iron center. *J Biol Chem* **284**, 4796-4805.
- Schubert, H. L., Wilson, K. S., Raux, E., Woodcock, S. C. & Warren, M. J. (1998).** The X-ray structure of a cobalamin biosynthetic enzyme, cobalt-precorrin-4 methyltransferase. *Nat Struct Biol* **5**, 585-592.
- Schubert, H. L., Rose, R. S., Leech, H. K., Brindley, A. A., Hill, C. P., Rigby, S. E. & Warren, M. J. (2008).** Structure and function of SirC from *Bacillus megaterium*: a metal-binding precorrin-2 dehydrogenase. *Biochem J* **415**, 257-263.
- Scott, A. I., Williams, H. J., Stolowich, N. J. & other authors (1989).** *Journal of the American Chemical Society* **111**, 1897.
- Scott, A. I., Stolowich, N. J., Wang, J., Gawatz, O., Fridrich, E. & Muller, G. (1996).** Biosynthesis of vitamin B₁₂: factor IV, a new intermediate in the anaerobic pathway. *Proc Natl Acad Sci U S A* **93**, 14316-14319.
- Scott, A. I. (2003).** Discovering nature's diverse pathways to vitamin B₁₂: a 35-year odyssey. *J Org Chem* **68**, 2529-2539.
- Serganov, A., Polonskaia, A., Phan, A. T., Breaker, R. R. & Patel, D. J. (2006).** Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature* **441**, 1167-1171.
- Shemin, D., Russell, C. S. & Abramsky, T. (1955).** The succinate-glycine cycle. I. The mechanism of pyrrole synthesis. *J Biol Chem* **215**, 613-626.
- 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.
- Smith, E. L. (1948).** Presence of cobalt in the anti-pernicious anaemia factor. *Nature* **162**, 144.

Spencer, J. B., Stolowich, N.J., Santander, P.J., Pichon, C., Kajiwara, M., Tokiwa, S., Takatori, K. and Scott, A.I. (1994). *J Am Chem Soc* **116**, 4991-4992.

Spencer, P., Stolowich, N. J., Sumner, L. W. & Scott, A. I. (1998). Definition of the redox states of cobalt-precorrinoids: investigation of the substrate and redox specificity of CbiL from *Salmonella typhimurium*. *Biochemistry* **37**, 14917-14927.

Stammen, S., Muller, B. K., Korneli, C., Biedendieck, R., Gamer, M., Franco-Lara, E. & Jahn, D. (2010). High-yield intra- and extracellular protein production using *Bacillus megaterium*. *Appl Environ Microbiol* **76**, 4037-4046.

Storbeck, S., Rolfes, S., Raux-Deery, E., Warren, M. J., Jahn, D. & Layer, G. (2010). A novel pathway for the biosynthesis of heme in Archaea: genome-based bioinformatic predictions and experimental evidence. *Archaea* **13**, 175050.

Sudarsan, N., Cohen-Chalamish, S., Nakamura, S., Emilsson, G. M. & Breaker, R. R. (2005). Thiamine pyrophosphate riboswitches are targets for the antimicrobial compound pyrithiamine. *Chem Biol* **12**, 1325-1335.

Thibaut, D., Blanche, F., Debussche, L., Leeper, F. J. & Battersby, A. R. (1990a). Biosynthesis of vitamin B₁₂: structure of precorrin-6x octamethyl ester. *Proc Natl Acad Sci U S A* **87**, 8800-8804.

Thibaut, D., Couder, M., Crouzet, J., Debussche, L., Cameron, B. & Blanche, F. (1990b). Assay and purification of *S*-adenosyl-L-methionine:precorrin-2 methyltransferase from *Pseudomonas denitrificans*. *J Bacteriol* **172**, 6245-6251.

Thibaut, D., Debussche, L. & Blanche, F. (1990c). Biosynthesis of vitamin B₁₂: isolation of precorrin-6x, a metal-free precursor of the corrin macrocycle retaining five *S*-adenosylmethionine-derived peripheral methyl groups. *Proc Natl Acad Sci U S A* **87**, 8795-8799.

Thibaut, D., Couder, M., Famechon, A., Debussche, L., Cameron, B., Crouzet, J. & Blanche, F. (1992). The final step in the biosynthesis of hydrogenobyric acid is catalyzed by the *cobH* gene product with precorrin-8X as the substrate. *J Bacteriol* **174**, 1043-1049.

Thorgersen, M. P. & Downs, D. M. (2007). Cobalt targets multiple metabolic processes in *Salmonella enterica*. *J Bacteriol* **189**, 7774-7781.

Tice, M. M. & Lowe, D. R. (2006). Hydrogen-based carbon fixation in the earliest known photosynthetic organisms. *Geology* **34**, 37.

Trott, O. & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**, 455-461.

Uzar, H. C., Battersby, A. R., Carpenter, T. A. & Leeper, F. J. (1987). *Journal of the Chemical Society, Perkins Transactions*, 1689.

