A study on the effect of B_{12} variants on the *E. coli btuB* riboswitch and the development of a system for the production of cobalt-free B_{12}

A thesis submitted to the University of Kent for the degree of PhD in Microbiology

2019

Naziyat Islam Khan

Abstract

Vitamin B_{12} is a water-soluble nutrient that is the largest and most structurally complex of all the known vitamins, requiring over thirty enzymatic steps for its biosynthesis. The work carried out here investigates the possibility of engineering the production of B_{12} variants, through the construction of an empty, metal-free B_{12} shell, and exploring what effect such compounds might have on cellular control elements, such as riboswitches. In order to study this, a method was designed and developed that utilises an '*in vitro*' system to determine the effect of B_{12} variants on a well-characterised B_{12}-dependent riboswitch known as the *btuB* riboswitch. This method allows a quick and easy way of testing the ability of B_{12} variants to directly bind to and deactivate the *btuB* riboswitch, without relying on indirect cellular effects.

The production of such a technique required the construction of a plasmid in which the *btuB* riboswitch was placed upstream of the eGFP reporter gene. This plasmid was combined with *in vitro* transcription/translation machinery, in the presence of different B_{12} variants, of both natural and unnatural origin, to see how each affects the levels of eGFP production. eGFP production was monitored in the presence of AdoCbl, AdoRbl, CNCbl, AdoCbi, CNCbi, Nibl, Znbl, EtPhCbl and F_2PhEtyCbl. The results showed that AdoCbl and AdoRbl had the greatest effect on the riboswitch, with AdoRbl closely mimicking the levels of inhibition caused by AdoCbl. CNCbl, AdoCbi and CNCbi were seen to have much lesser effects and required higher concentrations to cause similar effects to that of AdoCbl or AdoRbl. The remaining four B_{12} variants were seen to have little effect, if any. Overall, results showed that an adenosyl group in place of the upper ligand is crucial to riboswitch recognition. The demonstration that AdoRbl is able to bind to the riboswitch highlights the potential for such analogues, with the required components for recognition, to

interfere with and inhibit cellular processes. Thus, the research carried out here opens up an exciting new field of research into the use of B_{12} analogues as antibacterial agents.

Table of Contents

Figures index

Tables index

Abbreviations

AdoCbi - Adenosylcobinamide

AdoCbl – Adenosylcobalamin

A_M – Maximal activation (switching off) of the riboswitch by a particular ligand

AdoRbl – Adenosylrhodibalamin

ATP – Adenosine triphosphate

B_{12} or Cbl – Cobalamin

Cbi – Cobinamide

CNCbi - Cyanocobinamide

CNCbl – Cyanocobalamin

CoA – Coenzyme A

dH_2O – Distilled water

DMB - 5,6-dimethylbenzimidazole

DNA – Deoxyribonucleic acid

eGFP – Enhanced green fluorescent protein

EtPhCbl - Coß-(4-ethylphenyl)-cobalamin

F_2PhEtyCbl - Coß-(2,4-difluorophenyl)ethinyl-cobalamin

FMNH - Flavin mononucleotide hydride

FPLC – Fast protein liquid chromatography

GFP - Green fluorescent protein

GTP – Guanosine triphosphate

H_2O - Water

HBA - Hydrogenobyrinic acid

HBAD - Hydrogenobyrinic acid a,c-diamide

HBAH - Hydrogenobyric acid

HBAM - Hydrogenobyrinic acid monoamide

HEPES – 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HPLC – High performance liquid chromatography

IF – Intrinsic factor

IMAC – Immobilised metal affinity chromatography

IPTG - Isopropyl β-D-1-thiogalactopyranoside

K_D – Dissociation constant

K_i – Inhibition constant

K_{off} – Switching off constant (concentration at which there is half of the maximal activation, A_M)

LB - Luria-Bertani

MeCbl – Methylcobalamin

MetE – B_{12}-independent methionine synthase

MetH - B_{12}-dependent methionine synthase

MES – 2-ethanesulfonic acid

Mg/Ace – Magnesium acetate

MPD – 2-methyl-2,4-pentanediol

mRNA – Messenger ribonucleic acid

MS – Mass spectrometry

Na/Caco – Sodium cacodylate

NaMN - β-nicotinate mononucleotide

ncRNA – Non-coding ribonucleic acid

Nibl – Nibalamin

NTP – Nucleotide triphosphate

OHCbi - Hydroxycobinamide

OHCbl – Hydroxycobalamin

PAGE – Polyacrylamide gel electrophoresis

PBS – Phosphate-buffered saline

PCR – Polymerase chain reaction

PEG – Polyethylene glycol

RBS - Ribosome-binding site

RNA – Ribonucleic acid

SAH - S-adenosylhomocysteine

SAM - S-adenosylmethionine

SDS – Sodium dodecyl sulphate

SOC – Super optimal broth

TBDT - TonB-dependent transporter

TCN2 - Transcobalamin II

TFA – Trifluoroacetic acid

Uro'gen III - Uroporphyrinogen III

UTR - 5’ untranslated region

YE – Yeast Extract

Znbl - Zincobalamin

Acknowledgements

Firstly, I would like to thank Professor Martin Warren for providing me with the opportunity to work on this project. I am extremely grateful for your belief in me, as well as for all the support and encouragement you have provided me with throughout my three years on this project.

Next, I would like to thank my supervisors, Dr. Evelyne Deery and Dr. Andrew Lawrence, for all their support and guidance. I would also like to express my gratitude and appreciation to each and every member of the Warren Lab, who have not only created an enjoyable work environment, but also taught me a lot. Outside of my lab I would like to thank all my good friends and colleagues in the school of biosciences, who have given me so much support.

I have also received an immense amount of love from my grandparents, Wasiur Rahman, Fatima Afsar and Joybun Nessa. Thank you for keeping me in your prayers.

Another person I would like to thank is my guide and mentor, Mohammed Abul Kasem, who has given me an invaluable amount of support and guidance from the beginning to the end.

Finally, I would like to dedicate this thesis to my wonderful parents, Nurul Islam Khan and Halima Akter Khan, without whom I would not be where I am today. Thank you for your continuous love and support, as well as your belief in me to be successful and excel in my every endeavour.

Chapter 1: Introduction

1.1 What is B_{12}

1.1.1 B_{12} Overview

Vitamin B_{12} is a water-soluble vitamin that is the largest and most structurally complex of all the known vitamins. It is unique among the vitamins due to the fact that it is synthesised only by prokaryotes (Roth et al 1996; Warren et al, 2002). It was first identified as the anti-pernicious anaemia factor found in crude liver extract through the work of George Whipple, George Minot and William Murphy (Minot and Murphy, 1926; Whipple et al 1920) . Following this, the chemical structure was elucidated in 1955 by Dorothy Hodgkin and her team (Hodgkin et al, 1955) which was the largest crystal structure to be determined by crystallographic analysis at that time. It is now known to have a role in normal brain functioning (Weir and Scott, 1999) and the formation of red blood cells (Koury and Ponka, 2004). Technically, vitamin B_{12} is a specific form of cobalamin called cyanocobalamin. However, the term B_{12} will be used here to refer to cobalamin in general, largely to cover the more common versions of the molecule, adenosylcobalamin (AdoCbl), methylcobalamin (MeCbl) and cyanocobalamin (CNCbl).

1.1.2 Structure of B_{12}

B_{12} contains a modified tetrapyrrole that houses a cobalt ion at its centre (Figure 1.1), which is referred to as a corrin ring. The corrin ring is the core component of the vitamin with various attached side groups, including upper and lower axial ligands, the latter of which is connected to the corrin ring via the nucleotide loop (Figure 1.1) (Roth et al, 1996). The cobalt ion is able to form six bonds, four of which it forms with nitrogen atoms (from the four pyrrole subunits, named A-D, which are coordinated in plane to the central cobalt ion) and the remaining two with the upper and lower ligands found in the structure of B_{12} (Figure 1.1). The upper

ligand is generally variable, while the lower ligand of cobalamin is 5,6-dimethylbenzimidazole (DMB), which is attached to the periphery of the corrin ring (ring D) via an aminopropanol linker. The presence of DMB at the lower ligand position defines the molecule as cobalamin. When any other base is present, the molecule is a non-cobalamin. The general structure of B_{12} can be seen in Figure 1.1.

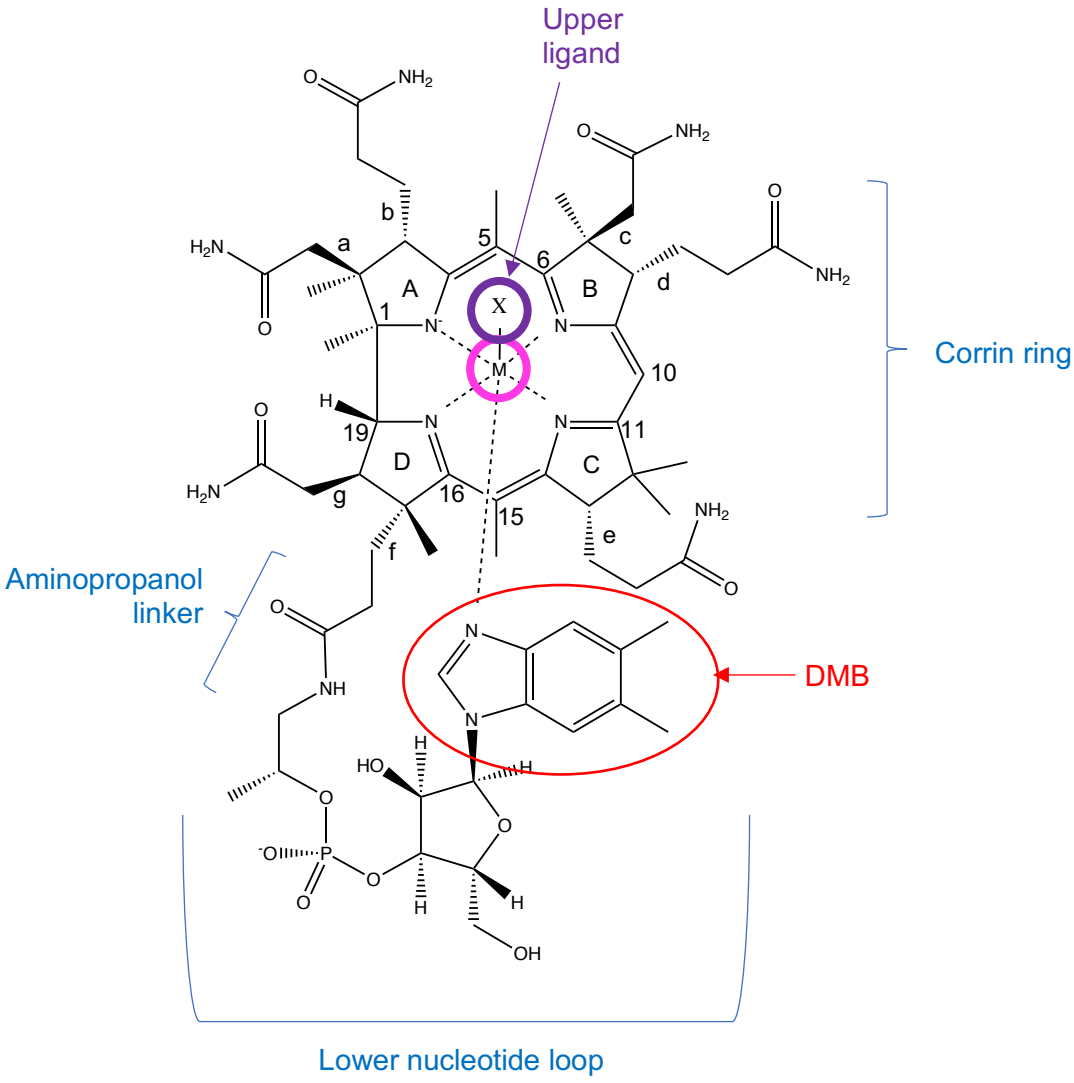


Figure 1.1. General structure of B_{12} (cobalamin). (DMB at the lower ligand position shown inside red circle). X = upper ligand, M = central metal ion. The four pyrrole groups are labelled A-D.

In addition to the variation created by the nature of the upper ligand, greater structural complexity arises from the ability of B_{12} to have two different conformations known as “base on” or “base off”. Determination of which of the two conformations the molecule is in depends on the presence of a bond between the cobalt ion and DMB, represented by the dotted line in Figure 1.1, which when present, represents the “base on” conformation (Mathews et al, 2007). In certain forms of B_{12}, this bond is broken and the molecule is said to be in the “base off” conformation. This broken bond is often a result of B_{12} binding to certain proteins, where conformational changes cause bond breakage (Wuerges et al, 2006).

1.1.3 Upper and lower ligands of B_{12}

Due to the variable nature of the upper ligand, there are a number of different forms of B_{12}. Most common forms include AdoCbl, MeCbl, and CNCbl with upper ligands of 5’deoxyadenosyl-, methyl-, or cyano-, respectively. Of these, AdoCbl and MeCbl are the biologically active forms in humans (Marsh, 1999), whose roles are discussed in Section 1.3.

Use of cobamides (cobalt-containing corrins with upper and lower ligands, of which cobalamin is an example) is far more widespread across the bacterial kingdom than its biosynthesis (Shelton et al, 2018). This suggests that the majority of cobamide-requiring bacteria take up incomplete corrinoids from their environment and synthesise lower ligands for completion of these molecules. Different bacteria use different lower ligands depending on their enzyme specificity as well as on the molecules or lower ligand forming components present in their environment (Crofts et al, 2013; Hazra et al, 2013). For example, *Lactobacillus reuteri* is a bacterium which assigns adenine at the lower ligand position, instead of DMB, in order to form pseudocobalamin (Santos et al, 2007). Research has shown that most bacteria

have enzymes that preferentially attach DMB as the lower ligand of B_{12}, but are capable of using a wide range of lower ligand substrates (Hazra et al, 2013). The lower ligand bases of cobamides can be divided into one of three chemical classes: benzimidazoles, purines and phenolics (Figure 1.2) (Hazra et al, 2013; Shelton et al, 2018).

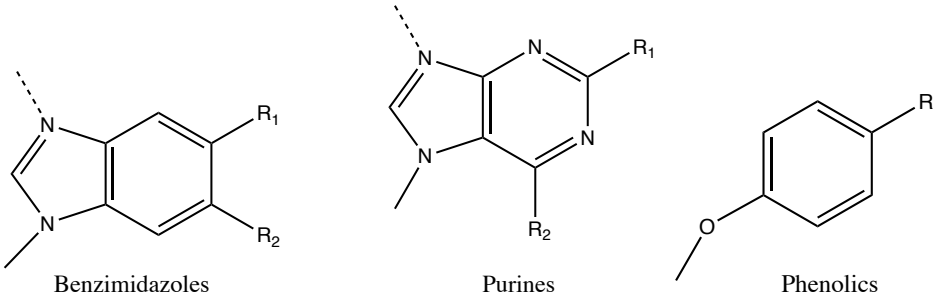


Figure 1.2. Lower ligand structure classes. Benzimidazoles: R_1, R_2 = H, OH, CH_3, OCH_3. Purines: R_1= H, CH_3, NH_2; R_2 = H, NH_2, OH, O. Phenolics: R = H, CH_3 Imaged adapted from (Shelton et al, 2018)

Different B_{12}-dependent enzymes require different forms of cobamides. This can be dependent upon either the presence of a certain type of upper ligand, a certain type of lower ligand or even dependent upon whether the molecule is in “base-on” or “base-off” conformation. For example, for enzymes such as class II ribonucleotide reductases, the cobamide has been shown to bind the enzyme in the “base-on” form (Abend et al., 1999; Lawrence et al., 1999). Meanwhile, enzymes such as methionine synthase and methylmalonyl CoA mutase require cobalamin to bind in the “base-off” conformation (Drennan et al 1994; Mancia et al, 1996). However, with both “base-on” and “base-off” requiring enzymes, it is thought that the structure of the lower ligand may also affect the binding of the cobamide (Hamza et al, 2005; Kräutler et al, 2003), for which reason bacteria have mechanisms to limit cobamide

production only to the forms required by the enzymes they contain (Chan and Semerena, 2011; Hazra et al, 2013).

1.2 B_{12} Biosynthesis

1.2.1 Biosynthesis Overview

Vitamin B_{12} is unique among the vitamins in that it is made solely by only certain bacteria and archaea. *De novo* synthesis of B_{12} is a highly complex process that requires ~30 enzymes (Warren et al, 2002). There are two genetically distinct pathways for B_{12} biosynthesis, one being the aerobic pathway, which is utilised by bacteria such as *Pseudomonas denitrificans* (Blanche et al 1995) and the other being the anaerobic pathway, which is employed by bacteria such as *Bacillus megaterium* (Moore and Warren, 2012). The aerobic B_{12} biosynthesis pathway is different to the anaerobic pathway in that it requires not only the presence of molecular oxygen, but also late cobalt insertion in the pathway (Heldt et al, 2005; Warren et al, 2002).

In both cases, B_{12} is derived from uroporphyrinogen III (uro'gen III) (Warren and Scott, 1990). Of the two biosynthesis pathways, the aerobic pathway has more stable intermediates and is therefore, more commonly studied. Similarly, the work involved in this report will also largely be based around the aerobic B_{12} biosynthesis pathway (Figure 1.3 and 1.4).

1.2.2 Aerobic Biosynthesis Pathway

The first enzyme involved in the aerobic biosynthesis of AdoCbl is CobA, which methylates the starting material, uro'gen III, at C2 and C7 in order to generate precorrin-2. This is followed by CobI which methylates precorrin-2 at C20, such that it is converted into precorrin-3A. The next enzyme in the pathway is CobG which catalyses hydroxylation of precorrin-3A at C20, resulting in the formation of precorrin-3B. Precorrin-3B is methylated by CobJ, at C17, following which the

crucial ring contraction step occurs and leads to the formation of precorrin-4 (Warren et al, 2002).

CobM is the subsequent enzyme in line, which methylates precorrin-4 at C11 to form precorrin-5, followed by CobF which methylates precorrin-5 at C1, to generate precorrin-6A. Next, CobK converts precorrin-6A to precorrin-6B, via C18-19 reduction, which is then acted upon by CobL, an enzyme catalysing C5 and C15 methylation as well as decarboxylation of the acetate group at C12, to generate precorrin-8. C11-C12 of precorrin-8 is subject to methyl rearrangement, by the action of CobH, in order to form hydrogenobyrinic acid, HBA. HBA is amidated by CobB, resulting in the formation of HBA-diamide (HBAD) (Warren et al, 2002).

Cobalt insertion is a major and crucial step in B_{12} biosynthesis. B_{12} rarely exists without the cobalt ion, except in certain photosynthetic bacteria, such as, *Allochromatium vinosum*, from which corrinoid compounds have been found containing no cobalt (Toohey, 1965), most likely a consequence of growing the organism in the absence of cobalt.

In the majority of bacteria, however, three proteins are involved in cobalt insertion; CobN, CobS and CobT, which convert HBAD into cob(II)yrinic acid *a,c*-diamide, before reduction of cobalt by CobR to form cob(I)yrinic acid *a,c*-diamide (Warren et al, 2002). The next step is the adenosylation of cob(I)yrinic acid *a,c*-diamide, by CobO, to give adenosylcobyrinic acid *a,c*-diamide. This is then subject to *b, d, e, g*-amidation by CobQ to give adenosylcobyric acid (Warren et al, 2002).

The final stage of B_{12} biosynthesis involves the synthesis and assembly of the lower nucleotide loop. As mentioned previously, the lower loop is attached to the corrin ring via an aminopropanol linker. The synthesis of this aminopropanol linker begins

with the actions of the CobC enzyme (referred to as CobD in the anaerobic pathway, for example in bacteria such as *Salmonella enterica*). CobC, a small dimeric enzyme known as L-threonine phosphate decarboxylase, catalyses the synthesis of aminopropanol phosphate from L-threonine phosphate, which is the crucial component that links the lower nucleotide loop to the corrin ring, at C17 of ring D (Cheong et al, 2002). L-threonine phosphate, itself, is synthesised from threonine by the action of a kinase called PduX (Fan and Bobik, 2008). In the aerobic pathway, CobC is thought to be part of a complex called 'protein β' with the cobinamide synthase, CobD (known as CbiB in the anaerobic pathway). Protein β (CobC and CobD complex) is thought to work together with another protein, protein α, to catalyse attachment of aminopropanol to adenosylcobyric acid in the presence of ATP and Mg^{2+} (Blanche et al, 1995), however, the exact function, identity and role of protein α is unknown. Together, CobC and CobD facilitate aminopropanol synthesis and its attachment at C17 of adenosylcobyric acid, in order to generate adenosylcobinamide (AdoCbi) (Warren et al, 2002).

Next, the bifunctional enzyme, CobP (known as CobU in the anaerobic biosynthesis pathway), initiates the process of assembling the lower loop by converting AdoCbi into AdoCbi-GDP (Warren et al, 2002). CobP is a small homodimeric protein with two different enzymatic functions, one as cobinamide kinase and the other as cobinamide phosphate guanylyltransferase. In the first step, CobP catalyses phosphorylation of the hydroxyl group of the aminopropanol in AdoCbi, using an ATP donor molecule, to form AdoCbi-phosphate (AdoCbi-P). In the second step, CobP transfers a GMP moiety, from a GTP donor, onto the phosphate group to form AdoCbi-GDP (Blanche et al, 1991).

Finally, lower ligand synthesis and assembly require two enzymes, nicotinate-nucleotide-dimethylbenzimidazole phosphoribosyltransferase, which is more commonly known as CobU (CobT in the anaerobic pathway) and cobalamin synthase, more commonly known as CobV (CobS in the anaerobic pathway) to form AdoCbl from AdoCbi-GDP (Warren et al, 2002). To begin with, CobU synthesises α-linked riboside monophosphates from a diverse range of nucleotide base substrates (generally one of benzimidazoles, purines or phenolics, as mentioned in Section 1.1.3) (Chan et al, 2014), for which reason there is such diversity in the possible lower ligands found within cobamides (Shelton et al, 2018). Generally, however, the preferred ligand which is attached and activated by the enzyme is DMB (Figure 1.1) (Hazra et al, 2013; Shelton et al, 2018). This activation involves the formation of a glycosidic bond between the ribose of β-nicotinate mononucleotide (NaMN) and DMB, in order to form α-ribazole phosphate (Blanche et al, 1991; Mander and Liu, 2010; Trzebiatowski et al, 1994). DMB, unless obtained from the environment, is synthesised from $FMNH_2$ by an enzyme called BluB (an oxidoreductase found in the aerobic biosynthesis pathway) (Mander and Liu, 2010). More recently, it has been found that DMB can also be synthesised anaerobically (Hazra et al, 2015). Following this synthesis of the lower loop (α-ribazole phosphate) is its assembly to AdoCbi-GDP by CobV, to form AdoCbl (Warren et al, 2002).

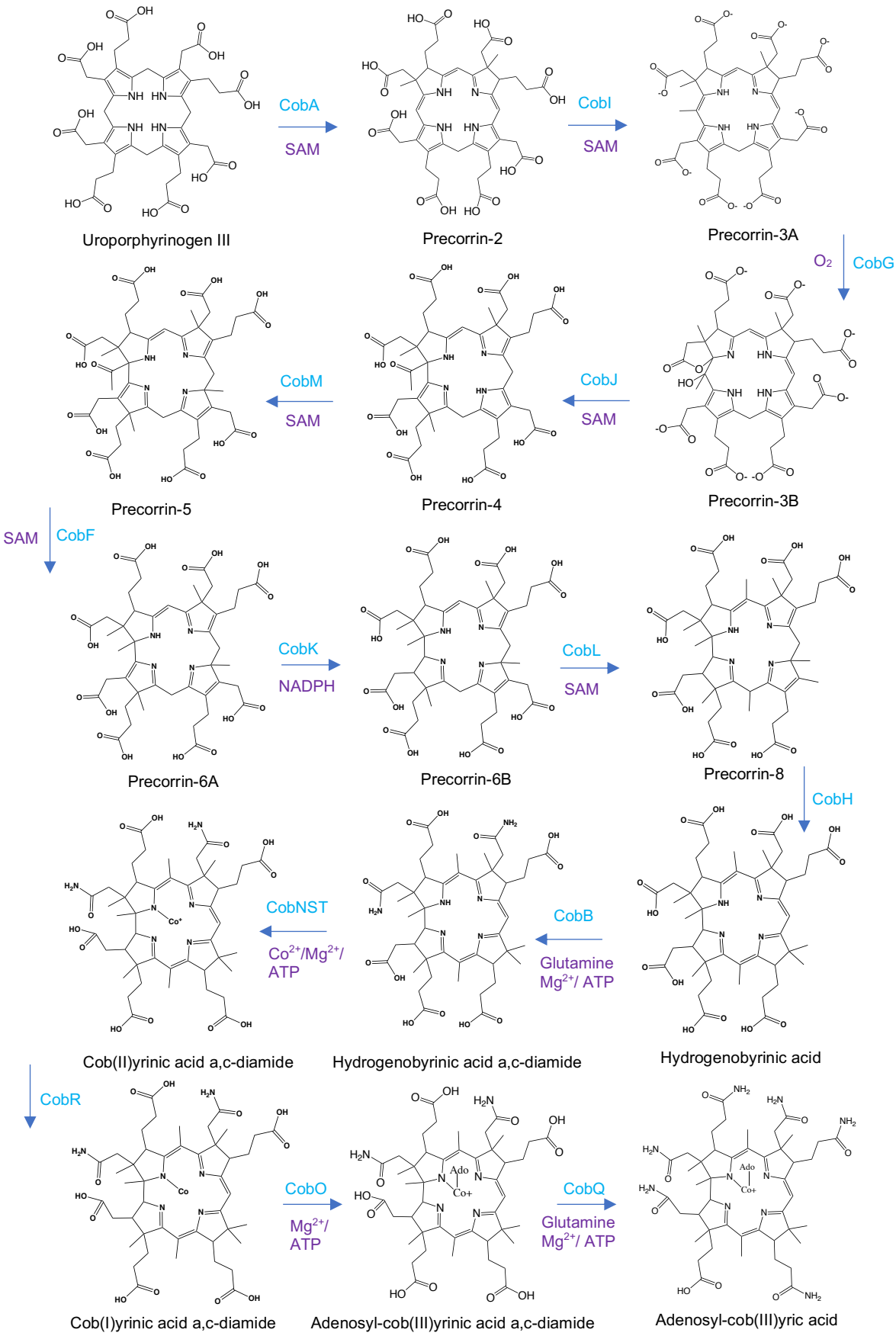


Figure 1.3. Aerobic biosynthesis of adenosyl-cob(III)yric acid from uro-gen

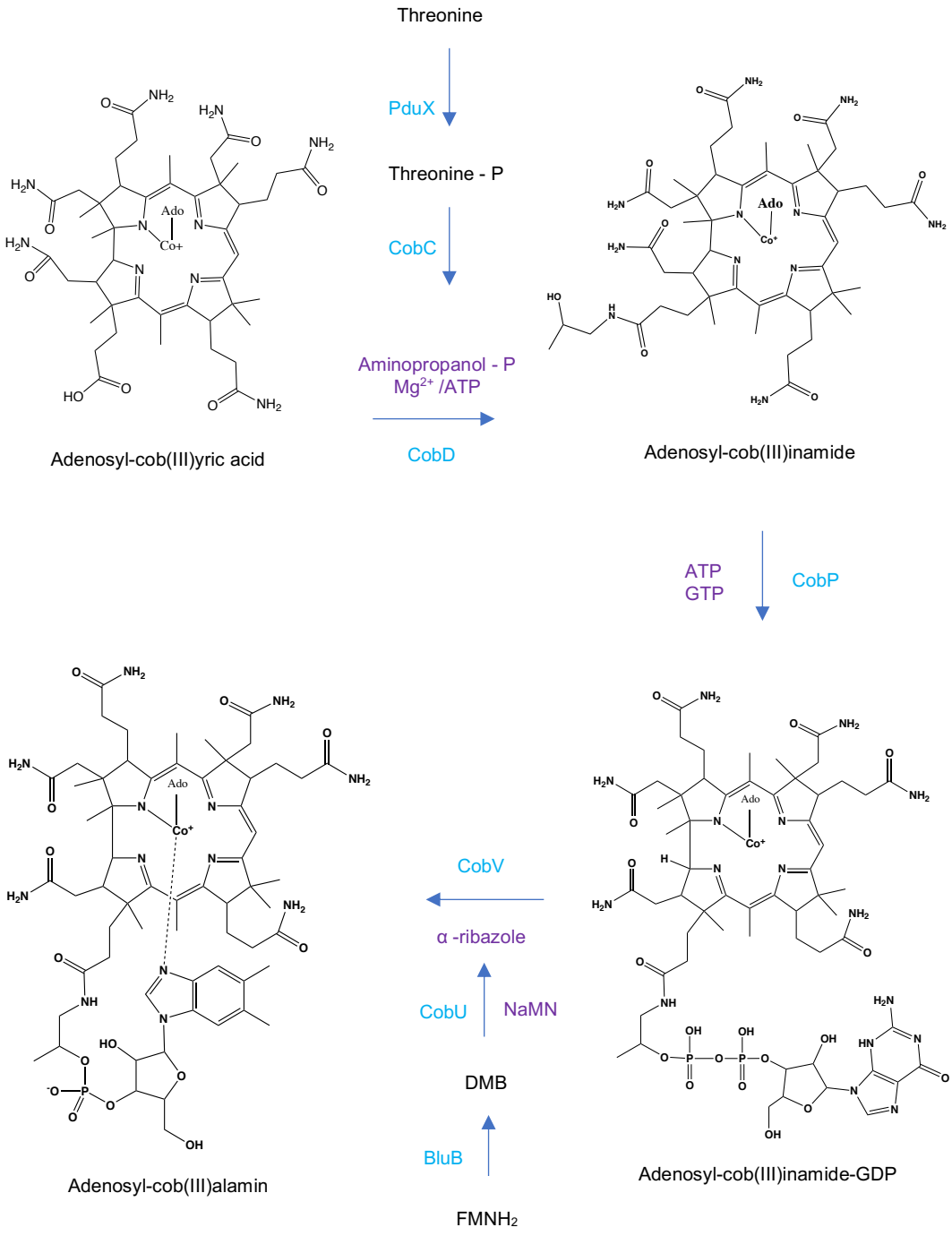


Figure 1.4. The conversion of adenosyl-cob(III)yric acid to adenosylcob(III)alamin

1.2.3 *Allochromatium vinosum*

Currently in this field very little is known about metal-free corrinoids in relation to both their structure and behaviour, in comparison to natural corrinoids. In 1965, metal-free corrinoids were first discovered and isolated from the photosynthetic bacterium, *A. vinosum* (Toohey, 1965). Four of these metal-free corrinoids were investigated for potential activity as B_{12} inhibitors, whereby growth inhibition was seen in *E. coli* (Perlman and Toohey, 1966). Following on from this, a whole new class of metal-substituted B_{12} analogues were synthesised using metal-free corrinoids isolated from *A. vinosum* (Koppenhagen and Piffner,1970).

One of the aims of this project involves the synthesis of hydrogenobalamin (a metal-free B_{12} shell) using enzymes from *A. vinosum*, to be used as the starting material for creating unnatural B_{12} analogues that act as B_{12} antagonists or inhibitors.

1.3 Role of B_{12}

1.3.1 Overview of the function of B_{12}

There are four main classes of B_{12}-dependent enzymes; the AdoCbl-dependent isomerases, the MeCbl-dependent methyltransferases, the radical SAM superfamily of B_{12}-dependent enzymes and the less well studied reductive dehalogenases (Banerjee and Ragsdale, 2003). It is the C-Co bond in B_{12} that gives the molecule its reactivity and is utilised by B_{12}-dependent enzymes. However, the bond is exploited differently depending on which enzyme is using the molecule, for example, the isomerases rupture the bond homolytically, while the methyltransferases cleave the bond heterolytically (Banerjee and Ragsdale, 2003).

1.3.2 Role in eukaryotes

B_{12} plays a key role in a number of essential biological processes. In humans, for example, there are two B_{12}-dependent enzymes, methionine synthase and methylmalonyl-CoA mutase (Banerjee et al, 2009).

MeCbl is a cofactor for methionine synthase, a methyltransferase which regenerates methionine from homocysteine (Brown, 2005; Stubbe, 1994). AdoCbl, on the other hand, is a coenzyme for methylmalonyl-CoA mutase, an isomerase which converts methylmalonyl-CoA to succinyl-CoA (Andrès et al, 2004; Brown, 2005; Clardy et al, 2011; Marsh, 1999). Without the actions of these two B_{12}-dependent enzymes (Figure 1.5) there would be abnormal functioning of the brain and nervous system, due to their requirement for nucleic acid synthesis and the methylation reactions of the brain (Weir and Scott, 1999). A lack of MeCbl leads to a build-up of homocysteine and therefore S-adenosylhomocysteine (SAH), which inhibits SAM-mediated methylation reactions. SAM (synthesised from methionine by

the actions of the enzyme, S-adenosylmethionine synthetase) is required for maintaining methylation reactions, since it is used by many methyltransferases as a methyl group donor. Inhibition of these methylation reactions is thought to cause cobalamin deficiency associated neuropathy. It is this maintenance of methylation reactions by SAM that is the main function of B_{12} in the brain and liver, as well as in rapidly dividing cells, such as the bone marrow, where DNA synthesis is hindered due to a lack of B_{12} (Weir and Scott, 1999). Since blood cells are generated in the bone marrow, an inhibition of DNA synthesis in bone marrow cells will prevent their division and in turn lead to anaemia due to the resultant low levels of red blood cells.

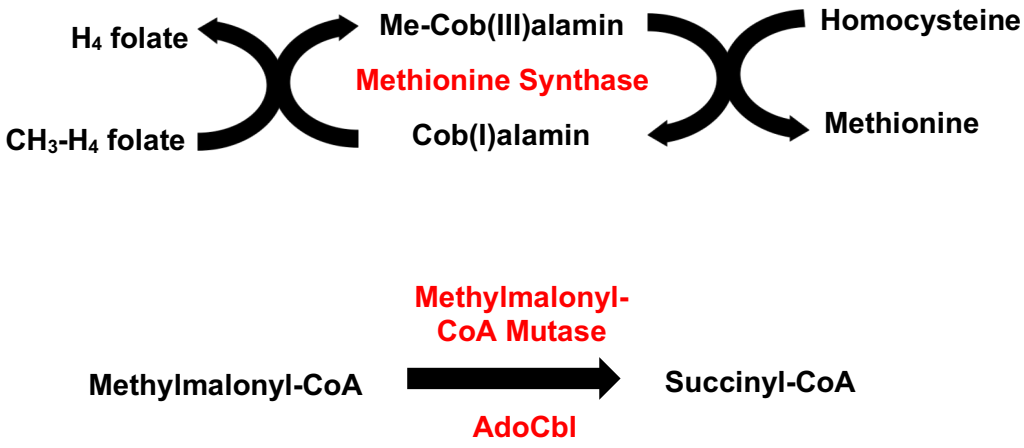


Figure 1.5. The reactions catalysed by the two B_{12} dependent enzymes found in humans, methionine synthase and methylmalonyl-CoA mutase. Methionine synthase catalyses the transfer of a methyl group from methyltetrahydrofolate (CH_3-H_4-folate) to homocysteine to form methionine. Methylmalonyl-CoA mutase catalyses the isomerisation of methylmalonyl-CoA to form succinyl-CoA.

Caenorhabditis elegans is a simple model organism, which has been demonstrated to be a useful tool for the study of B_{12} uptake in eukaryotic organisms (Bito et al, 2013). *C. elegans*, a nematode worm, has been found to contain the same two B_{12}-dependent enzymes as in humans and requires B_{12} for sufficient growth. A B_{12}-deficiency in worms can cause infertility, shorter lifespans and delayed growth (Bito and Watanabe, 2016). In more recent studies, through the use of fluorescent corrin analogues, it has been shown that the organism only takes up the complete cobalamin molecule and not incomplete corrinoids (Lawrence et al, 2018).

Plants are another interesting area of study in the field of B_{12} research, since a major cause of B_{12}-deficiency in humans is the plant-based diet consumed by vegetarians and vegans (Stabler and Allen, 2004) . In plants, there currently appears to be no known role of B_{12} and therefore no known import system developed for the molecule. However, certain photosynthetic organisms, such as microalgae, have been found to take up small amounts of the vitamin (Croft et al, 2005).

1.3.3 Role in prokaryotes

As well as in humans, B_{12} is a molecule that is required by a significant proportion of the bacterial world due to the presence of B_{12}-dependent enzymes in their cells (Shelton et al, 2018). One common subfamily of B_{12}-dependent enzymes in bacteria are the isomerases, which are generally found only in bacteria, except for methylmalonyl-CoA mutase which is found in both bacteria and humans (Banerjee and Ragsdale, 2003). These include the AdoCbl-dependent ribonucleotide reductases (RNRs), which catalyse the conversion of ribonucleotides to deoxyribonucleotides, a step which is essential for DNA replication and repair (Banerjee, 1999; Nordlund and Eklund, 1993), glutamate mutase, which is able to

convert glutamate into methylaspartate and vice versa (Reitzer et al, 1999), and methylmalonyl-CoA mutase which catalyses the isomerisation of methylmalonyl-CoA to form succinyl-CoA, (Figure 1.5). Other isomerases in this subfamily include methyleneglutarate mutase, isobutryl-CoA mutase and lysine 5,6 aminomutase (Banerjee and Ragsdale, 2003) .

Another major class of B_{12}-dependent enzymes in prokaryotes are the B_{12}-dependent methyltransferases, which catalyse the movement of a methyl group from a methyl donor to a methyl acceptor (Banerjee & Ragsdale, 2003). The most extensively studied enzyme of this class is the B_{12}-dependent version of methionine synthase, MetH, which, as mentioned previously, regenerates methionine from homocysteine (Figure 1.5). Other corrinoid-dependent methyltransferases include acyl-CoA synthases and coenzyme M methyltransferases (Matthews et al, 2008).

The radical SAM superfamily of enzymes is a major family of enzymes, of which the B_{12}-dependent enzymes is the largest class (Bridwell-Rabb et al, 2014). The radical SAM family B_{12}-dependent enzymes catalyse methyl transfer, using MeCbl as their methyl source (Frey et al, 2008). An example of one of these enzymes is BcpD, which is involved in the biosynthesis pathway of the herbicide, bialaphos (Frey et al, 2008).

The next well-known class of B_{12}-dependent enzymes includes the reductive dehalogenases, which are involved in the detoxification of aromatic and aliphatic chlorinated organics, including chlorinated phenols, chlorinated ethane and polychlorinated biphenyls (Copley, 1998; Janssen et al, 2001). Bacterial exposure to these chlorinated compounds results in the expression of genes encoding reductive dehalogenases, such that they can remove the chloride ion from the compounds, thus playing an important role in detoxification (Banerjee and

Ragsdale, 2003). This class of enzymes is surprisingly unique in that these enzymes use a form of Cbl that lacks an upper ligand. Recent advances have now shown another group of enzymes, known as the epoxyqueuosine reducatases (QueGs), also belong in this class along with the reductive dehalogenases. QueGs catalyse the reduction of epoxyqueuosine to form queuosine, an RNA nucleoside (Bridwell-Rabb & Drennan, 2017).

Other recent advances have revealed a completely new role of B_{12} as a light sensor. For example, within the bacterial genome of *Myxococcus xanthus* genes are contained which encode carotenoids, pigments that are able to protect the cells from photooxidative damage. Research has shown that the expression of these genes is down-regulated by a transcriptional repressor, CarH, whose activity is specifically directed by the light sensitive B_{12} molecule, AdoCbl. It is during dark conditions when AdoCbl binds to the B_{12}-binding motif found on CarH and forms a tetramer which interacts with the DNA and represses transcription. Upon light exposure, however, the C-Co bond of AdoCbl gets cleaved, causing the tetramer to fall apart, thereby allowing gene expression to occur (Jost et al, 2015; Ortiz-Guerrero et al, 2011). This demonstrates how the light-sensing ability of AdoCbl allows it to play an indirect role in the regulation of gene expression, via its actions upon a transcription factor.

In addition to playing a part in enzyme-catalysed reactions and having a role in light-sensing, B_{12} has also been found to be involved in more direct gene regulation, in many bacteria, via short regions of RNA known as riboswitches (Rodionov et al, 2003). More detail on riboswitches is given later in the chapter. A summary of the different B_{12}-dependent processes, and the versions of B_{12} they require to carry out their function, is seen in Table 1.1.

Table 1.1. Table to show the versions of B_{12} required by each B_{12}-dependent process.

B_{12}-dependent Process	**Required Version of B_{12}**
Isomerisation	AdoCbl
Methylation	MeCbl
Radical SAM reaction	MeCbl (sometimes AdoCbl)
Reductive dehalogenation	OHCbl
Light responsive transcription regulation	AdoCbl
Riboswitch regulation	AdoCbl, OHCbl

1.4 Uptake and transport of B_{12}

1.4.1 Uptake in eukaryotes

In humans, B_{12} is crucial to life but cannot be synthesised *de novo*, for which reason several B_{12}-binding proteins are involved in the process of B_{12} uptake from consumed food and its transport to the appropriate storage site. The first B_{12}-binding protein involved, known as haptocorrin (also referred to as R-protein, cobalophilin and transcobalamin I), is found in the saliva and binds B_{12} straight after it has been ingested in order to protect it from the acidic environment of the stomach (Braun, 2006; Lawrence and Roth, 1996). Despite its sensitivity to acid, the binding of B_{12} to haptocorrin allows the complex to pass through the stomach and into the small intestine. Pancreatic enzymes then digest haptocorrin, so that the released B_{12} molecule can bind the second B_{12}-binding protein known as intrinsic factor (IF). IF-B_{12} complex is absorbed by the epithelial cells of the ileum, following which B_{12} is dissociated once again so that it can bind to the third B_{12}-binding protein, transcobalamin II (TCN2) (Braun and Endriss, 2007) . Finally, the TCN2-B_{12} complex exits the epithelial cells and is transported to the liver (Postle and Larsen, 2007).

As mentioned previously, nematode worms are another eukaryotic organism that require B_{12} for growth and also for the B_{12}-dependent enzymes, methionine synthase and methylmalonyl-CoA mutase (Bito & Watanabe, 2016). They are unable to synthesise the molecule themselves and obtain B_{12} from the bacteria that they digest. However, the method of transport within *C. elegans* is currently not known. When a B_{12}-deficiency is induced in the organism, symptoms resemble those seen in mammalian B_{12}-deficiency, including infertility, delayed growth and a shorter lifespan (Bito & Watanabe, 2016). This makes nematode worms an organism of great interest in the study of B_{12}-deficiency in eukaryotic organisms.

In terms of plants, there is no known requirement for B_{12} and therefore no designated uptake system. Algae, on the other hand, have been shown by recent research to contain a protein that is implicated in B_{12} uptake, CBA1. However, its role and mechanism of action remain unknown (Helliwell et al, 2016). It is thought that half of all algae require B_{12} and likely get it by having a symbiotic relationship with B_{12}-producing bacteria (Croft et al, 2005). It has been known for years that bacteria are able to provide B_{12} to other bacteria or algae in mixed consortia. However, it is not known how bacteria transport B_{12} out of their cells. Since there is currently no known export system for B_{12} from bacterial cells, research into this area could lead the way towards finding one.

1.4.2 Uptake in prokaryotes

Around 86% of bacteria are thought to have at least one enzyme belonging to a B_{12}-dependent family of enzymes, yet only 37% are predicted to synthesize B_{12} *de novo* (Shelton et al, 2018). Those that are able to synthesize their own B_{12}, such as *Bacillus megaterium* (Moore & Warren, 2012), do not rely on uptake of the vitamin. However, since the molecule is energetically very expensive to synthesise, bacteria will acquire it from their environment when possible. Those that do not have this ability to synthesize B_{12} *de novo* rely on taking up the vitamin from their surroundings.

Bacteria such as *E. coli* are unable to make their own B_{12}, and therefore rely on the uptake of B_{12} from an external source using a tightly-regulated import system (Lewinson et al, 2010). Bacteria have evolved a number of B_{12} import systems over time, but currently the best characterized one is the BtuB import system in *E. coli*.

B_{12} uptake in *E. coli* relies on a large complex known as the TonB complex, consisting of three proteins (TonB, ExbD and ExbB) which span the periplasm and provide the proton motive force (from the cytoplasmic membrane) required to take up B_{12} via the TonB-dependent transporter (TBDT), an outer membrane transporter known as BtuB (Johnson Jr et al, 2012; Lundrigan et al, 1991; Nahvi et al, 2002). B_{12} is a complex molecule that needs to be actively transported across the membrane with the help of transporter proteins. BtuB is the first of these proteins which imports B_{12} from outside the cell, using energy provided by the TonB complex, and transports it to the periplasm (Cadieux et al., 2002). Once in the periplasm, the periplasmic protein BtuF binds B_{12} tightly and delivers it to the BtuC/D complex, found in the inner membrane (Lewinson et al., 2010) (Figure 1.6).