- Vary, P. S. (1994). Prime time for *Bacillus megaterium*. *Microbiology* **140** (Pt 5), 1001-1013.
- Vary, P. S., Biedendieck, R., Fuerch, T., Meinhardt, F., Rohde, M., Deckwer, W. D. & Jahn, D. (2007). *Bacillus megaterium*: from simple soil bacterium to industrial protein production host. *Appl Microbiol Biotechnol* **76**, 957-967.
- Vevodova, J., Graham, R. M., Raux, E., Warren, M. J. & Wilson, K. S. (2005). Crystallization and preliminary structure analysis of CobE, an essential protein of cobalamin (vitamin B₁₂) biosynthesis. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **61**, 442-444.
- Wadahama, H., Kamauchi, S., Ishimoto, M., Kawada, T. & Urade, R. (2007). Protein disulfide isomerase family proteins involved in soybean protein biogenesis. *FEBS J* **274**, 687-703.
- Wadia, R. S., Bandishti, S., and Kharche, M. (2000). B₁₂ and folate deficiency incidence and clinical features. *Neurology India* **48**, 302-304.
- Waldron, K. J., Rutherford, J. C., Ford, D. & Robinson, N. J. (2009). Metalloproteins and metal sensing. *Nature* **460**, 823-830.
- Wang, J., Stolowich, N. J., Santander, P. J., Park, J. H. & Scott, A. I. (1996). Biosynthesis of vitamin B₁₂: concerning the identity of the two-carbon fragment eliminated during anaerobic formation of cobyrinic acid. *Proc Natl Acad Sci U S A* **93**, 14320-14322.
- Wang, W., Hollmann, R., Furch, T., Nimtz, M., Malten, M., Jahn, D. & Deckwer, W. D. (2005). Proteome analysis of a recombinant *Bacillus megaterium* strain during heterologous production of a glucosyltransferase. *Proteome Sci* **3**, 4.
- Warburg, O., and Christian, W. (1942). Isolation and crystallization of enolase. *Biochemistry Z* **310**, 384-421.
- Warren, M. J., Roessner, C. A., Santander, P. J. & Scott, A. I. (1990a). The *Escherichia coli cysG* gene encodes S-adenosylmethionine-dependent uroporphyrinogen III methylase. *Biochem J* **265**, 725-729.
- Warren, M. J. & Scott, A. I. (1990). Tetrapyrrole assembly and modification into the ligands of biologically functional cofactors. *Trends Biochem Sci* **15**, 486-491.
- Warren, M. J., Stolowich, N. J., Santander, P. J., Roessner, C. A., Sowa, B. A. & Scott, A. I. (1990b). Enzymatic synthesis of dihydrosirohydrochlorin (precorrin-2) and of a novel pyrocorphin by uroporphyrinogen III methylase. *FEBS Lett* **261**, 76-80.
- Warren, M. J., Bolt, E. L., Roessner, C. A., Scott, A. I., Spencer, J. B. & Woodcock, S. C. (1994). Gene dissection demonstrates that the *Escherichia coli cysG* gene encodes a multifunctional protein. *Biochem J* **302** (Pt 3), 837-844.

Warren, M. J., Raux, E., Schubert, H. L. & Escalante-Semerena, J. C. (2002). The biosynthesis of adenosylcobalamin (vitamin B₁₂). *Nat Prod Rep* **19**, 390-412.

Winkler, W. C., Cohen-Chalamish, S. & Breaker, R. R. (2002). An mRNA structure that controls gene expression by binding FMN. *Proc Natl Acad Sci U S A* **99**, 15908-15913.

Wolfram, L., Friedrich, B. & Eitinger, T. (1995). The *Alcaligenes eutrophus* protein HoxN mediates nickel transport in *Escherichia coli*. *J Bacteriol* **177**, 1840-1843.

Woodson, J. D., Zayas, C. L. & Escalante-Semerena, J. C. (2003). A new pathway for salvaging the coenzyme B₁₂ precursor cobinamide in archaea requires cobinamide-phosphate synthase (CbiB) enzyme activity. *J Bacteriol* **185**, 7193-7201.

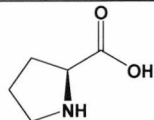
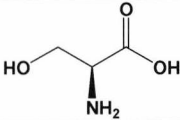
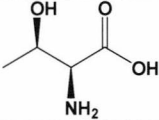
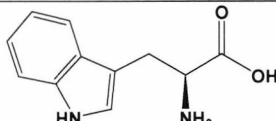
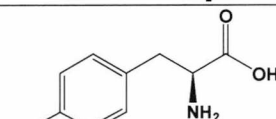
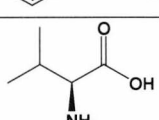
Woodward, R. B. (1973). The total synthesis of vitamin B₁₂. *Pure Appl Chem* **33**, 145-177.