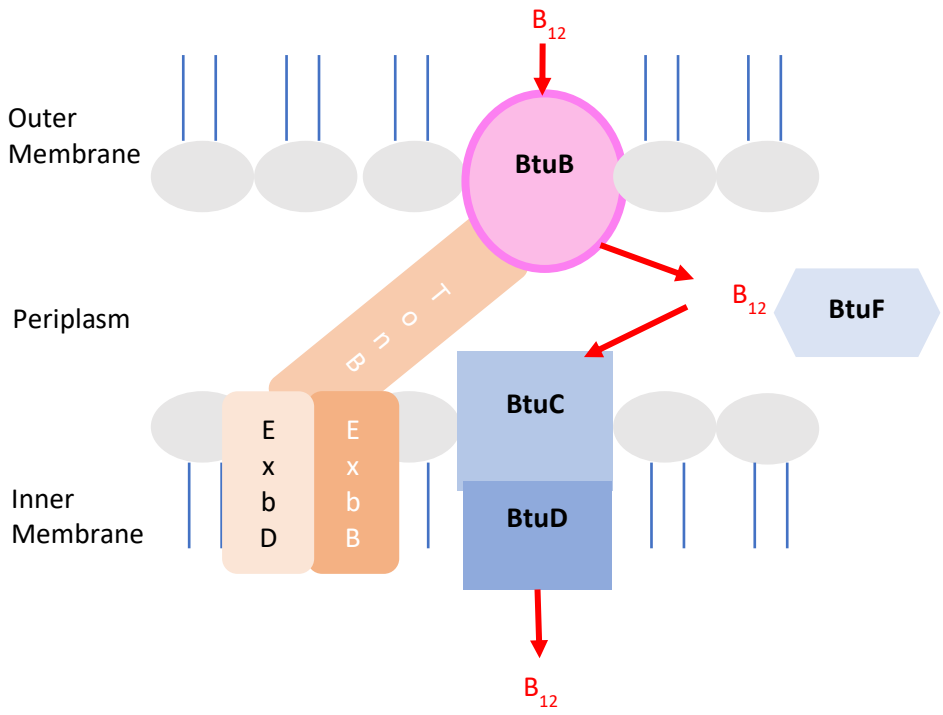


Figure 1.6. Diagram representing the B_{12} import system in *E. coli*. B_{12} is transported into the periplasm by the outer membrane transporter, BtuB, which is powered by the TonB complex (TonB, ExbB and ExbD). It is then bound by the periplasmic protein, BtuF, which transports it to the BtuC/D complex. The BtuC/D complex allows its import into the cytoplasm.

BtuC is a transmembrane protein and BtuD is an ATP Binding Cassette (ABC) transporter (Cadieux et al, 2002), and together they form the BtuC/D complex. BtuC/D cannot bind free B_{12} very well on its own, unlike BtuF which has a very high affinity for the molecule, and therefore requires BtuF to deliver B_{12} directly to its complex. Once the BtuF-B_{12} complex reaches BtuC/D, both complexes associate and undergo a conformational change whereby the BtuC/D complex causes BtuF to release B_{12} and allows its internalisation into the cytoplasm through BtuC/D. In the final stage, energy from ATP is required to release BtuF from the complex and this is provided when the ATP bound to BtuD is hydrolysed (Lewinson et al, 2010). BtuF is then free to re-initiate the cycle of B_{12} delivery.

1.5 B_{12} Riboswitch

1.5.1 Riboswitch overview

Riboswitches are highly structured regulatory elements, found within mRNA sequences, that are bound to by a ligand or metabolite to terminate transcription or translation (Breaker, 2012; Serganov and Nudler, 2013). One of the first discovered riboswitch types was the B_{12} riboswitch, where binding of B_{12} results in the switching off of expression of B_{12}-related genes. Following this, many other riboswitches, including the thiamine pyrophosphate (TPP) and S-adenosylmethionine (SAM) riboswitches, were also identified and studied.

There are three common mechanisms utilised by riboswitches to prevent expression of the downstream gene(s), which include transcription termination, prevention of translation initiation or splicing (in eukaryotes) (Breaker, 2012) (Figure 1.7).

In the first mechanism, the formation of a strong terminator stem followed by a run of uridine residues results in the generation of an intrinsic transcription terminator (Gusarov and Nudler, 1999; Yarnell and Roberts, 1999). The formation of this terminator stem is controlled by ligand binding to the riboswitch, which causes RNA polymerase to stall transcription and release any associated substrates or products (Breaker, 2012) (Figure 1.7A). Similarly, riboswitches that control translation initiation also work through a change in mRNA structure, where the formation of a hairpin structure, that sequesters the ribosome-binding site (RBS), prevents binding of the ribosome and therefore significantly reduces the likelihood of translation initiation (Nou and Kadner, 1998) (Figure 1.7B). Finally, another class of riboswitches found in eukaryotes, including some of those controlled by TPP, have

been shown to control gene expression through splicing (Cheah et al, 2007; Kubodera et al, 2003) (Figure 1.7C).

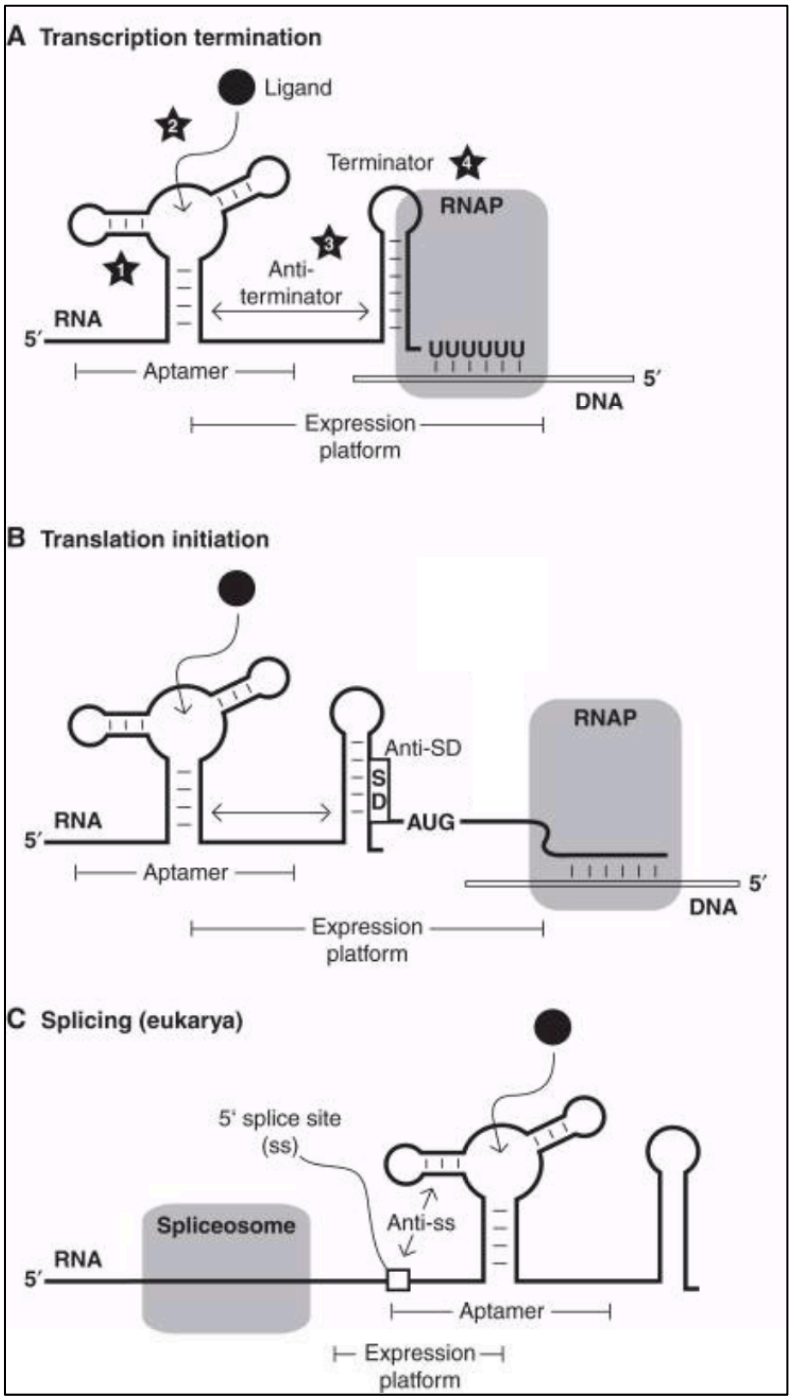


Figure 1.7. Diagram representing the three common mechanisms utilised by riboswitches to prevent expression of downstream gene(s). A) Transcription termination, B) termination of translation initiation, C) gene splicing (eukaryotes). Image adapted from Breaker, 2012.

1.5.2 B_{12}-dependent riboswitches

The *btuB* riboswitch is one of a number of different B_{12}-dependent riboswitches. Other examples of B_{12} binding riboswitches include the *metE* riboswitch, which controls production of a B_{12}-independent methionine synthase (encoded by *metE*) (Warner et al, 2007), and the Pacific Ocean metagenome env8 riboswitch controlled by the less common ligand, OHCbl (Johnson Jr et al, 2012).

The *metE* riboswitch is controlled by AdoCbl, which binds to a 'B_{12} element' on the riboswitch to switch it off via regulation at the transcriptional level (Rodionov et al, 2003). This results in a decrease in the levels of B_{12}-independent methionine synthase (MetE) found within cells. This is of particular importance when levels of AdoCbl are high in the cell, as it has been shown that the B_{12}-dependent version (MetH) is over 100 times more active than MetE and therefore, there is a metabolic advantage of using MetH instead of MetE when possible (i.e. when levels of B_{12} inside the cell are high) (Gonzalez et al, 1992; Jeter et al, 1984).

Currently, two B_{12} riboswitches have had their structures solved using X-ray crystallisation techniques, one from *Symbiobacterium thermophilum* (Peselis and Serganov, 2012) and another from *Thermoanaerobacter tengcongensis* (*Tte*) (Johnson Jr et al, 2012) (Figure 1.8). The AdoCbl riboswitch has a unique structure, allowing for AdoCbl to bind (Figure 1.8). The AdoCbl riboswitch has three main regions; the peripheral region which is thought to be involved in ligand recognition, the core region which contains the ligand binding site, and the accessory region (often referred to as the expression platform), which changes structure in response to ligand binding at the core region. Comparative sequence analysis studies suggest that the core region, containing the AdoCbl binding site, is highly conserved through evolution (Peselis and Serganov, 2012).

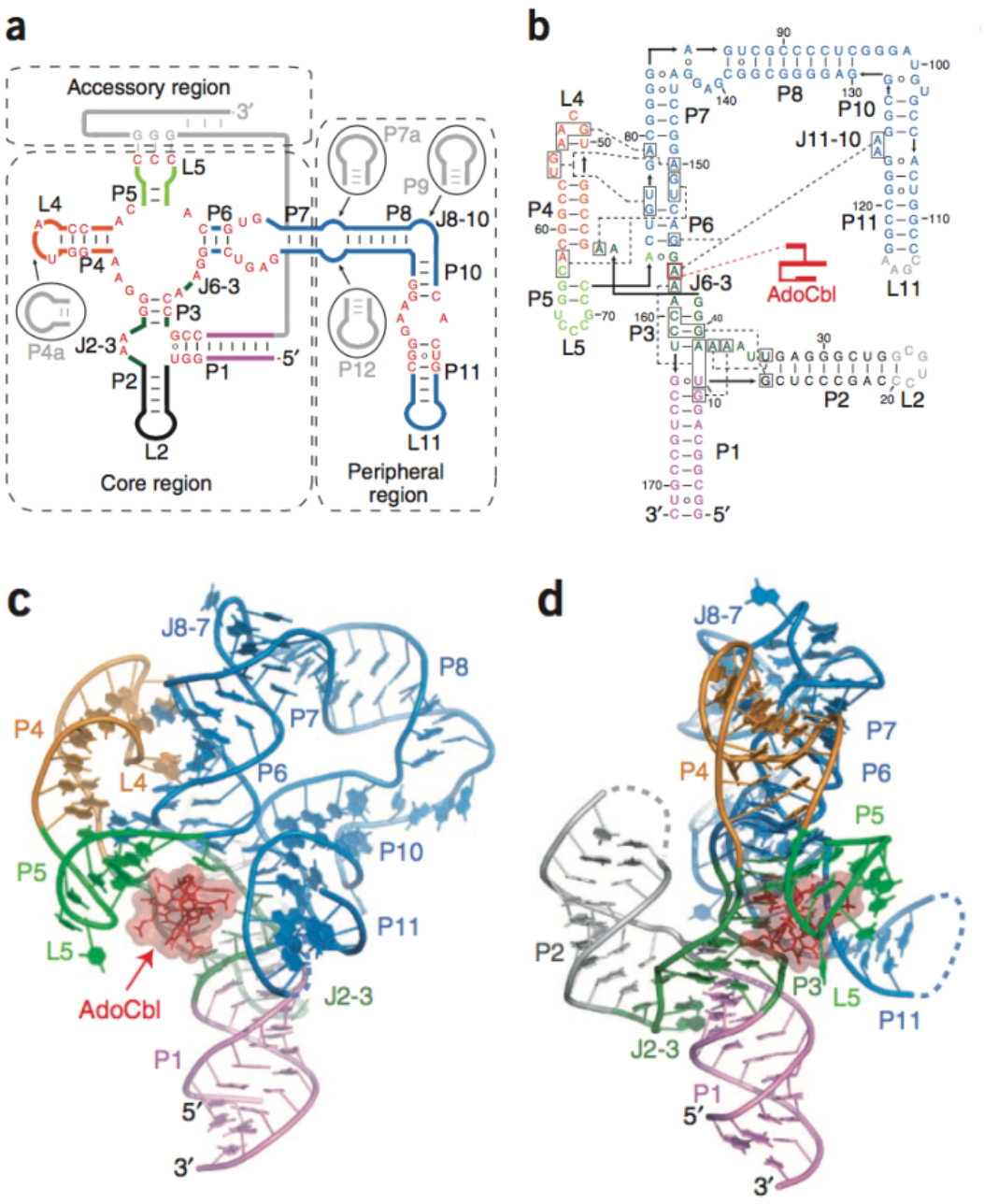


Figure 1.8. ***S. thermophilum*** **AdoCbl riboswitch structure**. (a) Consensus secondary structure of the AdoCbl riboswitch. Red nucleotides are conserved in more than 90% of species. (b) Schematic of the riboswitch fold observed in the crystal structure of the complex. Key long-distance tertiary base-pairing and stacking interactions are shown as dashed lines. (c, d) Ribbon representation of overall AdoCbl riboswitch structure. Adapted from: (Peselis and Serganov, 2012)

1.5.3 *E. coli btuB* riboswitch

Initial work on the *E. coli btuB* riboswitch was initiated when levels of *btuB* expression within *E. coli* cells were found to be controlled by an mRNA leader sequence starting 240 nucleotides upstream of the start of translation. In addition to this, *btuB* sequences between +156 and +365 were also seen to strongly influence gene expression in response to B_{12} levels (Lundrigan et al, 1991). At the same time, Lundrigan *et al* showed that despite the role of BtuB in uptake of B_{12} into cells, it does not appear to be directly involved in its own regulation. Within the mRNA leader sequence of the *btuB* gene lies the B_{12} box (Figure 1.9) which is the region to which B_{12} is thought to bind (Franklund and Kadner, 1997) and, in particular, with AdoCbl as the preferred ligand (Lundrigan and Kadner, 1989).

Regulation by B_{12} operates at both the transcriptional and translational level, with both types suggested to have inhibitory elements that act independently (Franklund and Kadner, 1997). However, primary control over *btuB* expression was found to be at the level of translation initiation, where the formation of a hairpin structure sequesters the RBS such that the ribosome can no longer bind (Nou and Kadner, 1998).

The mRNA leader sequence, in the 5' untranslated region (UTR) of the *btuB* gene, was later discovered to be an AdoCbl-responsive riboswitch. This *btuB* riboswitch controls *btuB* gene expression in response to AdoCbl binding. AdoCbl was found to bind directly to the mRNA and induce a structural reorganisation, which results in a change in expression levels (Nahvi et al, 2002).

The *btuB* riboswitch is thought to have two main domains; the aptamer domain (containing the peripheral region and core region, as mentioned above) and the

expression platform (Gruber et al, 2011) (Figure 1.9). The aptamer domain is highly conserved amongst various organisms and it is within the core region that the B_{12} box is located (Nahvi et al, 2004). Once AdoCbl binds at the B_{12}-box, the riboswitch structure is thought to change at the expression platform such that the RBS gets sequestered and is no longer accessible by the ribosome (Gruber et al, 2011) thus highly reducing the likelihood of translation initiation (Figure 1.9).

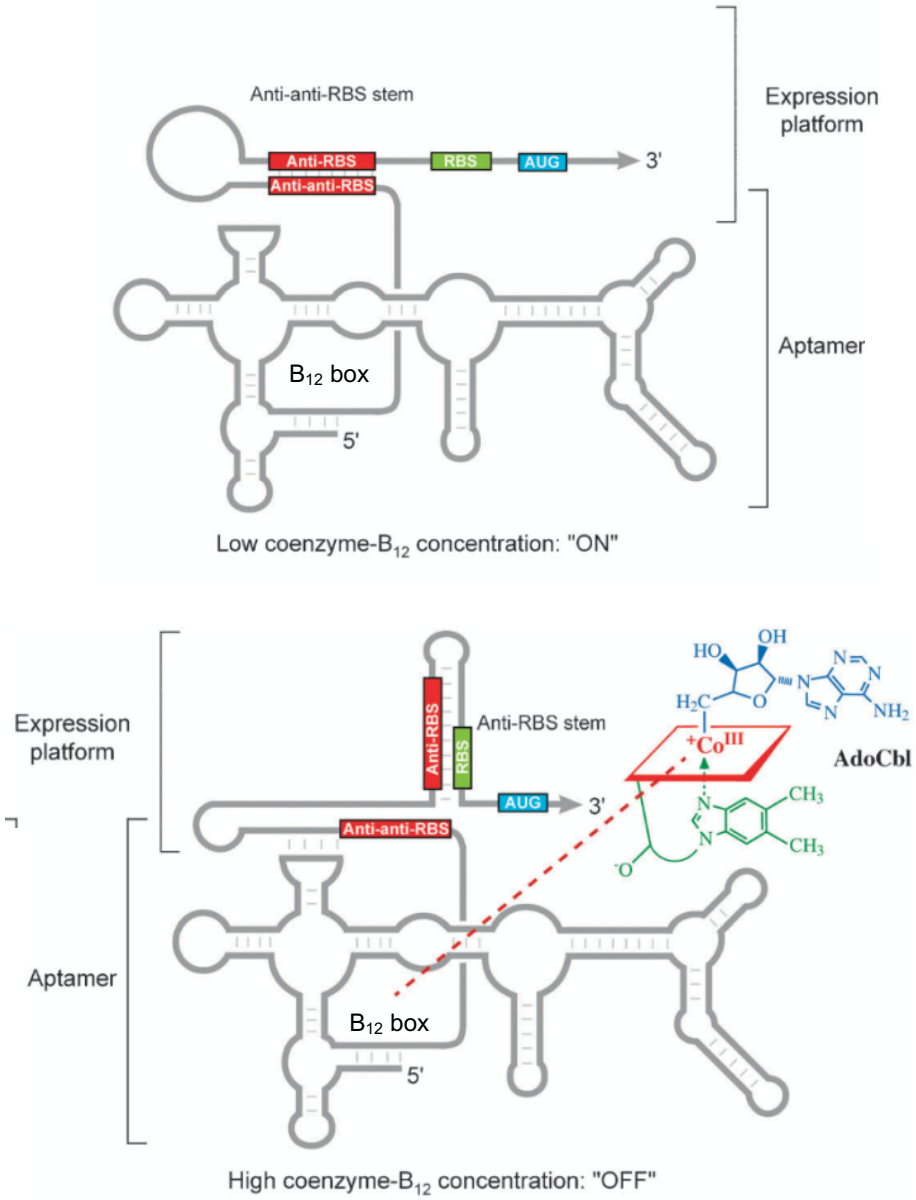


Figure 1.9. *E. coli btuB* riboswitch proposed model. Adapted from: (Gruber et al. 2011). The *btuB* riboswitch regulates translation of BtuB by AdoCbl-binding. At high AdoCbl (coenzyme B_{12}) concentrations, the aptamer domain changes structure such that the anti-anti-RBS stem is interrupted. As a result, the anti-RBS stem is formed that sequesters the RBS and prevents association with the ribosome. Thus, translation of mRNA to form BtuB is inhibited.

Previous work, largely involving in-line probing experiments (a method that comprises of cleaving RNA molecules under different conditions, for example, with or without a potential ligand, and comparing differences in cleavage patterns to see whether potential ligands have in fact bound to the RNA to protect it from degradation), demonstrates that the riboswitch directly binds B_{12} and has a preference for binding AdoCbl. Some of these experiments even suggest that certain B_{12} types, such as CNCbl, are unable to switch off the riboswitch (Nahvi et al, 2002; Nou and Kadner, 2000). This is thought to be due to a requirement for the 5'-deoxyadenosyl group at the upper ligand position, in order to bind to the B_{12}-binding pocket and allow structural modulation of the RNA (Nahvi et al, 2002).

Later studies used the same in-line probing technique to compare the binding affinities of a wider range of corrinoid compounds and also determine which particular region of AdoCbl is responsible for the structural changes undergone by the *E. coli btuB* riboswitch. Findings from these studies indicate that it is the corrin ring that is responsible for structural changes within the riboswitch, rather than the upper or lower ligand (Gallo et al, 2008), due to the fact that four different B_{12} variants (including CNCbl), with varying upper and lower ligands, were all found to induce essentially the same conformational changes in the riboswitch structure. These studies also indicate that binding to the riboswitch is completely independent of whether the molecule is in the "base-on" or "base-off" conformation. However, the upper and lower ligands have an important initial role in effectively binding the riboswitch and reaching a highly stable complex with the RNA (Gallo et al, 2008), demonstrated by the difference in binding affinities of the variants. This role of the axial ligands in binding affinity was confirmed in later studies (Gallo et al, 2011).