APPENDICES

Appendix I – Amino acid, abbreviations and structure

Amino acid	Three letter code	Single letter code	Structure
Alanine	Ala	A	
Arginine	Arg	R	
Asparagine	Asn	N	
Aspartate	Asp	D	
Cysteine	Cys	C	
Glutamate	Glu	E	
Glutamine	Gln	Q	
Glycine	Gly	G	
Histidine	His	H	
Isoleucine	Ile	I	
Leucine	Leu	L	
Lysine	Lys	K	
Methionine	Met	M	
Phenylalanine	Phe	F	

Appendix I – Amino acid, abbreviations and structure

Amino acid	Three letter code	Single letter code	Structure
Proline	Pro	P	
Serine	Ser	S	
Threonine	Thr	T	
Tryptophan	Trp	W	
Tyrosine	Tyr	Y	
Valine	Val	V	

Appendix II – *B. megaterium* DSM509 B₁₂ anaerobic pathway enzymes**Motifs**

4Fe-4S cluster	(C)
Polyhistidine rich-motif	(H)
SAM binding domain	(G)
Putative SAM/tetrapyrrole binding site	(X)

CbiX^L (Factor II cobaltochelatase)

MKSVLFVGHGSRDPEGNDQIREFISTMKHDWDASILVETCFLEFERPNVSQGIDTC
 VAKGAQDVVVIPIMLLPAGHSKIHIIPAAIDEAKEKYPHVNFVYGRPIGVHEEALEI
 LKTRLQESGENLETPAEDTAVIVLGRGSDPDANSPLYKITRLLWEKTNKIVETS
 FMGVTAPLIDEGVERCLKLGAKKVVILPYFLFTGVLIKRL EEMVKQYKMQHENIEF
 KLAGYFGFHPKLQTIKERAEEGLEGEVKMNCDT CQYRLGIMEHID HHHHH DHDH
 HGH HHH DHHH DHDH HEDKVGELK

CbiL (Cobalt-factor II C-20 methyltransferase)

MNMIGTLYGLGV GPG DPELITVKAFRKLKESPVIAYPKKQKGSKSYAQKIIDVYFS
 ANEKDMLGLVFPMTKDPAILERKWTETVERVWEKLQEGKDVAFVTE G D P M L Y S T F I
 HMMRLMQERHPEAPIQVIPIGISSINGAASRLGIPLADGDEHVAIVPAREDYETMKK
 VLMENDCVIFIKVAKVIDFMVGLLKELDLLEKASVVTKVTSDEEIIWKASELEGAD
 LQYLTLMVVRK

CbiH₆₀ (Cobalt-factor III C-17 methyltransferase, ring contraction enzyme)

MKGKLLVI G F G P G S F E H I T Q R A R E A I Q E S D M I I G Y K T Y V E L I Q G L L T N Q Q I I S T G M
 TEEVSRAQEAVKQAEAGKTVAVISS G D A G V Y G M A G L V Y E V L I E K G W K K E T G V E L E V
 I P G I S A I N S C A S L L G A P V M H D A C T I S L S D H L T P W E L I E K R I E A A A Q A D F V V A F Y N P
 KSGRRTRQIVEAQRILLKYRSPDTPVGLVKSAYRDREEVMTNLKMDLNHEIGMLT
 TVVVGNSSTFFYDDLMITPRGYQRKYTLNQTEQPLRPHQRLRKEAEPWALDQEETV
 KQTASAIEAVQTTPEETATSRALAEALQAILGEPTS AVVHQPIQSIFEVAVSPGL
 ANKKFTPVQMTTLAEVVGEEKTMEYTPDHQIKLQIPTTQPDV I I K K L Q A A S F L L S P
 VGDVFT I K A C D F C D G E K S D A I P H T E E L Q K R L G G M D M P K E L K L G I N G C G M A C Y G A V Q
 EDIGIVYRKGAFLFLGAKTVGRNAHSGQIVAEG I A P D D I V E I V E N I I H E Y K E K G H
 PNERFHKFFKRVKNVYGFYQDITPKIKVEPAPCGD

CbiF (Cobalt-precorrin-4 C-11 methyltransferase)

MKLYI I G A G P G D P D L I T V K G L K L L Q Q A D V V L Y A D S L V S Q E L V A K S K P G A E V L K T A G
 MHLEEMVGTMLERMREGKMVVRVHT G D P A M Y G A I M E Q M V L L K R E G V D I E I V P G V S S
 VFAAAAAEAELTIPDLTQTVILTRAEGRTVPPEFEKLTDLAKHKCTIALFLSATL
 TKKVMKEFISAGWSEDTPVVVVYKATWPDEKIVRTTVKDLDDAMRTNGIRKQAMIL
 AGWALDPHIHDKDYRSKLYDKTFTHGFRGK

CbiG (Cobalt-precorrin-4 deacylase)

MYMSKFEKGDQEQQNIQLFSGSVRMLLPSLFESYKGLIIIIISLGAVVRMIAPILK
 DKKTDPVVVIDDKGENVISVLSGHIGGANELTREVA AALEAHPVITTASDVQKTI
 PVDLFGKRFGVWWEAEKLT PVSASVVNEEEI AVVQESGEKNWWHYEHPVPANIKT
 YSSIQTALASPHAAALVTHRNLKKEEAILENGVLYRPKVLAIGMGCNRTSTAE
 IETVIEKTLAELQFSMKSVKALCTIELKKDEEGLLEVASKYGWEFVYYSPELNSI
 SIQQPSD TVFKYTGAYGVSEPAVM LYS GADKLELVKKKSGNVTISVALIPYD