1.5.4 Latest developments in riboswitch research

In more recent times, researchers have been able to exploit the fundamental knowledge acquired on riboswitches in order to design ones with medicinal or biotechnological applications. One such example involves the recent synthesis of functional riboswitches that are able to detect explosives, due to the design of an aptamer domain that is able to bind a component known as 2,4-dinitrotoluene (2,4-DNT), commonly found in explosives (Borujeni et al, 2015).

In other applications, riboswitches are being utilised as powerful RNA imaging tools to elucidate mechanisms such as those underlying the dynamics of the mRNA and ncRNA life cycle. Here, researchers developed an RNA imaging tool consisting of a cobalamin-linked fluorophore and a B_{12}-riboswitch containing RNA tag (Figure 1.10). Within the cobalamin linked fluorophore, cobalamin quenches the fluorescence of the fluorophore until cobalamin comes into close proximity with the riboswitch in the RNA tag. When cobalamin binds the riboswitch, it is no longer able to quench the fluorescence of the fluorophore and therefore the fluorescence can now be detected (Braselmann et al, 2018) (Figure 1.10).

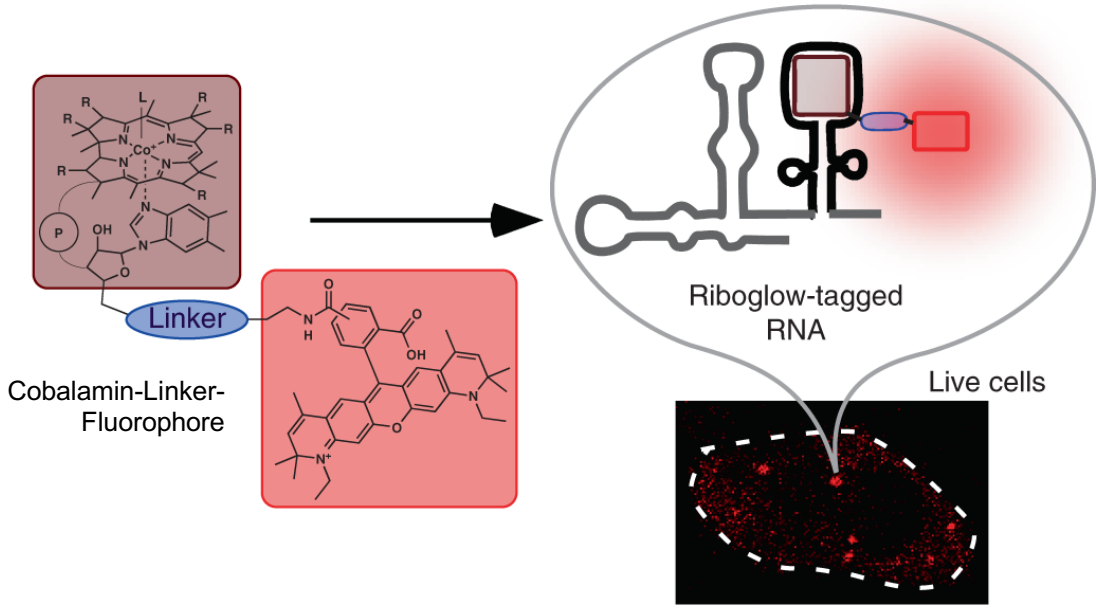


Figure 1.10. Diagram representing the use of a B_{12}-dependent riboswitch in RNA imaging. RNA imaging tool using a cobalamin-linked fluorophore and a riboswitch-containing RNA tag. Image adapted from Braselmann *et al*, 2018.

In addition to this, riboswitches controlling key metabolic systems in pathogenic organisms can be exploited through the use of metabolite/ligand analogues for the development of antibacterial compounds (Blount and Breaker, 2006). These advances suggest that the continued discovery and analysis of riboswitches will allow progress in creating more useful molecular tools.

1.6 Aims and objectives

The aim of this thesis is to investigate the possibility of engineering the production of B_{12} variants, through the construction of an empty, metal-free B_{12} shell, and exploring what effect such compounds might have on cellular control elements, such as riboswitches. In order to study this, a technique was designed and developed that utilises an *in vitro* system to determine the effect of B_{12} variants on the well-characterised B_{12}-dependent *btuB* riboswitch. This method allows a quick and easy way of testing the ability of B_{12} variants to directly bind to and deactivate the *btuB* riboswitch, without relying on indirect cellular effects.

The production of such a technique required the construction of a plasmid in which the *btuB* riboswitch was placed upstream of the eGFP reporter gene. This plasmid was combined with *in vitro* transcription/translation machinery in the presence of different B_{12} variants of both natural and unnatural origin, to see how each affects the levels of eGFP production. The effects of many of these variants on the riboswitch have not yet been well studied. Of these unnatural ligands, AdoRbl is one that has previously been suggested to interfere with the riboswitch. However, this idea lacked firm evidence (Widner et al, 2016). The idea was investigated in this project to determine whether AdoRbl does in fact interfere with the riboswitch, side by side with other B_{12} variants. Should AdoRbl, or any of the other unnatural B_{12} analogues, be able to do this will reflect on whether these analogues have antivitamin potential (the ability to interfere with and inhibit cellular processes, without carrying out the same functions as B_{12}). Thus, the research carried out here has the potential to open up an exciting new field of research into the use of B_{12} analogues in medical advances and drug design.

Chapter 2: Materials and Methods

2.1 Microorganisms and cloning

2.1.1 Bacterial strains

Table 2.1. Bacterial strains.

Strain	Genotype	Source
JM109	e14$^-$(McrA$^-$) recA1 endA1 gyrA96 thi-1 hsdR17 (rK$^-$ mK$^+$) supE44 relA1 Δ(lac-proAB) [F´ traD36 proAB laclqZΔM15]	Promega
MG1655	F$^-$, lambda$^-$, rph-1	Dr Rokas Juodeikis
Rosetta (DE3)	F$^-$ ompT hsdSB(rB$^-$ mB$^-$) gal dcm (DE3) pRARE (CamR)	Novagen
DH5α	F$^-$ φ80 lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17 (rK$^-$,mK$^+$) phoA supE44 λ$^-$ thi-1 gyrA96 relA1	Thermo Fisher
NovaF$^-$	F$^-$ endA1 hsdR17 (rK12$^-$ mK12$^+$) supE44 thi-1 recA1 gyrA96 relA1 lac	Novagen
BL21*(DE3)-pLysS	F$^-$ ompT hsdSB (rB$^-$, mB$^-$) gal dcm (DE3) pLysS (CmR)	Novagen
E. coli ED661	K-12 (DE3); λ$^-$ F$^-$ rph-1, T7RNAP. All the genes for HBAH synthesis were integrated within the genome under the control of the T7 promoter.	Dr Evelyne Deery

2.1.2 Plasmids

Table 2.2. Plasmids.

Plasmid	Description	Source
pET3a	Overproduction vector with T7 promoter, Amp^R. With added SpeI site	Dr Evelyne Deery
pET14b	Overproduction N- terminal His$_6$tag fusion protein vector with T7 promoter, Amp^R. With added SpeI site	Dr Evelyne Deery
cobC-pET14b	pET14b with *A. vinosum cobC* gene	Dr Andrew Lawrence
cobD-pET14b	pET14b with *A. vinosum cobD* gene	Dr Andrew Lawrence
cobD-pET23b	pET23b with *A. vinosum cobD* gene	Dr Andrew Lawrence
*cobAIG*JFMKLHBQ*$^{(Alv)}$-pET3a	pET3a with *R. capsulatus* cobAIJFMKLHB genes, *B. melitensis* cobG and *A. vinosum* cobQ	Dr Emi Nemoto-Smith
pETcocoR	Modified pETcoco plasmid. Protein overproduction vector, regulation of the	Dr Rokas

	copy number with glucose and arabinose: dual origins of replication control: single-copy (oriS) and medium copy (oriV) states, Amp^R	Juodeikis
eGFP-pET14b	pET14b with *eGFP* gene	Dr Matthew Lee

2.1.3 Primers

Table 2.3. Primers.

Name	Sequence	Restriction Site
CobP_For	CAT CAT ATG ATT GGC GAA AAC GAA AGG	NdeI
CobP_Rev	ACT ACT AGT CAG CGA TCG ATG GGA ACG	SpeI
CobP_Rev2	ATA CTC GAG GCG ATC GAT GGG AAC GCC	XhoI
CobV_For	ACT CAT ATG ACA CGA CTC CGC CCA C	NdeI
CobV_Rev	ACT ACT AGT CAC TCG ATG AGG ATC GC	SpeI
CobU_For	CAC CAT ATG AAC ATG ACG CTC GAC	NdeI
CobU_Rev	ACT ACT AGT CAC GCC TCG CTC ACG CC	SpeI
Rib_For	GCT TCT AGA GCC GGT CCT GTG AGT TAA TAG G	XbaI
Rib70_Rev	GAT CAT ATG CGG AAG ACG GCG CAG CAC	NdeI
Hyb2_For	GA TGC TTT ACA ATG AGC AAA GG	n/a
Hyb2_Rev	CC TTT GCT CAT TGT AAA GCA TC	n/a
eGFP_Rev	AGC ACG CGT CTT GTA GTT CCC GTC	MluI
M_loti_For	GCT TCT AGA CCT AAA TCC GCT CCA GAC GGT C	XbaI
M_loti_Rev	GAT CAT ATG GGC CAT GGC TTC GCC GTC GGC	NdeI
NKDNA_For	GTG CCA CCT GAC GTC TAA G	n/a
NKDNA_Rev	CGA CTC CTG CAT TAG GAA G	n/a

2.2 Media for growing microorganisms

2.2.1 Antibiotics

Table 2.4. Antibiotics.

Antibiotic	Stock concentration	Working concentration
Ampicillin	100 mg/mL in dH_2O	100 mg/L
Chloramphenicol	34 mg/mL in ethanol	34 mg/L

2.2.2 Luria-Bertani (LB) medium

10 g tryptone, 5 g yeast extract and 5 g NaCl, were made up to 1 L with distilled H_2O (dH_2O). This was then autoclaved (heated to 121°C for 15 minutes in an autoclave).

2.2.3 LB agar

Made up as LB medium, with 15 g technical agar also added. After autoclaving, LB agar was cooled to ~50°C and poured into Petri-dishes (~25 mL per plate). For LB agar with antibiotics, antibiotics are added after cooling to ~50°C, before pouring.

2.2.4 Mg^{2+} Stock (2 M)

2 g $MgCl_2$,$6H_2O$ and 2.5 g $MgSO_4$,$7H_2O$ were combined before making up to 10 mL with dH_2O. This was then filter-sterilised.

2.2.5 SOC medium

2 g tryptone, 0.5 g yeast extract, 1 mL 1 M NaCl and 0.25 mL 1 M KCl was made up to 100 mL with dH_2O. This was then autoclaved. After autoclaving, 1 mL 2 M Mg^{2+} stock (filter-sterilised) and 1 mL 2 M glucose solution (autoclaved) were added.

2.2.6 10X M9 salts

60 g Na_2HPO_4, 30 g KH_2PO_4, 10 g NH_4Cl and 5 g NaCl were weighed out and made up to a litre with dH_2O before autoclaving.

2.2.7 M9 minimal medium

100 mL 10X M9 Salts, 20 mL 20% (w/v) glucose (autoclaved), 2 mL 1 M $MgSO_4$ (filter-sterilised) and 1 mL 0.1 M $CaCl_2$ (filter-sterilised) were combined and made up to 1 L with 900 mL autoclaved dH_2O.

2.2.8 2YT medium

16 g tryptone, 10 g yeast extract (YE) and 5 g NaCl, were made up to 1 L with dH_2O. This was then autoclaved.

2.2.9 M9-YE medium

50 mL 10X M9 salts, 5 mL 50% glycerol, 5 g YE, 1 mL 1 M $MgSO_4$ and 0.5 mL 0.1 M $CaCl_2$ were made up to 500 mL with dH_2O before autoclaving.

2.3 Reagents and protocols for molecular biology

2.3.1 Preparation of competent *E. coli* cells (DH5α, NovaF$^-$, *E. coli 661*, BL21*(DE3)-pLysS, BL21*(DE3)-pLysS-*btuB* strains)

The appropriate bacterial strain was re-streaked from a single bacterial colony onto an LB agar plate with the appropriate antibiotics, aseptically, and incubated at 37°C overnight.

40 mL LB medium was inoculated with 1 bacterial colony and grown to OD_{600} = 0.3. This was chilled on ice for 10 minutes before centrifugation at 3000 rpm for 10 minutes, at 4°C. The pellet was resuspended in 20 mL ice-cold 0.1 M $CaCl_2$ (filter-sterilised 0.2 μm) and incubated on ice for 20 minutes. Cells are centrifuged again, as before, and gently resuspended in 2 mL ice-cold 0.1 M $CaCl_2$/15% glycerol (filter-sterilised 0.2 μm). Cells were dispensed as 50 μL aliquots into chilled cryotubes, flash-frozen on dry ice and stored at -80°C.

2.3.2 Transformation of competent cells

Competent cells (50 μL) were thawed and 5 μL 0.1 μg/μL plasmid (or ligation mixture) added. Cells were incubated on ice for 15 minutes, at 42°C for 90 seconds, then immediately on ice for 2 minutes. 200 μL SOC was added, followed by incubation at 37°C for 1 hour. Cells were then spread across LB agar plates with appropriate antibiotics and grown overnight at 37°C.

2.3.3 Preparation of overnight starter cultures

5 mL LB media is inoculated with a single bacterial colony and antibiotics at their appropriate working concentrations, before incubating at 37°C and 160 rpm in a shaking incubator, overnight.

2.3.4 Preparation of pETcocoR starter cultures

5 mL LB medium was inoculated with a single bacterial colony. Ampicillin was added at the appropriate working concentration as well as 0.2% glucose in order to maintain a low plasmid copy number. This was incubated at 30°C and 160 rpm in a shaking incubator, overnight.

The following day, 100 μL of the starter culture was transferred into a tube of fresh 5 mL LB (plus ampicillin) and grown at 37°C and 160 rpm in a shaking incubator until an OD_{600} of 0.2-0.4 is reached.

Once the appropriate OD was reached, L-arabinose was added at a final concentration of 0.01% in order to induce multiplication of the pETcocoR copy number. This was grown in the same conditions for 4-5 hours before centrifuging at 3000 rpm for 10 minutes. The pellet can be used for plasmid extraction.

2.3.5 TE buffer

10 mM Tris-HCl pH 8.0 and 1 mM EDTA were made up to 50 mL with sterile dH_2O, before autoclaving.

2.3.6 Plasmid DNA purification using the QIAprep Spin Miniprep Kit (Qiagen)

Bacterial cells from a starter culture were pelleted via centrifugation at 3000 rpm for 10 minutes. Plasmid DNA was then purified using the QIAprep Spin Miniprep Kit and a microcentrifuge, using the associated kit protocols (https://www.qiagen.com/handbooks). However, the final DNA elution stage was carried out using 50 µL 1X TE buffer.

2.3.7 DNA amplification using polymerase chain reaction (PCR)

PCR was set up as per Table 2.5.

Table 2.5. Standard PCR.

PCR reagents	Volumes (µL)
Fast Start High Fidelity buffer x10 (from Sigma)	5
DMSO	-
Rib_For Primer @ 10 uM	2
Rib70_Rev Primer @ 10 uM	2
dNTPs @ 2 mM	5
Template	20
H_2O	15
Fast Start High Fidelity enzyme (1U/µL, from Sigma)	1

Using a PCR machine, samples were heated as follows:

- 2 minutes at 96°C
- 20 cycles of: 30 s at 96°C, 30 s at 55°C, and x minutes at 72°C (x = 1 minute per 1000 bp).

- 7 minutes at 72°C

2.3.8 Preparation of 1% agarose gels

0.5 g agarose was dissolved in 50 mL 1X TAE buffer (diluted from commercially obtained 50X TAE buffer, from Thermo Scientific) by heating. Once cooled to ~60°C, 1 drop of ethidium bromide was added before pouring.

2.3.9 Agarose gel electrophoresis

The agarose gel was covered with 1X TAE buffer. DNA samples were then mixed with 6X loading dye (from the Qiagen Miniprep Kit) before 12-20 µL samples were loaded alongside 5 µL DNA ladder (Bioline hyperladder 1kb), and run at 140 V for 1 hour.

2.3.10 DNA purification by centrifugation

For gel extraction or PCR clean-up, the Qiagen Gel Extraction kit was used and the associated standard protocol was followed, with DNA eluted in 30 µL nuclease-free H_2O.

2.3.11 Restriction enzyme digest

5 µL DNA, 1 µL of appropriate 10X buffer and 0.5 µL 10 U/µL restriction enzyme(s), also from Promega, were combined in an Eppendorf tube and made up to 10 µL (with dH_2O) before incubating at 37°C for 60 minutes.

2.3.12 DNA ligation

5 μL 2X ligase buffer (Promega), 2.25 μL 100 ng/μL vector, 2.25 μL 100 ng/μL insert and 0.5 μL 10 U/μL T4 DNA ligase (Promega) were combined in an Eppendorf tube before incubating at room temperature for 2 hours.

2.3.13 'Link and Lock' multiple cloning strategy

This method allows the formation of multiple gene constructs by the consecutive cloning of genes into the pET3a plasmid, and was developed by Dr. Evelyne Deery and Dr. Helen McGoldrick (McGoldrick et al, 2005). This involves creating compatible cohesive ends on two different plasmids, using different restriction enzymes, which can fuse together to join two or more genes while creating a site that can no longer be cleaved by either of the original enzymes.

2.3.14 Test to verify cobalamin production

Overnight 10 mL LB starter cultures were prepared for each transformant strain. After overnight growth, 1 mL samples from each starter were centrifuged (5000 rpm, 2 minutes) and the supernatant removed. The pellet can then be checked for colour indicating cobalamin production by bacterial cells.

2.3.15 *In vitro* transcription/translation using the PURExpress *in vitro* protein synthesis kit

Plasmid DNA samples were prepared as normal, except that they were eluted in 25 μL EB buffer instead of 50 μL TE buffer. The lower elution volume allows the DNA samples to be more concentrated (150 ng/μL as per measured on a NanoVue

spectrophotometer), and the use of EB buffer allows them to be free of EDTA (which is thought to interfere with the protein synthesis kit components). Within the 25 μL reaction, there is a 2.5 μL allocation for nuclease-free water that can be substituted with cobalamin in order to investigate its effect on eGFP production and, in turn, on the riboswitch.

Reactions were prepared and undergone as per the associated kit protocol:

1. The necessary number of aliquots of solutions were thawed on ice. They were then centrifuged in a microfuge to collect solutions to the bottom of tubes.
2. The reactions were assembled on ice in a new tube in the following order:
    - 10 μL Solution A
    - 7.5 μL Solution B
    - 2.5 μL Nuclease-free H_2O (or B_{12} variant substitute)
    - 5 μL 150 ng/μL Template (plasmid) DNA containing a T7 promoter and T7 terminator
3. The reaction was mixed and then centrifuged in a microfuge.
4. The samples were incubated at 37°C for 4 hours.
5. The reaction was stopped by placing the tube(s) on ice.
6. The samples were used directly for analysis or frozen at –20°C for use at a later time.

Once the reactions were complete and ended by incubation on ice, 20 μL of reaction mix was removed and combined with the same volume of 2X Laemmli SDS buffer, before boiling for 5 minutes. Samples of 10 μL were then loaded onto a 12.5% SDS PAGE for analysis using SDS-PAGE.

2.3.16 Reverse riboswitch growth assays

Assays were set up in 5 mL minimal medium using transformants carrying one of each of the two plasmid variants. Each transformant was grown in two different conditions with each carried out in duplicate.

Following overnight growth, pellets were resuspended in 1 mL H_2O and boiled for 15 minutes to break open the cells. 200 μL samples were then removed and prepared for SDS-PAGE analysis. SDS-PAGE analysis was then carried out in order to compare eGFP levels in each sample.

2.3.17 Adenosylation of Cbi using BtuR

The assay was prepared in a universal Falcon tube using the following components; 100 μM Cbi, 1 mM ATP, 2.5 mM $MnCl_2$, 63 μM BtuR and 10 mM $NaBH_4$, before making up to a volume of 20 mL with PD10 buffer. The Falcon tube was then wrapped in foil, due to the light sensitivity of AdoCbi and to prevent removal of the adenosyl group soon after AdoCbi is made. This was then incubated overnight in a glove box.

2.3.18 '*In vivo*' CobP activity assay

BL21*(DE3)-pLysS-*btuB_CDPVU-pET3a* colonies were used to set up 2 different 50 mL LB cultures (with the appropriate antibiotics), and cells were grown at 37°C for 5 hours in a shaking incubator at 180 rpm, in the dark.

Following 5 hours of growth, dicyanocobinamide (obtained from Sigma-Aldrich) was added to the cultures at a final concentration of 100 pM.

After 2 hours of growth in the presence of dicyanocobinamide, protein production was induced with the addition of 0.4 mM IPTG and the cultures left to incubate overnight; one at 37°C and the second at 19°C.

On the following day, both cultures were centrifuged at 4000 rpm for 15 minutes in order to separate the supernatant and pellet. Samples were taken from both supernatants for analysis using HPLC-MS. In regards to the pellets, both were initially resuspended in 100 μL H_2O before the addition of 400 μL methanol (as cobalamin is soluble in methanol). These were then heated to 60°C, under a fume hood, for 15 minutes before centrifuging at 13000 rpm for 1 minute. The resulting supernatants were then removed and dried down in a vacuum centrifuge in order to remove all traces of methanol before running samples on a HPLC-MS instrument. Once dried down, they were resuspended in 100 μL ultra-pure H_2O for analysis using HPLC-MS.

2.4 Reagents and protocols for biochemistry

2.4.1 Solutions for immobilised metal ion affinity (IMAC)

Charge buffer - 50 mM $NiSO_4$

Binding buffer - 20 mM TRIS-HCl pH 8.0, 500 mM NaCl, 5 mM imidazole

Wash buffer I - 20 mM TRIS-HCl pH 8.0, 500 mM NaCl, 50 mM imidazole

Wash buffer II - 20 mM TRIS-HCl pH 8.0, 500 mM NaCl, 100 mM imidazole

Elution buffer - 20 mM TRIS-HCl pH 8.0, 500 mM NaCl, 400 mM imidazole

Strip buffer - 20 mM TRIS-HCl pH 8.0, 500 mM NaCl, 100 mM EDTA

PD10 buffer - 20 mM TRIS-HCl pH 8.0, 100 mM NaCl

2.4.2 Recombinant protein production

Day 1: *E. coli* BL21*(DE3)-pLysS cells were transformed with a plasmid containing a gene for the recombinant protein of interest and an antibiotic resistance marker.

Day 2: A 10 mL LB starter culture was prepared.

Day 3: 1 L autoclaved LB was inoculated with the 10 mL starter culture and appropriate antibiotics were added at working concentrations. The flask was then incubated at 37°C and 160 rpm, until the bacteria reached an OD_{600} of 0.6. Once this OD_{600} of 0.6 was reached, 400 μM IPTG was added and the culture was incubated at 19°C and 160 rpm, overnight.

2.4.3 Recombinant protein purification

Following the steps in 2.4.2, bacterial cells were collected by centrifugation at 4000 rpm for 20 minutes at 4°C. The bacterial pellet was then resuspended in 30 mL binding buffer. The cells were then transferred to a sonication vessel before

sonicating using a Sonics Vibracell Ultrasonic processor using the following settings:

- Timer – 6 minutes
- Pulse – 30 seconds on / 30 seconds off
- Amplitude – 50%

Following sonication, the cells were centrifuged at 18000 rpm for 20 minutes at 4°C to separate the cell debris (pellet) from the soluble fraction (supernatant). The protein of interest was then purified from the supernatant, using IMAC.

Preparation of nickel column:

An empty column was fitted with a filter before adding 5 mL Chelating Sepharose™ resin. This was allowed to set before charging with 15 mL charge buffer. The column was then equilibrated with 20 mL binding buffer.

Purification of protein:

Supernatant was loaded onto the column. The column was then washed with 20 mL binding buffer, followed by 10 mL wash buffer I and 5 mL wash buffer II. The protein was then eluted with 6 mL elution buffer in 1 mL fractions, collected in 6 Eppendorf tubes.

Column wash:

The resin was then regenerated by running through 15 mL of strip buffer and washing with 50 mL dH_2O.

2.4.4 Semi-quantitative BioRad assay

Using the BioRad assay, it can be seen which of the elutions contain the majority of the protein. Drops containing 10 μL H_2O and 5 μL Biorad reagent were prepared per protein sample. 5 μL of protein solution, from each of the elutions, was added to each of the Biorad assays. The more the Biorad reagent changes from red to blue, the more protein is present in the sample. The two samples that contained the most protein were combined.

2.4.5 Buffer exchange using PD10 column (GE Healthcare)

The column was equilibrated with 25 mL PD10 buffer, before loading on 2.5 mL of purified protein. The flow-through was discarded. 3.5 mL PD10 buffer was then added and the flow-through (containing protein) was collected.

2.4.6 A_{280} protein concentration estimation using the extinction coefficient

The protein concentration can be estimated using the Beer-Lambert law:

Concentration (M) = A_{280} / protein's extinction coefficient (ε value)

ε values used:

CobU – 12490 $Lmol^{-1}cm^{-1}$

CobP – 37470 $Lmol^{-1}cm^{-1}$

Concentration (mg/ml) = concentration (M) x estimated protein mass (Daltons)

2.4.7 2X Laemmli SDS buffer

Table 2.9. 2X Laemmli SDS buffer components.

Reagent	Stock Concentration	Final Concentration	Volume required to make up 10 mL
Tris-HCl pH 6.8	0.5 M	0.125 M	2.5 mL
Glycerol	100%	20%	2 mL
SDS	10%	4%	4 mL
β-mercaptoethanol	100%	10%	1 mL
Bromophenol blue	0.08%	0.004%	0.5 mL

2.4.8 1X Running buffer

288 g glycine and 60 g Tris were made up to 2 L with dH_2O, to prepare 10X running buffer. For the preparation of 1X running buffer, 100 mL 10X running buffer was combined with 10 mL 10% SDS and made up to 1 L with dH_2O.

2.4.9 Coomassie blue stain

250 mL 100% trichloroacetic acid (TCA), 0.6 g Coomassie blue R, 0.1 g SDS, 0.25 g trizma base, and 0.1 g glycine were combined and made up to 500 mL with dH_2O.

2.4.10 Analysis of recombinant protein by SDS-PAGE (polyacrylamide gel electrophoresis)

Protein samples were collected, mixed with equal volumes of 2X Laemmli SDS buffer and boiled for 10 minutes. Polyacrylamide gels were set up as outlined in Table 2.10.

Table 2.10. SDS gel recipe.

Resolving Gel:	12.5%	15%	Stacking Gel:	5%
dH_2O (mL)	3.4	2.2	dH_2O (mL)	3.4
30% acrylamide (mL)	6.3	7.5	30% acrylamide (mL)	1.5
1.5M Tris-HCl pH 8.8 (mL)	3.8	3.8	0.5M Tris-HCl pH 6.8 (mL)	1.9
10% SDS (mL)	1.5	1.5	10% SDS (mL)	0.75
10% APS (mL)	0.15	0.15	10% APS (mL)	0.075
TEMED (mL)	0.01	0.01	TEMED (mL)	0.01

Gels were run in running buffer at 200 V for 1.5 hours and stained with Coomassie blue stain. Gels were de-stained overnight in dH_2O, under gentle agitation.

2.4.11 Transfer buffer

100 mL 10X running buffer was combined with 200 mL 100% methanol and made up to 1 L with dH_2O.

2.4.12 Blocking solution

5% (w/v) non-fat dried milk in PBS.

2.4.13 Phosphate-free blocking solution

5% (w/v) nonfat dried milk, 150 mM NaCl, 50 mM Tris-HCl pH 7.5, made up with 30 mL dH_2O.

2.4.14 Western blot analysis

Samples were initially run on SDS-PAGE (as described in 2.5.10). However, instead of staining on completion, gels were rinsed in transfer buffer for 5 minutes. Western blot apparatus was set up and run for 1 hour at 100 V.

Once protein was transferred onto the nitrocellulose membrane, this was then covered in blocking solution and incubated for 2 hours at room temperature, on a platform shaker.

Blocking solution was then discarded and the membrane was immediately incubated with the appropriate primary antibody (1:1000 dilution in PBS), for 2 hours at room temperature, on a platform shaker.

The membrane was then washed in PBS for 10 minutes, 3 separate times. Following this, the membrane was equilibrated in a buffer containing 150 mM NaCl and 50 mM Tris-HCl, before incubating for 1 hour in enzyme-coupled secondary antibody (1:5000 dilution in phosphate-free blocking solution).

The membrane was then washed in phosphate-free blocking solution for 10 minutes, 3 separate times. Following this, incubation in substrate 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium (BCIP/NBT) allowed visualisation of bands.

2.4.15 Antibodies

Mouse monoclonal anti-polyHistidine antibody (Sigma Aldrich) 1:1000

Mouse anti-GFP antibody (Sigma Aldrich) 1:1000

Anti-mouse IgG (H+L), AP Conjugate (Promega) 1:5000

Rabbit anti-TetR antibody (Sigma Aldrich) 1:1000

Goat anti-rabbit IgG (H+L) secondary antibody, AP, Invitrogen 1:5000

2.4.16 Digestion buffer

10 mM NH_4HCO_3, 10% acetonitrile in dH_2O.

2.4.17 Gel filtration chromatography

Gel filtration chromatography was carried out using an Äkta FPLC system. Samples were resolved on a SuperdexTM 200 10/300 GL column (GE Healthcare) which was equilibrated in 20 mM Tris HCl, pH8, 100 mM NaCl. A 1 mL protein sample was loaded onto the column and run at a flow rate of 0.5 mL/min. Calibration of the column was achieved using the MWGF-1000 gel filtration molecular weight marker kit from Sigma (Table 2.11).

Table 2.11. Sigma molecular weight marker kit components.

Protein	MW (kDa)	Concentration (mg/mL)
Carbonic anhydrase, bovine erythrocytes	29	3
Albumin, bovine serum	66	10
Alcohol dehydrogenase, yeast	150	5
β-amylase, sweet potato	200	4
Apoferritin, horse spleen	443	10
Thyroglobulin, bovine	669	8
Blue dextran	2000	2

2.4.18 Reverse phase chromatography (on an RP18 column)

LiChroprep RP18 resin from Merck was activated in hexane before washing with ethanol 3 times. Activated resin was then poured into an appropriate column, as required, before washing with acetonitrile and equilibrating with 0.1% TFA. Prior to application of the sample, samples must be acidified using 0.5 M HCl to a pH below 4. The sample is then applied and the column washed with 0.1% TFA. Bound molecules can then be eluted with 50% ethanol and dried using a vacuum centrifuge.

2.4.19 High performance liquid chromatography (HPLC)

Samples were run through a HPLC (Agilent 1100 system) using the following method:

- Column: Ace 5 AQ (2.1 x 250 mm; 5 µM; 100 Å)
- Column temperature: 30°C
- Flow rate: 0.2 ml min^{-1}
- Solvent A: 0.1% TFA
- Solvent B: Acetonitrile

Table 2.12. HPLC gradient method.

Time (minutes)	Solvent B (%)
0	0
5	0
45	100
50	100
60	0
70	0

2.5.20 Electrospray ionization (ESI) - mass spectrometry (MS)

Samples were run through the HPLC column, before analysing by ESI-MS using a Bruker MicroTOF-QII machine. ESI-MS involves the application of a high voltage to samples, in order to produce an electrospray of ions which can be used to obtain a mass spectrum.

2.4.21 Preparation of crystallisation screens

Protein crystallisation was attempted using the hanging drop method. 24-well crystallisation trays were greased for the screening process. Initial screens for each protein involved the use of conditions from Molecular Dimensions Structure Screens 1 and 2. 500 µL of each condition were pipetted into each well. On top of a siliconized glass coverslip, 1 µL of each condition was pipetted down followed by 1 µL of protein sample directly on the condition drop. The reservoir was then sealed

with the coverslip corresponding to that condition, such that the drops are facing down. Once the required number of wells was complete, the trays were incubated at 19°C and drops were observed periodically under a light microscope for the presence of crystals, nucleation or precipitation. A 1:3 ratio was used for protein:substrate mixes.

2.4.22 Optimisation screens

Table 2.13. Optimisation screen of condition 8 (SS2). Component optimised: MPD (2-methyl-2,4-pentanediol).

Tray 1	**1**	**2**	**3**	**4**	**5**
A	pH 7.0 35% v/v MPD	pH 7.5 35% v/v MPD	pH 8.0 35% v/v MPD	pH 8.5 35% v/v MPD	pH 9.0 35% v/v MPD
B	pH 7.0 40% v/v MPD	pH 7.5 40% v/v MPD	pH 8.0 40% v/v MPD	pH 8.5 40% v/v MPD	pH 9.0 40% v/v MPD
C	pH 7.0 45% v/v MPD	pH 7.5 45% v/v MPD	pH 8.0 45% v/v MPD	pH 8.5 45% v/v MPD	pH 9.0 45% v/v MPD
D	pH 7.0 50% v/v MPD	pH 7.5 50% v/v MPD	pH 8.0 50% v/v MPD	pH 8.5 50% v/v MPD	pH 9.0 50% v/v MPD
Tray 2	**1**	**2**	**3**	**4**	**5**
A	pH 7.0 50% v/v MPD	pH 7.5 50% v/v MPD	pH 8.0 50% v/v MPD	pH 8.5 50% v/v MPD	pH 9.0 50% v/v MPD
B	pH 7.0 55% v/v MPD	pH 7.5 55% v/v MPD	pH 8.0 55% v/v MPD	pH 8.5 55% v/v MPD	pH 9.0 55% v/v MPD
C	pH 7.0 60% v/v MPD	pH 7.5 60% v/v MPD	pH 8.0 60% v/v MPD	pH 8.5 60% v/v MPD	pH 9.0 60% v/v MPD
D	pH 7.0 65% v/v MPD	pH 7.5 65% v/v MPD	pH 8.0 65% v/v MPD	pH 8.5 65% v/v MPD	pH 9.0 65% v/v MPD

Table 2.14. Optimisation screen of condition 25 (SS2).

Tray 1	1	2	3	4	5	6
A	pH 5.5 1.2 M $(NH_4)_2SO_4$	pH 5.5 1.4 M $(NH_4)_2SO_4$	pH 5.5 1.6 M $(NH_4)_2SO_4$	pH 5.5 1.8 M $(NH_4)_2SO_4$	pH 5.5 2.0 M $(NH_4)_2SO_4$	pH 5.5 2.2 M $(NH_4)_2SO_4$
B	pH 5.7 1.2 M $(NH_4)_2SO_4$	pH 5.7 1.4 M $(NH_4)_2SO_4$	pH 5.7 1.6 M $(NH_4)_2SO_4$	pH 5.7 1.8 M $(NH_4)_2SO_4$	pH 5.7 2.0 M $(NH_4)_2SO_4$	pH 5.7 2.2 M $(NH_4)_2SO_4$
C	pH 6.0 1.2 M $(NH_4)_2SO_4$	pH 6.0 1.4 M $(NH_4)_2SO_4$	pH 6.0 1.6 M $(NH_4)_2SO_4$	pH 6.0 1.8 M $(NH_4)_2SO_4$	pH 6.0 2.0 M $(NH_4)_2SO_4$	pH 6.0 2.2 M $(NH_4)_2SO_4$
D	pH 6.5 1.2 M $(NH_4)_2SO_4$	pH 6.5 1.4 M $(NH_4)_2SO_4$	pH 6.5 1.6 M $(NH_4)_2SO_4$	pH 6.5 1.8 M $(NH_4)_2SO_4$	pH 6.5 2.0 M $(NH_4)_2SO_4$	pH 6.5 2.2 M $(NH_4)_2SO_4$

Table 2.15. Finer optimisation screen of condition 8 (SS2). Component optimised: MPD (2-methyl-2,4-pentanediol).

Tray 1	1	2	3	4	5	6
A	pH 7.1 35% v/v MPD	pH 7.1 38% v/v MPD	pH 7.1 40% v/v MPD	pH 7.1 42% v/v MPD	pH 7.1 45% v/v MPD	pH 7.1 50% v/v MPD
B	pH 7.3 35% v/v MPD	pH 7.3 38% v/v MPD	pH 7.3 40% v/v MPD	pH 7.3 42% v/v MPD	pH 7.3 45% v/v MPD	pH 7.3 50% v/v MPD
C	pH 7.5 35% v/v MPD	pH 7.5 38% v/v MPD	pH 7.5 40% v/v MPD	pH 7.5 42% v/v MPD	pH 7.5 45% v/v MPD	pH 7.5 50% v/v MPD
D	pH 7.7 35% v/v MPD	pH 7.7 38% v/v MPD	pH 7.7 40% v/v MPD	pH 7.7 42% v/v MPD	pH 7.7 45% v/v MPD	pH 7.7 50% v/v MPD
Tray 2	**1**	**2**	**3**	**4**	**5**	**6**
A	pH 7.9 35% v/v MPD	pH 7.9 38% v/v MPD	pH 7.9 40% v/v MPD	pH 7.9 42% v/v MPD	pH 7.9 45% v/v MPD	pH 7.9 50% v/v MPD
B	pH 8.1 35% v/v MPD	pH 8.1 38% v/v MPD	pH 8.1 40% v/v MPD	pH 8.1 42% v/v MPD	pH 8.1 45% v/v MPD	pH 8.1 50% v/v MPD
C	pH 8.3 35% v/v MPD	pH 8.3 38% v/v MPD	pH 8.3 40% v/v MPD	pH 8.3 42% v/v MPD	pH 8.3 45% v/v MPD	pH 8.3 50% v/v MPD
D	pH 8.5 35% v/v MPD	pH 8.5 38% v/v MPD	pH 8.5 40% v/v MPD	pH 8.5 42% v/v MPD	pH 8.5 45% v/v MPD	pH 8.5 50% v/v MPD

Table 2.16. Optimisation screen of condition 14 (SS1). Components optimised: Na/Caco (sodium cacodylate) and PEG (polyethylene glycol).

Tray 1	1	2	3	4	5	6
A	0.1M Na/Caco 0.06M $(NH_4)_2SO_4$ 15% PEG	0.1M Na/Caco 0.06M $(NH_4)_2SO_4$ 20% PEG	0.1M Na/Caco 0.06M $(NH_4)_2SO_4$ 25% PEG	0.1M Na/Caco 0.06M $(NH_4)_2SO_4$ 30% PEG	0.1M Na/Caco 0.06M $(NH_4)_2SO_4$ 35% PEG	0.1M Na/Caco 0.06M $(NH_4)_2SO_4$ 40% PEG
B	0.1M Na/Caco 0.09M $(NH_4)_2SO_4$ 15% PEG	0.1M Na/Caco 0.09M $(NH_4)_2SO_4$ 20% PEG	0.1M Na/Caco 0.09M $(NH_4)_2SO_4$ 25% PEG	0.1M Na/Caco 0.09M $(NH_4)_2SO_4$ 30% PEG	10.1M 0.1M Na/Caco 0.09M $(NH_4)_2SO_4$ 35% PEG	0.1M Na/Caco 0.09M $(NH_4)_2SO_4$ 40% PEG
C	0.1M Na/Caco 0.12M $(NH_4)_2SO_4$ 15% PEG	0.1M Na/Caco 0.12M $(NH_4)_2SO_4$ 20% PEG	0.1M Na/Caco 0.12M $(NH_4)_2SO_4$ 25% PEG	0.1M Na/Caco 0.12M $(NH_4)_2SO_4$ 30% PEG	0.1M Na/Caco 0.12M $(NH_4)_2SO_4$ 35% PEG	0.1M Na/Caco 0.12M $(NH_4)_2SO_4$ 40% PEG
D	0.1M Na/Caco 0.15M $(NH_4)_2SO_4$ 15% PEG	0.1M Na/Caco 0.15M $(NH_4)_2SO_4$ 20% PEG	0.1M Na/Caco 0.15M $(NH_4)_2SO_4$ 25% PEG	0.1M Na/Caco 0.15M $(NH_4)_2SO_4$ 30% PEG	0.1M Na/Caco 0.15M $(NH_4)_2SO_4$ 35% PEG	0.1M Na/Caco 0.15M $(NH_4)_2SO_4$ 40% PEG

Table 2.17. Optimisation screen of condition 15 (SS1). Components optimised: Na/Caco (sodium cacodylate), Mg/Ace (magnesium acetate) and PEG (polyethylene glycol).

Tray 1	1	2	3	4	5	6
A	0.1M Na/Caco 0.1M Mg/Ace 10% PEG	0.1M Na/Caco 0.1M Mg/Ace 15% PEG	0.1M Na/Caco 0.1M Mg/Ace 20% PEG	0.1M Na/Caco 0.1M Mg/Ace 25% PEG	0.1M Na/Caco 0.1M Mg/Ace 30% PEG	0.1M Na/Caco 0.1M Mg/Ace 35% PEG
B	0.1M Na/Caco 0.15M Mg/Ace 10% PEG	0.1M Na/Caco 0.15M Mg/Ace 15% PEG	0.1M Na/Caco 0.15M Mg/Ace 20% PEG	0.1M Na/Caco 0.15M Mg/Ace 25% PEG	0.1M Na/Caco 0.15M Mg/Ace 30% PEG	0.1M Na/Caco 0.15M Mg/Ace 35% PEG
C	0.1M Na/Caco 0.2M Mg/Ace 10% PEG	0.1M Na/Caco 0.2M Mg/Ace 15% PEG	0.1M Na/Caco 0.2M Mg/Ace 20% PEG	0.1M Na/Caco 0.2M Mg/Ace 25% PEG	0.1M Na/Caco 0.2M Mg/Ace 30% PEG	0.1M Na/Caco 0.2M Mg/Ace 35% PEG
D	0.1M Na/Caco 0.25M Mg/Ace 10% PEG	0.1M Na/Caco 0.25M Mg/Ace 15% PEG	0.1M Na/Caco 0.25M Mg/Ace 20% PEG	0.1M Na/Caco 0.25M Mg/Ace 25% PEG	0.1M Na/Caco 0.25M Mg/Ace 30% PEG	0.1M Na/Caco 0.25M Mg/Ace 35% PEG

Table 2.18. Optimisation screen of condition 30 (SS1).

Tray 1	1	2	3	4
A	0.5M $(NH_4)_2SO_4$ 0.1M HEPES 1% PEG	1M $(NH_4)_2SO_4$ 0.1M HEPES 1% PEG	1.5M $(NH_4)_2SO_4$ 0.1M HEPES 1% PEG	2M $(NH_4)_2SO_4$ 0.1M HEPES 1% PEG
B	0.5M $(NH_4)_2SO_4$ 0.1M HEPES 2% PEG	1M $(NH_4)_2SO_4$ 0.1M HEPES 2% PEG	1.5M $(NH_4)_2SO_4$ 0.1M HEPES 2% PEG	2M $(NH_4)_2SO_4$ 0.1M HEPES 2% PEG
C	0.5M $(NH_4)_2SO_4$ 0.1M HEPES 3% PEG	1M $(NH_4)_2SO_4$ 0.1M HEPES 3% PEG	1.5M $(NH_4)_2SO_4$ 0.1M HEPES 3% PEG	2M $(NH_4)_2SO_4$ 0.1M HEPES 3% PEG
D	0.5M $(NH_4)_2SO_4$ 0.1M HEPES 4% PEG	1M $(NH_4)_2SO_4$ 0.1M HEPES 4% PEG	1.5M $(NH_4)_2SO_4$ 0.1M HEPES 4% PEG	2M $(NH_4)_2SO_4$ 0.1M HEPES 4% PEG

2.4.23 X-Ray diffraction experiment

Crystals were picked on a loop and dipped into the cryo-protectant consisting of the corresponding well solution and 25% glycerol (w/v), before vitrification by transferring into a 100 K nitrogen stream. This procedure was done by Marie Anderson. Vitrified crystals were then taken to Synchroton radiation source (Diamond Light Source, Oxfordshire, UK) where datasets were collected.

2.4.24 eGFP fluorescence analysis

Fluorescence intensities from *in vitro* reaction samples were measured, using a Perkin Elmer Luminescence Spectrometer LS50B. Excitation and emission wavelengths of 488 nm and 510 nm were used, respectively, in order to obtain fluorescence intensity values from each sample. These values were then used to work out eGFP concentrations, relative to a standard curve prepared from the intensity values of eGFP standards of known concentration.

Additional settings:

- Photomultiplier tube (PMT) voltage – 900 V
- Slit width – 10 mm round
- Averaging time – 0.5 s

2.4.25 CobP *in vitro* assays

Purified CobP protein (variable concentration) was combined with 1 M $MgCl_2$, 14 mM GTP or 20 mM ATP, and a cobinamide of choice (variable concentration) before making up to a final volume of 100 µL with pH 8 PD10 buffer and incubating in the dark overnight.