CbiD (Cobalt-precorrin-5 C-1 methyltransferase)

MKEVPKEPKKLRGYTTGACATAATRAALLTLISGEVQDESTIYLPVGRFATFHLE
 ECEYRTSSAVASVIKDAGDDPDATHGALIISEVSWCNGAGI IIDGGVGVGRVTKPG
 LPVPVGEAAINPVPRKMLKETAEQLLAEYNIQKGVKIVISVPEGEEMAKKTLNARL
 GILGGISILGTRGIVVPFSTAAKASIVQAI SVAKASNCEHVVITTTGGRSEKYGMK
 QFSELPEEAFIEMGDFVGFLLKQCKKQGMKKVSLVGMGKFSKVAQGVMMVHSSKA
 PIDFNFLAKAASESGASAELVEEIKGANTASQVGDLMTQSGHHQFFEKLCYCCLS
 ALKEVGDGIDVDTSLYTLKGDFLGQAVQHGN

CbiJ (Cobalt-precorrin-6A NADH dependent corrin ring reductase)

MILLLAGTSDARALAVQVKKAGHEVTATVVTENAAIELQRAEVKVKIGRLTKEDMM
 DFINEHRVKAIVDASHPFAEEASKNAIGAAAETAIPYIRYERASQAFAYDNMTMVS
 TYEEAAEVAAEKKGVI MLTTGSKTLQVFTEKLLSLSDVRLVARMLPRLDNMEKCQQ
 LGFPQKNIIAIQGPFTKEFDRALYKQYGVTVMVTKESGKVGSVDEKVEAAKELGLD
 IIMIGRPKIEYGTVYSTFEVVHALVNQNR

CbiET (Cobalt-precorrin-6B C-5, C-15 methyltransferase and decarboxylase)

MAIKI I **GIG** DDGKLSLLPMYEKWIYESDVLVGGKRHLDDFFQDFKGEKVAIEGGLFS
 LVERLKNEEGNVVVLAS **GD** PLF **Y** GIGSYLSTKLDVEIYPYLSSIQLAFSRLKERWQ
 DAYFTSVHGRS I KGLAQRIDGHKKVAILTDEQNSPSALANYLLSFGMTEYKMFVAE
 NLGGETERCQLLSLEEAAANQLFSPLNVVILKQVEKSPVWPLGIEDDEFIQRPDKG
 LITKKEVRTLSISALQLKKDSVVDIGTCTGSVAIEAAKIAREGQIFAVEKNEADL
 ENCRENLLKFRVDAHTVHGKAPEGLSEFADPDAVFIGGTAGGMETILDVCCSRLNA
 GGRIVLNAVTVENLAEAMKAFKERGFETAVTLAQISRSPILHLTRFDALNPIYII
 TAKRGE

CbiC (Cobalt-precorrin-8X methylmutase)

MIEMDFRTEFKPLTVQPQQIEGKSFEMITEELGPHPTDEQYPIVQRVIHASADFE
 LGRSMLFHPDAIQAGIKAIRSGKQVVADVQMVQVGTNKQRIEKHGGEIKVYISDPD
 VMEEAKRLNTRAIISMRKAIKEADGGIFAIGNAPTALLELIRLIKEGEAKPGLVI
 GLPVGFVSAAESKEELAKLDVPFITNIGRKGGSTVTVAALNAISILADQV