On the following day, each assay was heated at 60ºC for 10 minutes to denature all protein, before centrifuging at 13000 rpm for 5 minutes. 100 µL samples were removed for analysis using HPLC-MS.

Table 2.19. CobP assays 1st attempt

Sample:	CobP (µM)	$MgCl_2$ (mM)	GTP (mM)	ATP (µM)	AdoCbi (µM)	OHCbi (µM)
1	-	5	2	-	41.4	-
2	-	5	2	-	-	50
3	-	5	-	2	41.4	-
4	-	5	-	2	-	50
5	0.6	5	-	-	41.4	-
6	0.6	5	-	-	-	50
7	0.6	5	2	-	41.4	-
8	0.6	5	2	-	-	50
9	0.6	5	-	2	41.4	-
10	0.6	5	-	2	-	50

Table 2.20. CobP assays 2nd attempt

Sample:	CobP (µM)	$MgCl_2$ (mM)	GTP (mM)	AdoCbi (µM)	OHCbi (µM)	Buffer pH
1	6	5	2	20	-	8
2	6	5	2	-	20	8

3	6	5	2	20	-	9
4	6	5	2	-	20	9

Table 2.21. CobP assays 3rd attempt

Sample:	Species (Rc/Av)	Purified/ Lysate	CobP (μM)	$MgCl_2$ (mM)	GTP (mM)	AdoCbi (μM)	DTT (mM)
1	Av	P	60	5	2	10	-
4	Av	L	60	5	2	10	-
7	Rc	P	6	5	2	10	-
10	Rc	L	6	5	2	10	-
13	Av	P	60	5	2	10	10
16	Rc	P	6	5	2	10	10

Table 2.22. CobP assays 4th attempt

Sample:	Species (Rc/Av)	Purified/ Lysate	CobP (μM)	$MgCl_2$ (mM)	GTP (mM)	AdoCbi (μM)	*E. coli* Lysate
1	Av	P	600	5	2	10	-
2	Av	L	n/a	5	2	10	-
Controls:							
3	-	-	-	5	2	10	Yes
4	Av	P	600	5	2	10	Yes
5	-	-	-	5	2	10	-

Chapter 3: Biosynthesis of cobalt-free B_{12} (hydrogenobalamin)

3.1 Introduction

As a consequence of over half a century of B_{12} research it is now possible to persuade *E. coli* to make B_{12} through metabolic engineering, a compound it cannot make on its own. It has also been discovered that certain organisms, such as *A. vinosum*, have the ability to produce cobalt-free corrinoids when grown in low cobalt conditions (Toohey, 1965). One of the aims of this project was to see whether it would be possible to combine these two key factors, the ability to recombinantly produce complex molecules in *E. coli* with the ability of *A. vinosum* enzymes to produce cobalt-free corrinoids. The goal therefore was to be able to produce cobalt-free corrinoids, including cobalt-free cobalamin (hydrogenobalamin), in *E. coli*. This can then be used as the starting point for creating unnatural B_{12} variants that act as B_{12} antagonists or inhibitors, especially if they have different transition metals inserted into the empty corrin framework. These metal variants of cobalamin (metbalamins), have a very similar structure to cobalamin and will likely interfere with B_{12}-dependent processes within the cell, thereby acting as B_{12} inhibitors.

To achieve such a goal, the construction of a plasmid containing a gene set encoding enzymes that together have the capacity to make hydrogenobalamin was attempted. In order to do this certain B_{12} biosynthesis genes, *cobQ, cobC*, *cobD*, *cobP*, *cobV* and *cobU*, were amplified from the *A. vinosum* genome. These particular genes were chosen for a number of reasons. *cob*Q was chosen to be amplified from the *A. vinosum* genome because the enzyme, when taken from standard aerobic organisms, is known to recognise ado-cobyrinic acid *a,c*-diamide as its substrate (Crouzet et al, 1991), whereas *A. vinosum* CobQ is able to use HBAD as its substrate (Lawrence et al, 2018). The rest of the chosen genes all have gene products which are involved in the synthesis and assembly of the lower

loop, following the cobalt chelation stage, and since *A. vinosum* is one of the few bacteria capable of producing metal-free corrinoids (Toohey, 1965), as discussed in Chapter 1, it is thought that the enzymes from this bacterium should be able to successfully work with metal-free substrates.

Following the amidation of ado-cobyrinic acid *a,c*-diamide by CobQ, CobC and CobD work together to carry out the synthesis and attachment of the aminopropanol linker. Next, the bifunctional enzyme, CobP, initiates the process of assembling the lower loop by converting AdoCbi into AdoCbi-GDP. The final stage involves lower ligand synthesis and assembly, and requires two enzymes, CobU and CobV, to form AdoCbl from AdoCbi-GDP. CobU carries out the synthesis of the lower loop (α-ribazole phosphate), and CobV assembles this onto AdoCbi-GDP to form AdoCbl (Warren et al, 2002). These steps are outlined in greater detail in Chapter 1 (Figure 1.4).

The work described in this chapter was focussed on attempting to generate new metal-free intermediates through the cloning and recombinant expression of the final six enzymes from *A. vinsoum*. The biosynthetic steps are outlined in Figure 3.1.

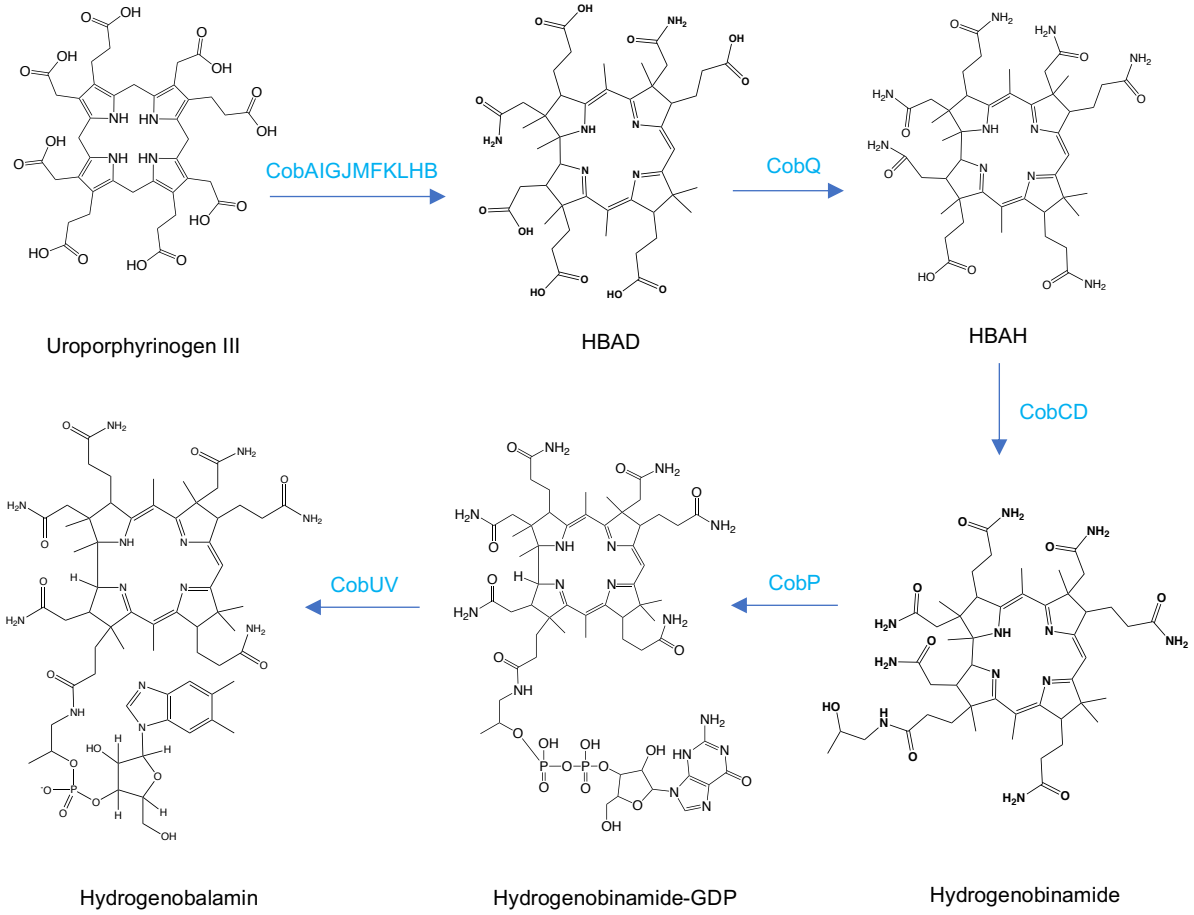


Figure 3.1. The conversion of uroporphyrinogen III to hydrogenobalamin.

3.2 Results

3.2.1 Amplification of B_{12} biosynthesis genes, involved in the synthesis and attachment of the lower loop of B_{12}

The *A. vinosum cobQ*, *cobC* and *cobD* had been previously amplified from the *A. vinosum* genome (by Dr. Emi Nemoto-Smith and Dr. Andrew Lawrence), meaning that *cobP*, *cobV* and *cobU* still had to be cloned. The final three genes were amplified by PCR, as described in Chapter 2, with varying concentrations of DMSO. During amplification, an NdeI site was introduced at the 5' end of each gene and an SpeI site at the 3' end, via use of the appropriate forward and reverse primers (Table 2.3). Introduction of these restriction sites allowed them to be cloned into individual pET3a vectors following PCR, and for them to subsequently be cloned consecutively within a single plasmid using the 'Link and Lock' multiple cloning strategy, a method developed by Dr. Evelyne Deery and Dr. Helen McGoldrick (McGoldrick et al, 2005).

Agarose gel electrophoresis was used to separate the amplified DNA bands from each of the individual PCR samples. The PCR bands from the 2 μL DMSO samples (~562 bp *cobP*, ~754 bp *cobV* and ~1057 bp *cobU*), (Figure 3.2), were each cut out and gel extraction procedures were performed (as described in Chapter 2, Section 2.3.6).

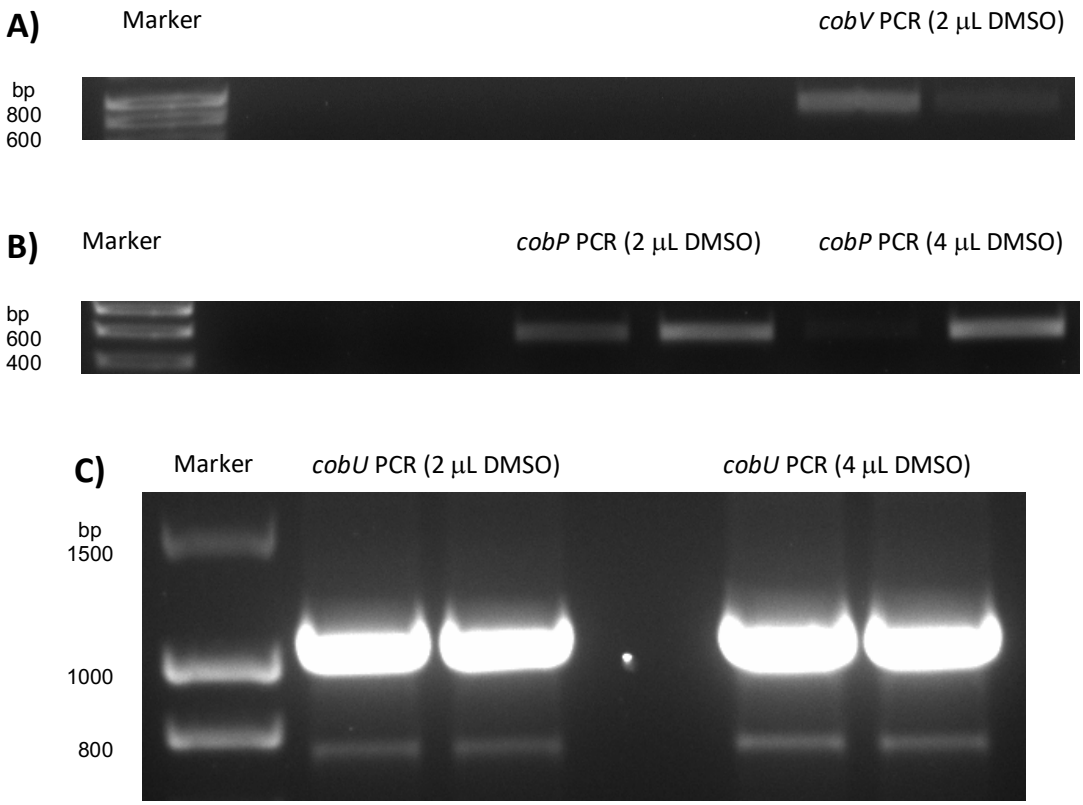


Figure 3.2. Agarose gel electrophoresis showing PCR products from the amplification of *A. vinosum* genes. PCR samples for; A.) *cobV* B.) *cobP* C.) *cobU.* Successful PCR bands are seen in the 2 μL DMSO samples (~562 bp *cobP*, ~754 bp *cobV* and ~1057 bp *cobU*), and certain 4 μL DMSO samples.

The three DNA fragments were digested with NdeI and SpeI restriction enzymes, to allow for ligation with NdeI and SpeI cut pET3a. Transformation of *E. coli* JM109 was carried out using the new ligation product. Plasmid was extracted from successful transformants (as described in Chapter 2, Section 2.3.6) and was sent for sequencing to ensure that the relevant genes had been successfully amplified and cloned into pET3a.

Five of the six required *A. vinosum* genes were cloned into the same plasmid to generate a *cobCDPVU-pET3a* construct using the 'Link and Lock' multiple cloning

strategy, (McGoldrick et al, 2005). Initially, *cobC* and *cobD* were cloned together via 'Link and Lock' into pET3a to give *cobCD-pET3a*. Similarly, *cobP* and *cobV* were also cloned together to give *cobPV-pET3a.* Following this, the *cobCD* genes were removed and cloned next to the *cobPV* genes in the *cobPV*-pET3a vector, to generate *cobCDPV-pET3a*. Finally, *cobCDPV*-pET3a was digested with SpeI and EcoRI to allow the cloning of *cobU* via 'Link and Lock' to give *cobCDPVU*-pET3a, the five gene plasmid containing the *A. vinosum* genes required for synthesis of the lower loop of B_{12}. The plasmids generated during this investigation were also used to allow a study of the individual encoded proteins.

3.2.2 Characterising the lower loop enzymes using recombinant protein production and purification methods

To determine whether the *A. vinosum cobD, cobP, cobV* and *cobU* genes allow for recombinant protein production, each of these four genes was cloned individually into pET14b using the restriction enzymes NdeI and SpeI. This generated four new plasmids; *cobD*-pET14b, *cobP*-pET14b, *cobV*-pET14b and *cobU*-pET14b. Transformation of *E. coli* BL21*(DE3)-pLysS competent cells was carried out using each of the individual plasmids.

Recombinant protein production and protein purification, using IMAC, were undertaken (as described in Chapter 2, Section 2.4.2 and 2.4.3) in order to demonstrate whether the proteins, encoded by each of the newly amplified *A. vinosum* genes, had been overproduced and whether they were soluble. Previous work, by Dr. Andrew Lawrence, showed that *A. vinosum* CobC expresses well and is found in the soluble fractions.

The first two proteins that were investigated, CobP and CobU, have expected masses (including the His tag) of 22.5 kDa and 39.1 kDa, respectively. Fractions from the purifications of CobP and CobU were analysed using SDS-PAGE, with results suggesting both proteins are soluble (Figure 3.3). Bands of the correct size are seen for CobP in not only the elution fractions, but also in the total, supernatant and flow-through fractions (Figure 3.3).

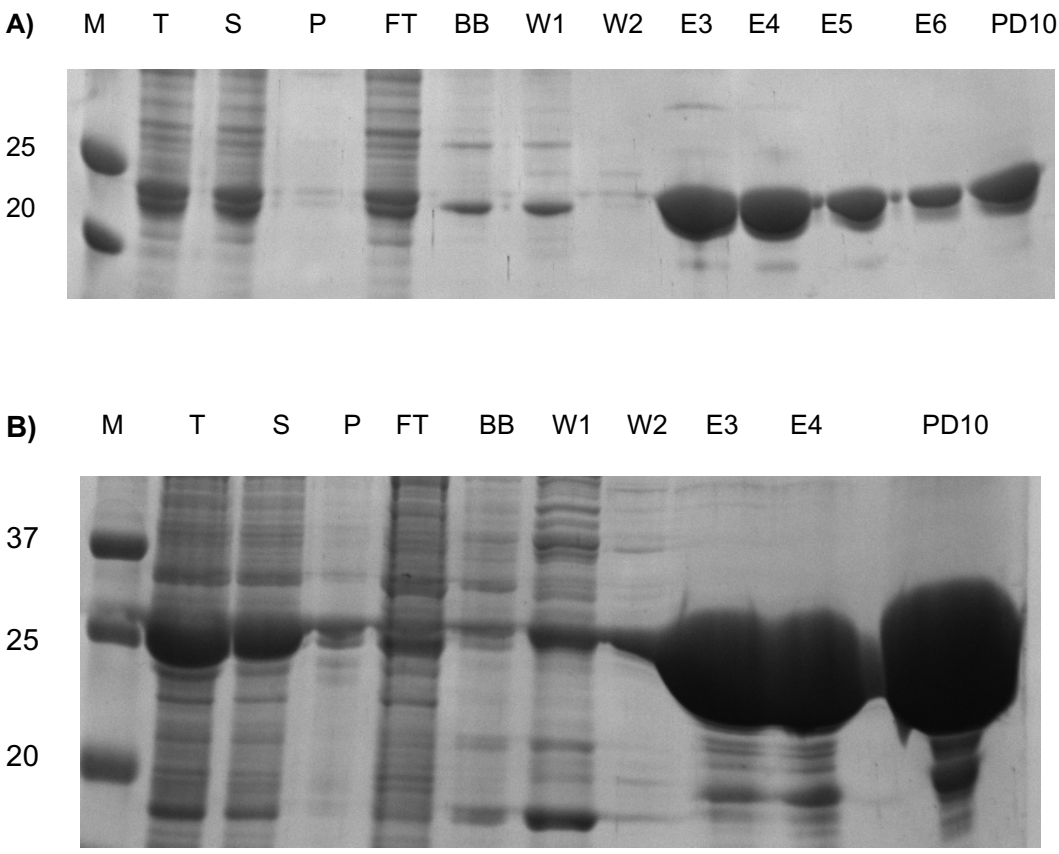


Figure 3.3. SDS-PAGE analysis of A) CobP (22.5 kDa) and B) CobU (39.1 kDa) purifications. CobP and CobU are visible in the elutions. (M – marker, T- total, S – supernatant, P – pellet, FT – supernatant flow-through, BB – binding buffer, W1 – wash buffer 1, W2 – wash buffer 2, E3 – elution 3, E4 – elution 4, PD10 – PD10

CobV and CobD purification fractions were analysed in the same way, using SDS-PAGE. Including the His-tag, CobV and CobD are predicted to have masses of 28.8 kDa and 37.2 kDa, respectively. However, following recombinant protein production and purification from *E. coli* BL21*(DE3)-pLysS, proteins of this size were not detected on an SDS gel in either the soluble or insoluble fractions (Figure 3.4).

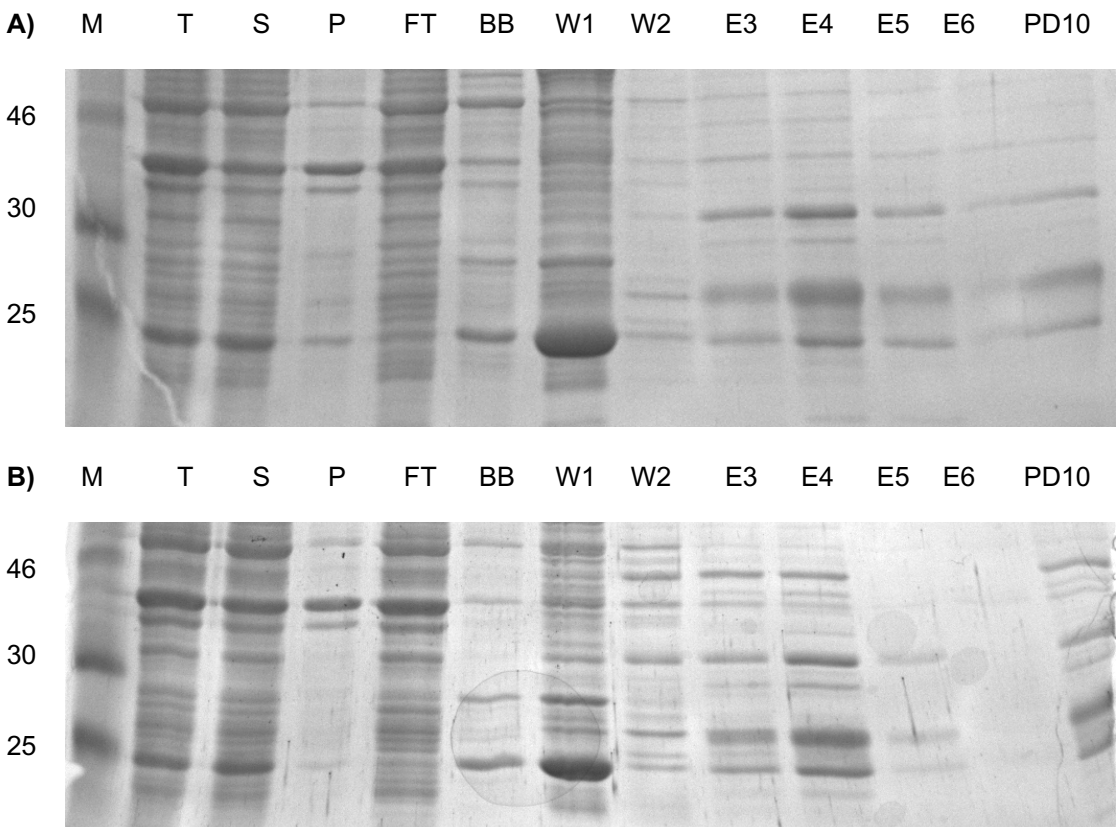


Figure 3.4. SDS-PAGE analysis of fractions from the purification of CobV and CobD from *E. coli* BL21*DE3-pLysS. A) CobV protein (28.8 kDa with the His tag) purification, B) CobD protein (37.2 kDa with the His tag) purification.