Appendix III – Molecular dimensions, structure screen 1

	Salt	Buffer	pH	Precipitant
1	0.02 M CaCl ₂	0.1 M Na acetate	4.6	30% MPD
2	0.2 M (NH ₄) ₂ acetate	0.1 M Na acetate	4.6	30% PEG 4K
3	0.2 M (NH ₄) ₂ sulfate	0.1 M Na acetate	4.6	25% PEG 4K
4	None	0.1 M Na acetate	4.6	2.0 M Na formate
5	None	0.1 M Na acetate	4.6	2.0 M (NH ₄) ₂ sulfate
6	None	0.1 M Na acetate	4.6	8% PEG 4K
7	0.2 M (NH ₄) ₂ acetate	0.1 M Na citrate	5.6	30% PEG 4K
8	0.2 M (NH ₄) ₂ acetate	0.1 M Na citrate	5.6	30% MPD
9	None	0.1 M Na citrate	5.6	20% 2-propanol, 20% PEG 4K
10	None	0.1 M Na citrate	5.6	1.0 M (NH ₄) ₂ phosphate
11	0.2 M CaCl ₂	0.1 M Na acetate	4.6	20% 2-propanol
12	None	0.1 M Na cacodylate	6.5	1.4 M Na acetate
13	0.2 M Na citrate	0.1 M Na cacodylate	6.5	30% 2-propanol
14	0.2 M (NH ₄) ₂ sulfate	0.1 M Na cacodylate	6.5	30% PEG 8K
15	0.2 M Mg acetate	0.1 M Na cacodylate	6.5	20% PEG 8K
16	0.2 M Mg acetate	0.1 M Na cacodylate	6.5	30% MPD
17	None	0.1 M imidazole	6.5	1.0 M Na acetate
18	0.2 M Na acetate	0.1 M Na cacodylate	6.5	30% PEG 8K
19	0.2 M Zn acetate	0.1 M Na cacodylate	6.5	18% PEG 8K
20	0.2 M Ca acetate	0.1 M Na cacodylate	6.5	18% PEG 8K
21	0.2 M Na citrate	0.1 M Na HEPES	7.5	30% MPD
22	0.2 M MgCl ₂	0.1 M Na HEPES	7.5	30% 2-propanol
23	0.2 M calcium chloride	0.1 M Na HEPES	7.5	28% PEG 400
24	0.2 M CaCl ₂	0.1 M Na HEPES	7.5	30% PEG 400
25	0.2 M Na citrate	0.1 M Na HEPES	7.5	20% 2-propanol
26	None	0.1 M Na HEPES	7.5	0.8 M K/Na tartrate
27	None	0.1 M Na HEPES	7.5	1.5 M Li sulfate
28	None	0.1 M Na HEPES	7.5	0.8 M Na/K phosphate
29	None	0.1 M Na HEPES	7.5	1.4 M tri-sodium citrate
30	None	0.1 M Na HEPES	7.5	2% PEG 400, 2.0 M (NH ₄) ₂ sulfate
31	None	0.1 M Na HEPES	7.5	10% 2-propanol, 20% PEG 4K
32	None	0.1 M Tris	8.5	2.0 M (NH ₄) ₂ sulfate
33	0.2 M MgCl ₂	0.1 M Tris	8.5	30% PEG 4K
34	0.2 M Na citrate	0.1 M Tris	8.5	30% PEG 400
35	0.2 M Li sulfate	0.1 M Tris	8.5	30% PEG 4K
36	0.2 M (NH ₄) ₂ acetate	0.1 M Tris	8.5	30% 2-Propanol
37	0.2 M Na acetate	0.1 M Tris	8.5	30% PEG 4K
38	None	0.1 M Tris	8.5	8% PEG 8K
39	None	0.1 M Tris	8.5	2.0 M (NH ₄) ₂ phosphate
40	None	None	-	0.4 M K/Na tartrate
41	None	None	-	0.4 M (NH ₄) ₂ phosphate
42	0.2 M (NH ₄) ₂ sulfate	None	-	30% PEG 8K
43	0.2 M (NH ₄) ₂ sulfate	None	-	30% PEG 4K
44	None	None	-	2.0 M (NH ₄) ₂ sulfate
45	None	None	-	4.0 M Naformate
46	0.05 M K phosphate	None	-	20% PEG 8K
47	None	None	-	30% PEG 1.5K
48	None	None	-	0.2 M Mg formate
49	1.0 M Li sulfate	None	-	2% PEG 8K
50	0.5 M Li sulfate	None	-	15% PEG 8K