In another attempt to detect whether CobV was being produced, the protein purification was re-attempted in the presence and absence of 50 mg/mL CNCbl. This was done so that any CobV present within the bacterial cells could bind to the CNCbl. The strongly coloured nature of CNCbl should then allow easier detection of CobV. Unfortunately, CobV was still not detected, with no difference between fractions from both sets of purifications (with and without CNCbl), suggesting that it was not being overproduced.

Since both CobD and CobV could not be detected, the *cobV*-pET14b and *cobD*-pET14b plasmids were individually transformed into Rosetta competent cells (an *E.*

coli BL21(DE3) derivative designed to enhance the expression of proteins that contain codons rarely used in *E. coli)*, supplied by Novagen. This meant that if the *A. vinosum* cobV or *cobD* genes contain rare codons, for which there are no appropriate anti-codons in *E. coli* BL21*(DE3)-pLysS, then this would no longer be a problem since Rosetta contains anti-codons to the rare codons. Bacterial cells were cultured, harvested and the protein purification process was repeated. Fractions were analysed using SDS-PAGE (Figure 3.5).

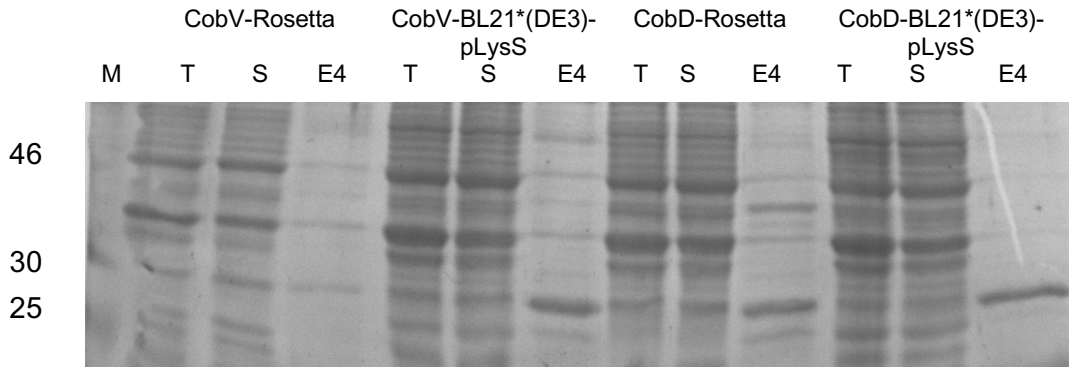


Figure 3.5. SDS-PAGE comparing purification fractions from Rosetta and BL21*(DE3)-pLysS strain *E. coli* for CobV (28.8 kDa with the His tag) and CobD (37.2 kDa with the His tag). Neither protein is seen to be overproduced.

From the SDS-PAGE analysis, it appeared that the two proteins were still not being overproduced and therefore the presence of rare codons was not the problem. A western blot analysis was therefore undertaken to see if either protein, CobV or CobD, could be detected as a result of their His-tags. No bands were seen within 15 minutes of developing, although bands were detected in the CobV sample after overnight development. However, the bands were smaller than the expected size of 28.8 kDa (Figure 3.6). The smaller bands could be due to some form of degradation

of the protein or, alternatively, could represent some form of background reaction and not be CobV at all. No bands were detected in the CobD sample (Figure 3.6).

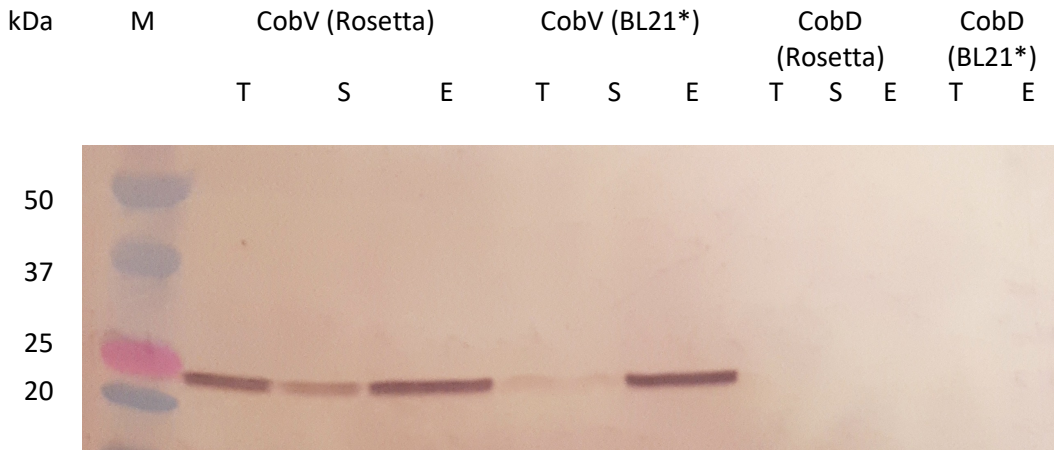


Figure 3.6. Western blot analyzing insoluble (T), soluble (S) and elution (E) fractions from CobV and CobD purifications (from both *E. coli* BL21*DE3-pLysS and Rosetta). Bands can be seen in the CobV samples, however, these are much smaller than the expected size of 28.8 kDa

Despite further attempts at recombinant protein production and purification of both proteins, the western blot results remained the same with a smaller than expected band for CobV and no bands for CobD. For CobD it is highly likely that the protein is not being made by the cells. However, the results remain inconclusive for CobV and it is not clear whether the observed band is actually CobV.

As numerous attempts of trying to detect CobD were unsuccessful, it was thought that perhaps it may be due to the particular location of the His-tag on the protein that was affecting its ability to bind to the nickel column. It could be that while the His-tag was at the N-terminus, it was being folded within the protein in such a way

that it was not easily accessible. Therefore, *cobD*-pET23b (provided by Dr. Andrew Lawrence) was used for the transformation of *E. coli* Rosetta and cultured. By placing the gene in a pET23b vector, the resultant gene product should have a C-terminal His-tag. In this case, CobD should now have a His-tag at the C-terminus of the protein which may perhaps be more accessible by the nickel column.

Protein purification was carried out and fractions analysed using SDS-PAGE. However, despite the change in the position of the His-tag, CobD was still not detected by SDS-PAGE. Therefore, samples from this purification were analysed on a western blot alongside samples from the purification of N-terminal His-tagged CobD from *E. coli* Rosetta. No bands were seen on the western blot, suggesting that the *A. vinosum* CobD protein is not being produced. This was repeated several times yet the results remained the same.

3.2.3 Generation of a plasmid for the production of cobalt-free B_{12} (hydrogenobalamin)

The cob*CDPVU* gene set was required to be cloned into a plasmid containing all of the remaining B_{12} biosynthesis genes needed for the production of hydrogenobalamin. These include the *cobAIG*JFMKLHBQ*$^{(Alv)}$ genes, found in the *cobAIG*JFMKLHBQ*$^{(Alv)}$-pET3a plasmid (see Table 2.2). Cloning of cob*CDPVU* into this plasmid was attempted through use of the 'Link and Lock' strategy.

Numerous cloning attempts were made to obtain the *cobAIG*JFMKLHBQ*$^{(Alv)}$[CDPVU]$^{(Alv)}$-pET3a plasmid. However, this was not successful. This was most likely due to the desired plasmid being too large, at over 20000 bp, causing reduced transformation efficiency. A new strategy was therefore attempted using a more stable pETcoco variant plasmid, pETcocoR (see Table 2.2). *cobAIG*JFMKLHBQ*$^{(Alv)}$ genes were cloned out of pET3a and into pETcocoR to give *cobAIG*JFMKLHBQ*$^{(Alv)}$-pETcocoR (Figure 3.7).

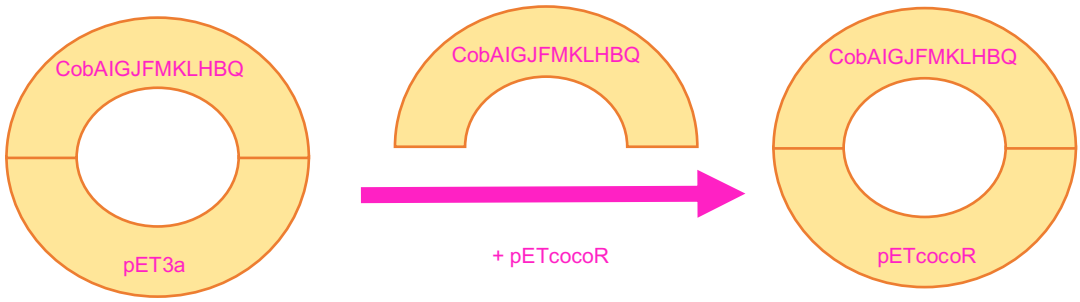


Figure 3.7. New cloning strategy. The *cobAIG*JFMKLHBQ*$^{(Alv)}$ genes were cloned out of pET3a and into pETcocoR to get *cobAIG*JFMKLHBQ*$^{(Alv)}$-

E. coli BL21*(DE3)-pLysS was transformed with the new *cobAIG*JFMKLHBQ*$^{(Alv)}$-pETcocoR plasmid and cultured in 2YT, in order to test whether the plasmid allows for the production of hydrogenobyric acid (HBAH). Cells were first boiled to break them apart and then centrifuged in order to separate the cell debris from the supernatant. Next, the supernatant was collected and desalted on an RP18 column (as described in Chapter 2, Section 2.4.18), a process based on reversed-phase chromatography, in order to obtain HBAH. Any HBAH (orange-coloured) present was eluted with 50% ethanol. The RP18 elution was vacuum centrifuged to remove ethanol from the sample. A sample of 20 μL volume was removed and run on a 20% acrylamide gel, against hydrogenobyrinic acid (HBA) and hydrogenobyrinic acid *a,c*-diamide (HBAD) standards. Although there appeared to be HBAH present in the RP18 elution sample, the molecule did not appear very pure, containing a mixture of amidated HBA forms including hydrogenobyrinic acid monoamide (HBAM) and HBAD (Figure 3.8).

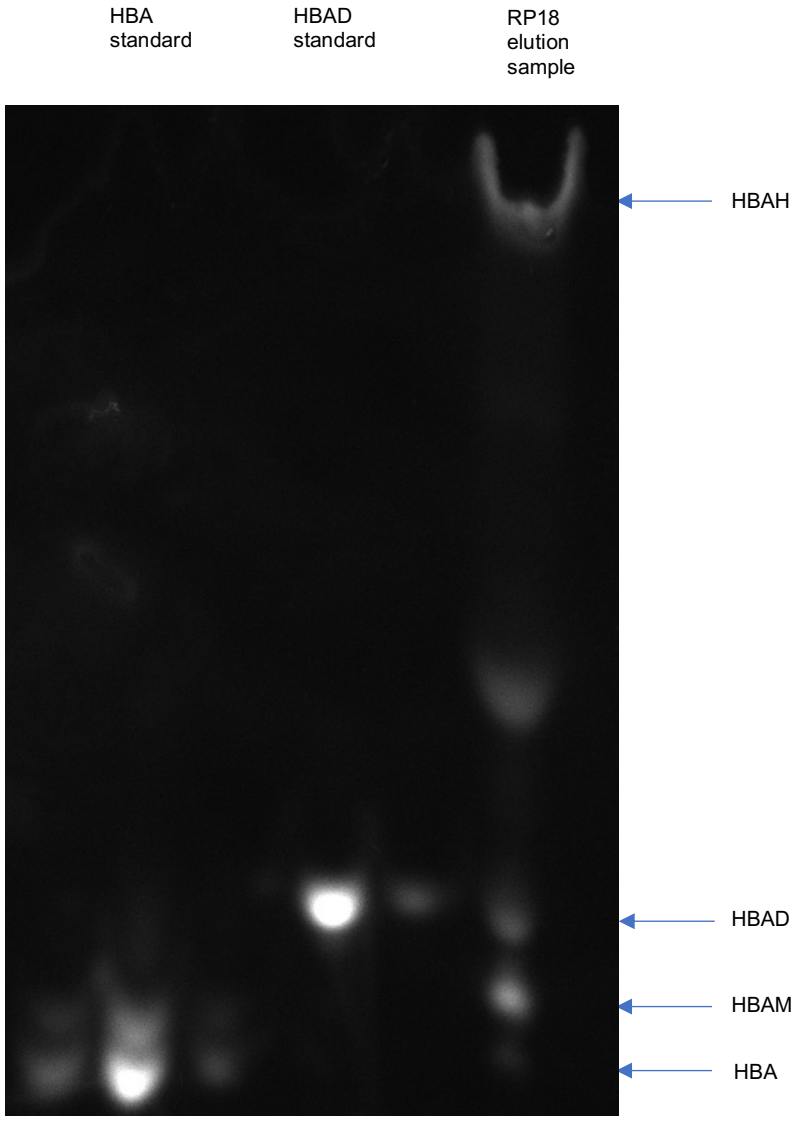


Figure 3.8. HBAH purification sample run on 20% acrylamide gel against HBA and HBAD standards, seen by the two brightest bands in the HBA and HBAD lanes, respectively. The HBAH-containing RP18 elution was run against HBA and HBAD standards. HBAH, which is uncharged and therefore does not travel down the gel, was found to be present in the RP18 elution sample. This suggests that the bacteria are able to successfully make HBAH. However, the mixture is not very pure, containing other amidated HBA forms including HBAM and HBAD.

Since it was confirmed that the *cobAIG*JFMKLHBQ*$^{(Alv)}$-pETcocoR plasmid allowed the production of HBAH, attempts were made to insert *cobCDPVU* into the new *cobAIG*JFMKLHBQ*$^{(Alv)}$-pETcocoR plasmid, such that the gene products of the *cobCDPVU* genes could allow for the production of hydrogenobalamin from HBAH. However, the *cobAIG*JFMKLHBQ*$^{(Alv)}$[CDPVU]$^{(Alv)}$-pETcocoR plasmid could not be constructed. This could either be due to a failure in the ligation of both DNA fragments or due to problems at the transformation stage, despite the use of various different competent bacterial cells (JM109, DH10β and NovaF^{-}). In either case, a lack of success was most likely due to the desired plasmid being too large, at over 28000 bp, causing reduced ligation and transformation efficiency.

3.2.4 Use of an alternative method to generate hydrogenobalamin

As it was not possible to construct a single plasmid that encoded all the enzymes for hydrogenobalamin production, an alternative method was investigated. This involved the use of a genetically engineered *E. coli* strain (*ED661)*, developed by Dr. Evelyne Deery, in which certain B_{12} biosynthesis genes have been integrated into the *E. coli* genome. *E. coli ED661* contains the *cobAIG*JFMKLH-B*^{Alv}*Q*^{Alv}*E** genes (with *cobG** and *cobE** from *Brucella melitensis*, and *cobB* and *cobQ* from *A. vinosum*) under the control of the T7 promoter. *E. coli ED661* also contains the T7 RNA Polymerase (T7RNAP) gene under the control of the lactose promoter (P_{Lac}). Together, the integrated genes are capable of producing HBAH, as is the *cobAIG*JFMKLHBQ*$^{(Alv)}$-pETcocoR plasmid that could not be combined with *cobCDPVU*-pET3a. It is for this reason that *E. coli ED661* was used for the preparation of ultra-competent cells and transformed with *cobCDPVU*-pET3a, with the expectation that it becomes capable of producing hydrogenobalamin. Additionally, *E. coli BL21*(DE3)-pLysS* was transformed with *btuF*-pET14b, a plasmid containing the *btuF* gene, in order to overproduce a His-tagged form of the B_{12}-binding protein, BtuF, which binds various different corrinoids (Lewinson et al, 2010). This His-tagged BtuF was then be used to purify corrinoids on a nickel column.

E. coli ED661-cobCDPVU-pET3a was cultured in 1 L 2YT medium, from an initial 10 mL LB medium starter culture, with added ammonium chloride (1 g/L). Meanwhile, *E. coli BL21*(DE3)-pLysS-btuF-pET14b* was cultured in 1 L LB medium, from an initial 10 mL LB medium starter culture.

Both cultures were harvested and cells were lysed. After cell lysis, both pellets were combined together in order that BtuF would bind to all of the corrinoids produced by *E. coli ED661-cobCDPVU*-pET3a. The sample was applied onto a charged nickel column to allow the His-tagged BtuF protein to bind to the column, thereby allowing all corrinoids to be separated from the remaining sample. Once all the sample was run through the column, BtuF was denatured using 8 M urea in order to allow any corrinoids to be released and eluted from the column. The elution obtained from the *E. coli ED661-cobCDPVU*-pET3a bacterial pellet was collected.

Similarly, the supernatant from the *E. coli ED661-cobCDPVU*-pET3a bacterial culture was also combined with a BtuF containing pellet and the same procedure carried out to obtain another corrinoid mix. Both elutions, from the pellet and supernatant, were desalted on an RP18 column. Samples were eluted in 50% ethanol and dried down using a vacuum centrifuge. The samples were resuspended in 100 μL 0.1% TFA before analysis using HPLC-MS (Chapter 2, Sections 2.4.19 and 2.4.20).

The HPLC-MS detected one major peak in the chromatogram of both the bacterial cell pellet elution sample and the supernatant sample. The mass spectrum of each of these peaks both show the same results (Figure 3.9), in which two distinct peaks can be seen with m/z values of 438 and 875, representing double charged and single charged HBAH, respectively. No peaks representing hydrogenobalamin (which has a mass of around 1270 Da) or any of the intermediates after HBAH (see Figure 3.1), could be detected in either sample, suggesting that the bacteria were unable to make any changes to HBAH. This indicates that there is a problem with the next step after HBAH, catalysed by CobC and CobD (see Figure 3.1). It is not clear whether these enzymes are inactive or whether the conditions were not

optimal for their activity. However, the real issue is likely to be a lack of CobD, since production of the protein could not be detected in previous experiments.

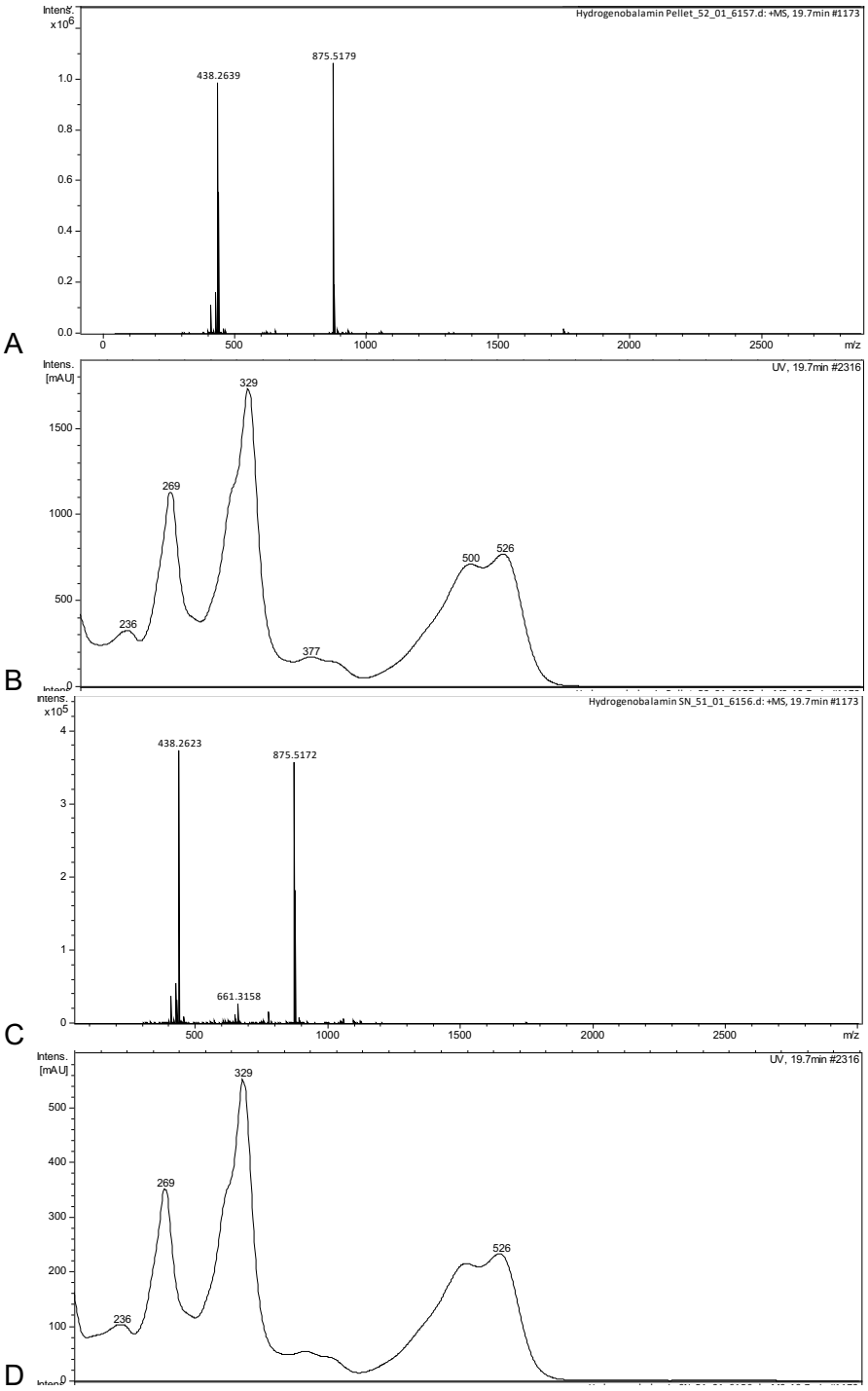


Figure 3.9. Mass spectrometry trace from the *E. coli ED661-cobCDPVU-pET3a* cell pellet and supernatant elution samples. A) Mass spectrum of the major peak from the cell pellet chromatogram. B) UV scan of the major peak from the cell pellet chromatogram. C) Mass spectrum of the major peak from the supernatant chromatogram. D) UV scan of the major peak from the supernatant chromatogram. HBAH double (438) and single charged (875) peaks were

Due to the high HBAH-producing nature of the *E. coli* ED661 strain, it is possible that the bacteria were metabolically compromised and all of the required enzymes were not synthesised. The experiments were therefore repeated using *E. coli* BL21*(DE3) for protein expression. Since *E. coli* BL21*(DE3) is unable to make HBAH, the molecule was provided to the culture medium at a concentration of 10 nM. In order for HBAH to be taken up by the cells, the outer membrane active transporter protein BtuB is required. However, *E. coli* BL21*(DE3) do not have a native functional BtuB. This is due to a nonsense mutation in the *btuB* gene (Studier et al, 2009), which causes translation to be terminated before the protein is fully synthesised. Therefore, to ensure maximum uptake of HBAH, a double transformation was carried out in which *E. coli* BL21*(DE3) was transformed with *btuB*-pLysS, a plasmid containing the *btuB* gene, and *cobCDPVU*-pET3a. Bacteria were then cultured, harvested and corrinoids purified from the bacterial cell pellet using a BtuF-bound nickel column and an RP18 column. The elution sample was analysed using HPLC-MS.