Appendix IV – 2D-PAGE

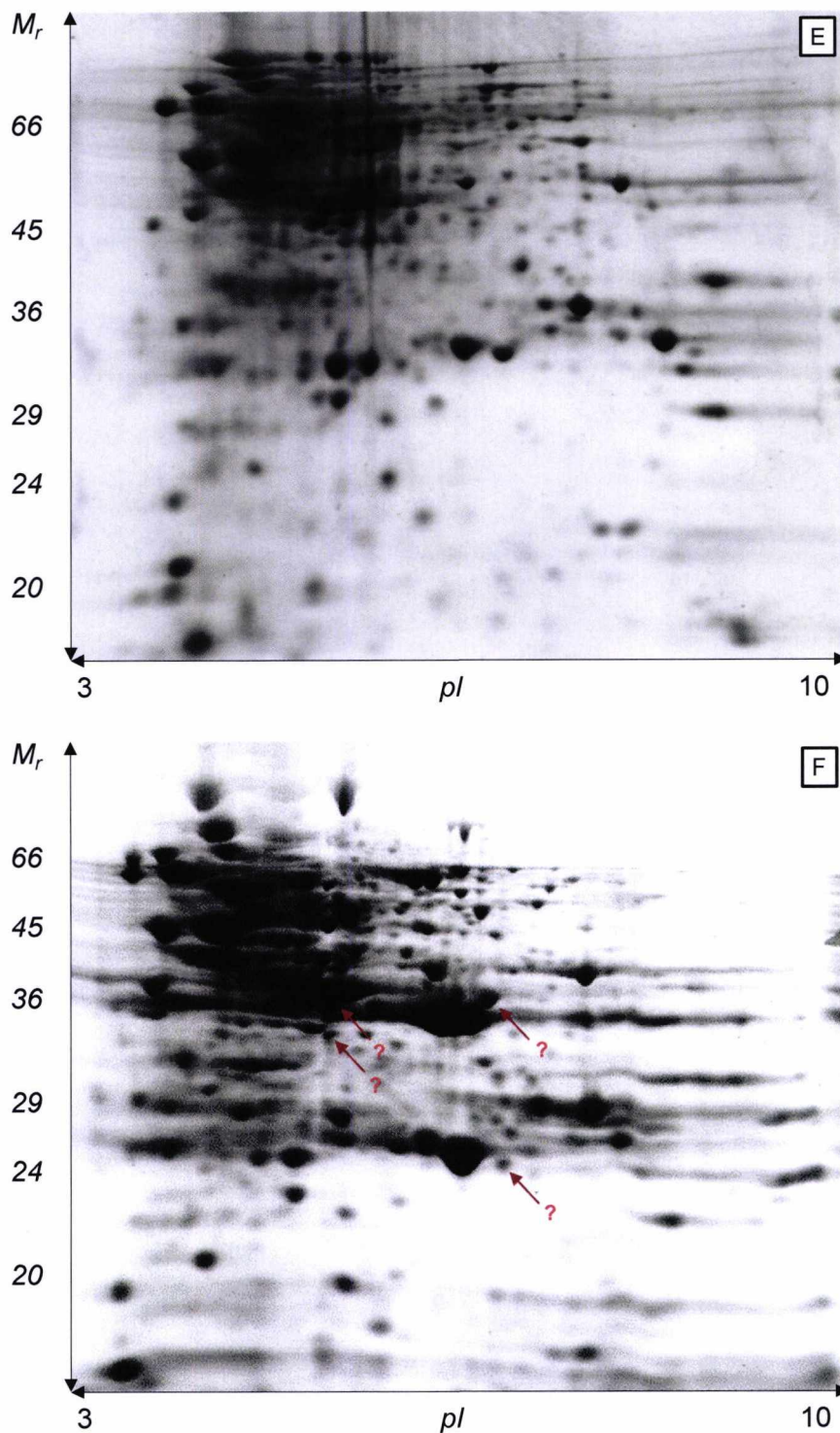


Figure 3.3 (continued). 2D-PAGE of recombinant strains of *B. megaterium* DSM319. 2-D IEF/SDS-PAGE gel electrophoretic separation of 100 μ g of cell lysate from *B. megaterium* DSM319 carrying the pSJM129 (E) or two plasmids pSJM169+pSJM236 (F). Cells were grown in LB broth at 37°C, induced with 0.5% D-xylose (w/v) and grown for 24 hrs.

Appendix V

Protocol for the synthesis of the anaerobic vitamin B₁₂ pathway intermediates

For the synthesis of the anaerobic pathway intermediates a detailed protocol is provided. All enzyme purifications and incubations were conducted under anaerobic conditions (2-10 ppm). To begin with, a detailed protocol for the synthesis of cobalt(II)-factor III is shown.

Preparation of cobalt(II)-factor III

Production of the enzymes HemB, HemC, HemD, CobA, SirC, CbiX^S and CbiL in *E. coli*

The enzymes HemB (*M. thermoautotrophicus* ALA dehydratase), HemC (*B. megaterium* PBG deaminase), HemD (*B. megaterium* uroporphyrinogen III synthase), CobA (*P. denitrificans* SUMT), SirC (*B. megaterium* precorrin-2 dehydrogenase), CbiX^S (*M. thermoautotrophicus* cobaltochelataase) and CbiL (*M. thermoautotrophicus* cobalt-factor II C-20 methyltransferase) can be overproduced in *E. coli* BL21Star(DE3), which contains the plasmids pLysS and pETcoco-2-cobA-hemB-hemC-hemD-sirC-cbiX^S-cbiL (kept as a glycerol stock at -80 °C). All genes contain a His₆-tag fusion at the N-terminus, allowing the expressed gene products to be overproduced in *E. coli* and then purified by metal affinity chromatography.

A pre-culture was prepared by inoculating the strain directly from the glycerol stock into 50 ml of LB containing kanamycin (30 µg ml⁻¹) and incubated overnight at 37 °C, 150 rpm for 16 hrs. 10 ml of pre-culture was then added to 1 litre of LB medium containing kanamycin (30 µg ml⁻¹). The culture was grown for 6 hrs, 37 °C, 150 rpm. Protein production was induced with 0.4 mM IPTG and the culture is left overnight at 18 °C, 150 rpm.

Purification of the enzymes HemB, HemC, HemD, CobA, SirC, CbiX^S and CbiL

For enzyme purification, cultures were centrifuged at 4000 rpm for 10 min at 4 °C. The supernatant was discarded and the pellet was re-suspended in 15 ml (per litre of

culture) of binding buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 5 mM imidazole). Cells were sonicated (5 min at 65% amplitude: 30 second ON, 30 second OFF), followed by centrifugation (17500 rpm for 20 min at 4 °C). After this, the His₆-tagged proteins can be purified aerobically on a His₆-bind column. Washes included 50 ml of binding buffer and 25 ml of wash buffer I (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 30 mM imidazole). Proteins were eluted in 20 mM Tris-HCl pH 8.0, 500 mM NaCl, 500 mM imidazole.

Synthesis of cobalt(II)-factor III

The purified enzymes and chemicals (20 mg ALA, 60 mg NAD⁺, 100 mg SAM, 2.4 mg CoCl₂.6H₂O, 40 mg NaOH) were purged with nitrogen and taken into the glovebox. A PD-10 column was pre-equilibrated with 25 ml of Buffer A (20 mM Tris-HCl pH 8.0, 100 mM NaCl).

1. 2.5 ml of purified proteins were added to the PD-10 column and buffer exchanged into 3 ml of Buffer A.
2. 20 mg ALA, 60 mg NAD⁺, 100 mg SAM were dissolved in 7 ml of Buffer A (substrate solution) in a clean glass vial
3. 40 mg of NaOH was dissolved in 1 ml of ddH₂O (1 M stock)
4. 100 µl of 1 M NaOH was added to the substrate solution
5. 900 µl of 1 M Tris-HCl pH 8.0 was added to the substrate solution
6. 3 ml of purified enzymes were added to the substrate solution
7. Incubate substrate solution for 30 min at 37 °C
8. A small stir bar (“flea”) was added to the substrate solution
9. The substrate solution was positioned in a clamp over the magnetic stirrer. The speed was increased until a vortex had formed
10. 2.4 mg of CoCl₂.6H₂O was dissolved in 1 ml of Buffer A (10 mM stock solution)
11. 100 µl of 10 mM CoCl₂.6H₂O was added to the substrate solution
12. This was incubated for 5 min with constant stirring at room temperature
13. Steps 11-12 were repeated ten times
14. Incubate substrate solution for 30 min, 37 °C
15. Steps 11-12 were repeated five times
16. Leave incubation overnight at 37 °C

17. 20 μl of concentrated cobalt(II)-factor III was added to 980 μl of ddH₂O and the absorbance at 392 nm ($A_{392\text{nm}} \epsilon \approx 150000 \text{ M}^{-1} \text{ cm}^{-1}$) was measured to calculate an approximate concentration of cobalt(II)-factor III

When calculating the concentration of cobalt(II)-factor III, if the main absorbance peak had shifted to 414-425 nm, 10 mM DTT was added and left for 30 min. This converted any cobalt(III)-factor III into cobalt(II)-factor III. This protocol should produce approximately 0.5-1 mM of concentrated cobalt(II)-factor III.

Synthesis of intermediates from cobalt-factor IV to cobyrinic acid

Enzyme production and purification in *B. megaterium* DSM319

Glycerol stocks (stored at -80 °C) of *B. megaterium* DSM319 recombinant strains were streaked onto LB agar plates containing 10 $\mu\text{g ml}^{-1}$ tetracycline and grown overnight (< 16 hrs) at 30 °C.

Plasmid strains

B. megaterium DSM319 pC-His1622-*cbiH*₆₀ (pSJM188)
B. megaterium DSM319 pC-His1622-*cbiF* (pSJM116)
B. megaterium DSM319 pC-His1622-*cbiG* (pSJM73)
B. megaterium DSM319 pC-His1622-*cbiD* (pSJM195)
B. megaterium DSM319 pC-His1622-*cbiJ* (pSJM78)
B. megaterium DSM319 pC-His1622-*cbiET* (pSJM234)
B. megaterium DSM319 pN-His-TEV1622-*cbiC* (pSJM231)

After overnight incubation, cell paste was scraped from the overnight growth and inoculated directly into 50 ml of LB containing 10 $\mu\text{g ml}^{-1}$ of tetracycline. After re-suspension of the cells, the starting OD₅₇₈ was approximately 0.1-0.3. Cultures were grown at 30 °C, 200 rpm for 4-6 hrs (OD₅₇₈ ~ 2-5). 10 ml of the culture was then transferred to 1 L of LB containing 10 $\mu\text{g ml}^{-1}$ of tetracycline and grown for 3-4 hrs at 30 °C, 150 rpm. Protein production was induced with 0.25% (w/v) D-xylose. Cultures were grown overnight at 30 °C, 150 rpm for less than 16 hrs.

For enzyme purification, cultures were centrifuged at 5000 rpm for 10 min at 4 °C. The supernatant was discarded and the pellet was re-suspended in 15 ml (per litre of culture) of binding buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 5 mM imidazole). Cells were sonicated (5 min at 65% amplitude: 30 second ON, 30

second OFF), followed by centrifugation (17500 rpm for 20 min at 4 °C). Following this, the supernatant (in a glass beaker) was taken into the glovebox after purging with nitrogen. The His₆-tagged enzymes were purified separately using a His₆-bind column. Washes included 25 ml of binding buffer, 25 ml of wash buffer I (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 30 mM imidazole), followed by elution in 20 mM Tris-HCl pH 8.0, 500 mM NaCl, 500 mM imidazole. Enzyme concentration was estimated by absorbance at 280 nm. Purified enzymes were kept at 4 °C.

Synthesis of cobalt-factor IV

CbiH₆₀ (in elution buffer) was incubated for 10 min with cobalt(II)-factor III in a total volume of 2.5 ml. Typical concentrations ranged from 20-60 µM of both the enzyme, CbiH₆₀, and the substrate, cobalt(II)-factor III. This mixture was then applied to a PD-10 column and then buffer-exchanged into 3 ml of Buffer H (50 mM Tris-HCl pH 8.0, 400 mM NaCl). The total volume was then adjusted to 4.9 ml with Buffer H. 50 µL of 0.1 M SAM and 50 µL of 1 M DTT was added and the reaction was incubated at 37 °C for 24 hrs in a glass vial (wrapped in aluminium foil). The expected yield with a 1:1 ratio of enzyme and substrate would be approximately 75%. This could be improved by adding an excess of CbiH₆₀. To purify the cobalt-factor IV, the enzyme and cobalt-factor III could be removed by heat (80 °C, 30 min) or acid denaturation [1% (v/v) acetic acid]. Cobalt-factor IV could be further purified by anion-exchange (sample diluted in ddH₂O to lower NaCl to < 100 mM) and reverse-phase (RP18) chromatography methods.

Synthesis of cobalt-precorrin-5

CbiH₆₀, CbiF and CbiG (in elution buffer) was incubated for 10 min with cobalt(II)-factor III in total volume of 2.5 ml. Typical concentrations ranged from 20-60 µM of both the enzymes, CbiH₆₀, and the substrate, cobalt(II)-factor III. An example of this mixture is shown below.

Enzyme/Substrate	Stock concentration (μM)	Volume added to 2.5 ml (μL)	Approximate concentration after buffer exchange (μM)
CFIII	340	440	43
CbiH ₆₀	200	750	43
CbiF	245	610	43
CbiG	395	380	43
Buffer H added		320	

This mixture was then applied to a PD-10 column and then buffer-exchanged into 3 ml of Buffer H (50 mM Tris-HCl pH 8.0, 400 mM NaCl). The total volume was then adjusted to 4.9 ml with Buffer H. 50 μL of 0.1 M SAM and 50 μL of 1 M DTT was added and the reaction was incubated at 37 °C for 24 hrs in a glass vial (wrapped in aluminium foil). Typically 3 litres of CbiH₆₀, 1 litre of CbiF and 1 litre of CbiG cultures could provide enough enzyme preparation for a total volume of 15 ml of an enzyme-cocktail mixture that could produce cobalt-precorrin-5. The expected yield would be approximately 75%. This could be improved by adding an excess of CbiH₆₀. To purify the cobalt-precorrin-5, the enzymes and cobalt-factor III could be removed by heat (65 °C, 30 min) or acid denaturation [1% (v/v) acetic acid]. Cobalt-precorrin-5 could be further purified by anion-exchange (sample diluted in ddH₂O to lower NaCl to < 100 mM) or by reverse-phase (RP18) chromatography methods. However, cobalt-precorrin-5 would be insoluble after freeze-drying, if it was eluted in ethanol containing acetic acid. To prevent this, prior to elution the intermediate was equilibrated with ddH₂O and then eluted in 40% (v/v) ethanol.

Synthesis of cobalt-precorrin-6A

Cobalt-precorrin-6A could be synthesised by a similar method described for the synthesis of cobalt-precorrin-5. However it was more efficient to first synthesise cobalt-precorrin-5 and then add CbiD. CbiD was purified from 2-4 litres of culture and buffer exchanged into Buffer H. From a typical preparation, after buffer exchange, approximately 6 ml of 40 μM CbiD could be obtained. Using the mixture containing cobalt-precorrin-5 that is previously described, 10 ml of cobalt-precorrin-5, 6 ml of 40 μM CbiD, 200 μL of 0.1 M SAM, 200 μL of DTT were incubated at 37 °C for 24 hrs. The estimated yield was > 95%. This reaction could

also be modified by first purifying cobalt-precorrin-5 and incubating with CbiD, with the substrate concentration in excess.

Synthesis of cobalt-precorrin-6B

Cobalt-precorrin-6B could be synthesised from cobalt-precorrin-6A, by adding purified CbiJ (10 μM) with an excess of NADH (250 μM) to a mixture containing cobalt-precorrin-6A. The estimated yield was $> 95\%$. This reaction could also be modified by first purifying (by anion-exchange) cobalt-precorrin-6A and incubating with CbiJ, with the substrate concentration in excess.

Synthesis of cobalt-precorrin-7

To obtain cobalt-precorrin-7, the CbiE from *M. thermoautotrophicus* was produced in *E. coli* and purified. Using a similar method to the synthesis of cobalt-precorrin-6B, CbiE was incubated as an enzyme-cocktail as described in Section 6.2.6.

Synthesis of cobalt-precorrin-8 and cobyric acid

These intermediates were only briefly characterised and the enzyme-cocktail based method is described in Section 6.2.7.