Results showed that HBAH was indeed present, indicated by the 875 m/z peak, but with no signs of its modification since no larger peaks were seen (Figure 3.10) These results correspond with those seen previously, when using *E. coli* ED661, where no modifications to HBAH could be detected.

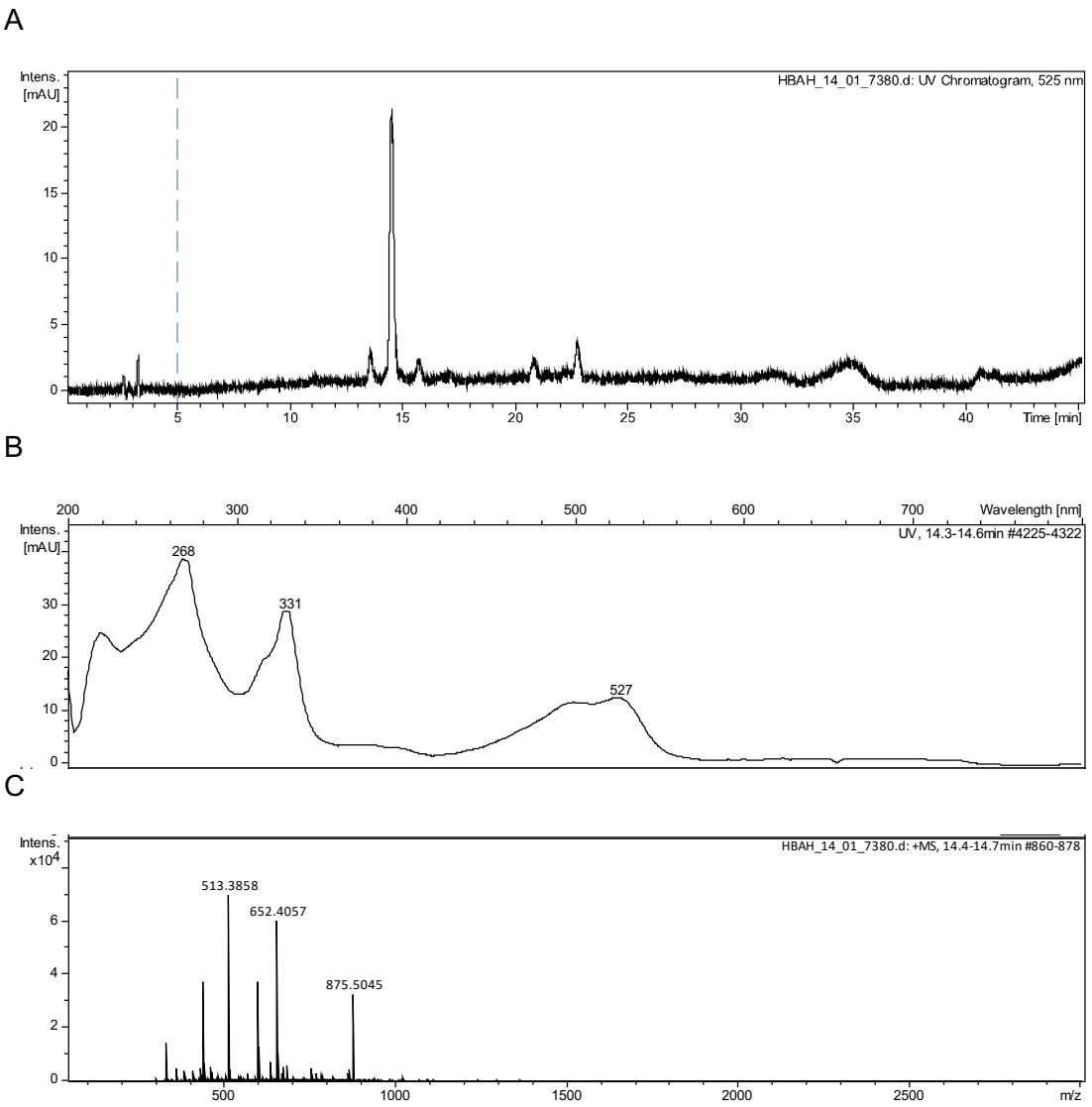


Figure 3.10. Mass spectrometry trace from the *E. coli* BL21*(DE3)-*btuB*-pLysS_*cobCDPVU*-pET3a cell pellet elution sample. A) UV chromatogram showing the different peaks coming off at different timepoints. B) UV scan of the major peak from the cell pellet chromatogram. C) Mass spectrum of the major peak (coming off at 14.7 minutes) from the cell pellet chromatogram. HBAH with a 875 m/z peak was detected. However, no peaks of greater mass were

3.3 Discussion

Currently, very little is known about metal-free corrinoids in relation to both their structure and behaviour, in comparison to natural corrinoids. In 1965, metal-free corrinoids were first discovered and isolated from the photosynthetic bacterium *A. vinosum* (Toohey, 1965). Some of these metal-free corrinoids were investigated for potential activity as B_{12} inhibitors, including a neutral descobaltocobamide (I) and a positively charged descobaltocobinamide (II), whereby growth inhibition was seen in *E. coli* (Perlman and Toohey, 1966). Following on from this, a whole new class of metal-substituted B_{12} analogues, including those with copper, zinc or rhodium at their centre, were synthesised using the extracted metal-free corrinoids (Koppenhagen and Piffner,1970).

The aim here was to investigate the possibility of engineering the production of B_{12} variants, through the construction of an empty, metal-free B_{12} shell, and exploring what effect such compounds might have on cellular control elements, such as riboswitches. To construct this empty, metal-free B_{12} shell, a set of instructions that is able to be read and interpreted by *E. coli* was produced, in the form of a series of genes that are tied together into a single plasmid. Since no central metal ion was desired in this shell, the plasmid did not require the cobalt chelatase genes, *cobNST*, or the cobalt reductase, *cobR.* Additionally, the gene encoding the adenosyltransferase, CobO, was not required, because no upper ligand can be coordinated without a central metal ion. This meant that the plasmid was required to have the following genes, *cobAIGJFMKLHB[QCDPVU]*$^{A/v}$, with the final six from *A. vinosum*. This is because the final six gene products are required to work on metal-free intermediates.

Unfortunately, despite the use of different cloning strategies and the use of various different competent cells, the desired plasmid containing the

cobAIGJFMKLHB[QCDPVU]^{Alv} gene set, could not be constructed. This was most likely due to the plasmid being too large (over 20000 bp), making it toxic to the bacterial cells and reducing transformation efficiency. It was also thought that perhaps the plasmid had too many copies in the cell. Therefore, attempts were made to clone all of the genes into a more stable, low copy number plasmid, pETcocoR (Table 2.2). The majority of genes were successfully transferred to give *cobAIGJFMKLHBQ*^{Alv}-pETcocoR. The presence of the plasmid allowed bacteria to synthesise HBAH and therefore attempts were made to link the *cob[CDPVU]*^{Alv} gene cluster to the genes in the new plasmid, *cobAIGJFMKLHBQ*^{Alv}*-pETcocoR*. Unfortunately, this was unsuccessful and a new approach was undertaken for the biosynthesis of hydrogenobalamin.

This new approach involved transformation of *E. coli* ED661 (genetically engineered to contain *cobAIGJFMKLH[BQ]*^{Alv}) with cob*CDPVU*-pET3a. However, following transformation, the strain was only able to synthesise HBAH and not hydrogenobinamide or hydrogenobalamin. Since it was thought that the cells may be metabolically compromised, resulting in limited expression of the cob*CDPVU* genes, experiments were repeated using *E. coli* BL21*(DE3) instead. HBAH was added externally this time. However, results were the same, showing no sign of hydrogenobalamin production or anything further than HBAH.

Results indicated that there is a problem with the next step after HBAH, catalysed by CobC and CobD (Figure 3.1). It is not clear whether these enzymes are inactive or whether the conditions were not optimal for their activity. However, the real issue is likely to be a lack of CobD, since production of the protein could not be detected in previous experiments. The absence of CobC and CobD activity could also be due to a lack of protein α in *E. coli*, a protein thought to be required for CobC and CobD enzyme activity. In the aerobic pathway, CobC is supposedly part of a complex

called ‘protein β’ with the cobinamide synthase, CobD. Protein β (CobC and CobD complex) is thought to work together with another protein, protein α, to catalyse attachment of aminopropanol to adenosylcobyric acid in the presence of ATP and Mg^{2+} (Blanche et al, 1995), however, the exact identity or role of protein α is unknown. Together, CobC and CobD facilitate aminopropanol synthesis and attachment at C17 of adenosylcobyric acid (Warren et al, 2002; Cohen, 2014) in order to generate adenosylcobinamide (AdoCbi). In this case, *A. vinosum* CobC and CobD were expected to use HBAH as a substrate instead of adenosylcobyric acid, in order to generate hydrogenobinamide. However, this was not the case. Further analysis into CobD production would need to be carried out in order to find out the reason behind the lack of activity. One such method of further analysis could involve the codon optimisation of *A. vinosum cobD,* in order to increase the likelihood of CobD production.

Since there appeared to be no sign of CobC or CobD activity, experiments were carried out to investigate the activity of CobP, CobV and CobU. These experiments are discussed in Chapter 4.

Chapter 4: Structure and activity analysis of the *A. vinosum* CobP and CobU

4.1 Introduction

The final stages of B_{12} biosynthesis involve the assembly of the lower nucleotide loop, with the first step initiated by the bifunctional enzyme, CobP. CobP initiates the process of assembling the lower loop by converting AdoCbi into AdoCbi-GDP (Figure 4.1) (Warren et al, 2002) . CobP is a small, homodimeric protein with two different enzymatic functions, one as cobinamide kinase and the other as cobinamide phosphate guanylyltranserase. In the first step, CobP catalyses phosphorylation of the hydroxyl group of the aminopropanol in AdoCbi, using an ATP donor molecule, to form AdoCbi-phosphate (AdoCbi-P). In the second step, CobP transfers a GMP moiety, from a GTP donor, onto AdoCbi-P to form AdoCbi-GDP (Blanche et al., 1991; Cohen, 2014).

Finally, lower ligand synthesis and assembly requires two enzymes, CobU and CobV, to form AdoCbl from AdoCbi-GDP (Figure 4.1) (Warren et al, 2002). To begin with, CobU catalyses the formation of a glycosidic bond between β-nicotinate mononucleotide (NaMN) and DMB, in order to form α-ribazole phosphate (Blanche et al, 1991; Trzebiatowski et al, 1994; Mander and Liu, 2010). DMB, unless obtained from the environment, is synthesised from $FMNH_2$ by an enzyme called BluB (an oxidoreductase that requires molecular oxygen) (Taga et al, 2007). The α-ribazole phosphate is attached to AdoCbi-GDP by CobV, with the displacement of GDP, to form AdoCbl (Warren et al, 2002).

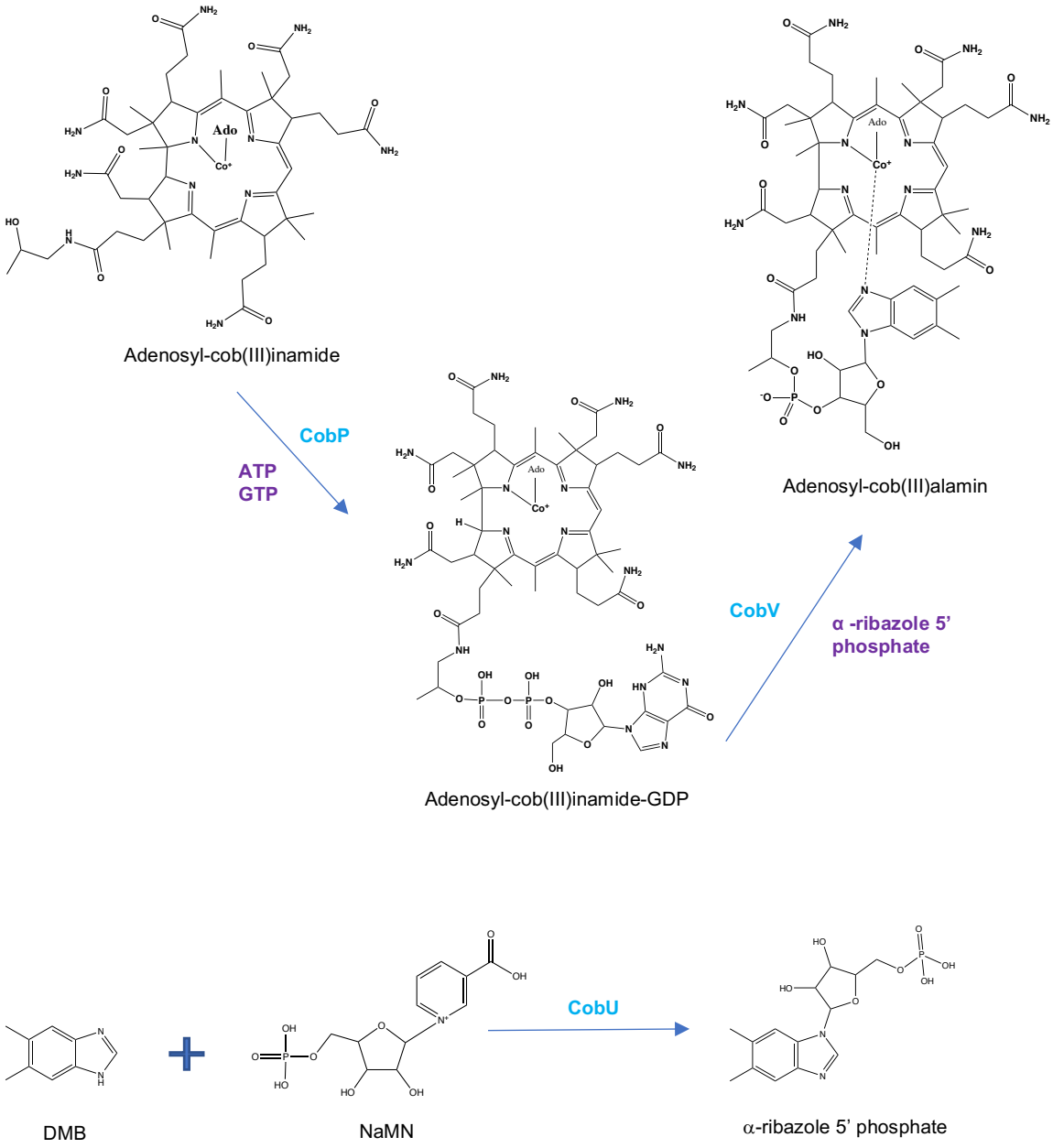


Figure 4.1. Enzymes involved in the conversion of adenosyl-cob(III)inamide to adenosylcob(III)alamin. This process requires the actions of three enzymes in the aerobic biosynthesis pathway; CobP, CobV and CobU.

In Chapter 3, it was observed that the *cobCDPVU*-3a plasmid was unable to provide bacteria with the ability to convert HBAH into hydrogenobalamin or even modify it slightly. However, evidence presented in Chapter 3 strongly indicates that the problem here lies with CobD, as it was not possible to detect any protein being produced. In contrast, both CobP and CobU appeared to be produced in comparatively large quantities. The aim of the research undertaken in this chapter was to initiate structure/function studies on these enzymes.

4.2 Results

4.2.1 Recombinant production and purification of CobP

Previously in Chapter 3, *A. vinosum* CobP was successfully overproduced recombinantly in *E. coli* and purified. Results showed that the 22.5 kDa protein (which includes the 2.2 kDa N-terminal His-tag) was soluble and naturally found in the cytosolic region of the cell. However, upon further inspection, during which a smaller volume of eluted protein sample was loaded onto an SDS gel, CobP appeared to run as two protein bands instead of a single band as initially expected (Figure 4.2).

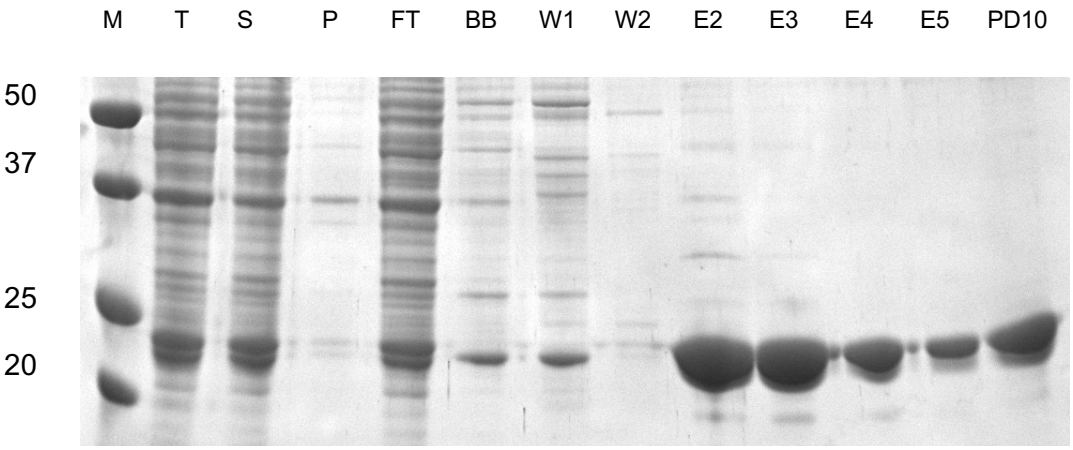


Figure 4.2. SDS-PAGE analysis of protein fractions collected during the purification of CobP from *E. coli* BL21*(DE3)-pLysS. CobP (22.5 kDa with His tag), is visible in the elutions. (Annotations: M – marker, T- total (resuspended sample), S – supernatant, P – pellet, FT –flow-through, BB – binding buffer, W1 – wash buffer 1, W2 – wash buffer 2, E3 – elution 3, E4 – elution 4, PD10 – PD10

From the data, it was not clear whether both bands represent the CobP protein, and if they do, whether one is a degraded form of the other. In order to confirm this was

not just the result of an unwanted impurity, a fresh culture was prepared and CobP protein was purified in the same way. Again, two overproduced protein bands were seen. The most likely explanation for the presence of two protein bands is the degradation of CobP. This was investigated through the use of protease inhibitors.

Due to the potential degradation seen with CobP, a frozen pellet of 1 L bacterial cells (prepared in the same way as before) was defrosted and protease inhibitor (10X SIGMAFAST protease inhibitor tablet dissolved in 10 mL double distilled water) was immediately added prior to sonication. CobP was then purified and fractions analysed using SDS-PAGE (Figure 4.3). Results showed two bands again, suggesting that the protein was still being degraded.

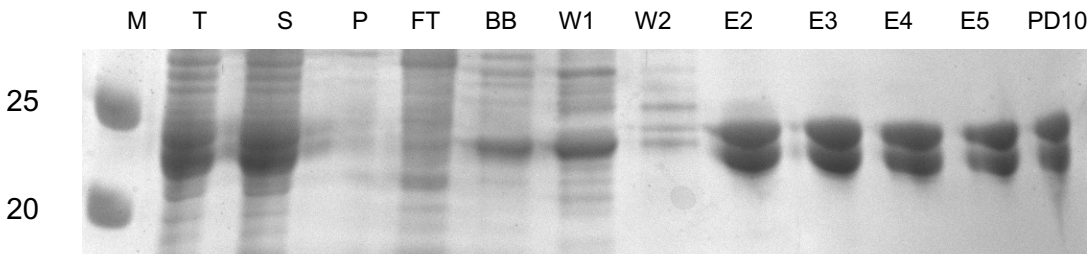


Figure 4.3. SDS-PAGE analysis of fractions from the purification of CobP from *E. coli* BL21*(DE3)-pLysS in the presence of protease inhibitor complex. CobP (22.5 kDa) is visible in the elutions (soluble fractions. Two bands can be seen in T, S and elution lanes.

Since this approach was unsuccessful, further procedures were tested. In the first, cells were cultured and grown fresh prior to the addition of protease inhibitor, as opposed to the addition of protease inhibitor to defrosted cells. In the second approach, cells were given a shorter induction time of 3 hours after the addition of

IPTG, as opposed to overnight incubation. Protein was then purified and fractions collected as previously described (Figure 4.4).

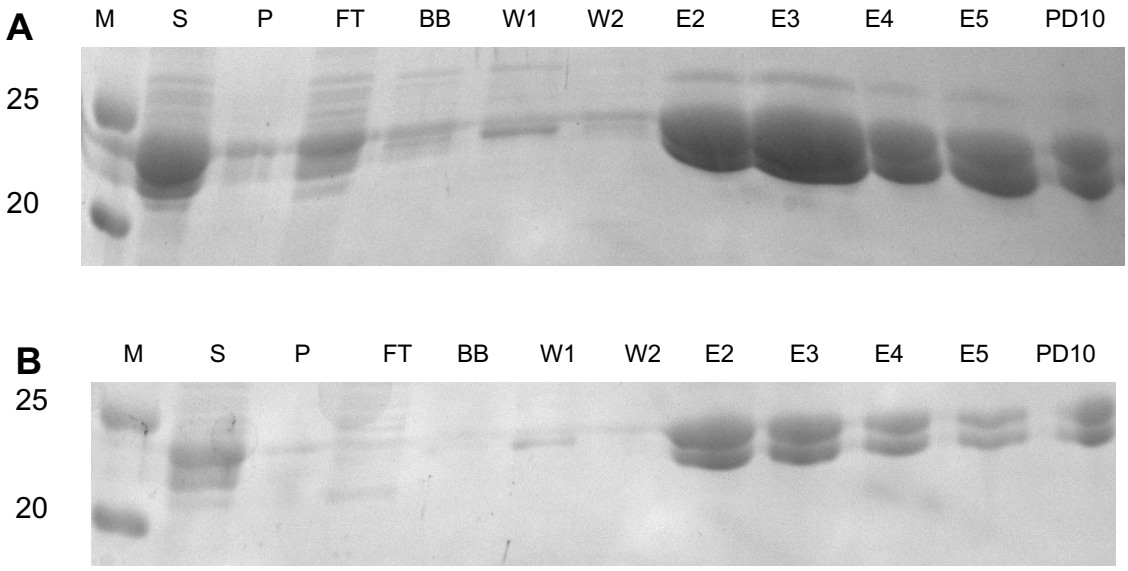


Figure 4.4. SDS-PAGE analysis of fractions from the purification of CobP from *E. coli* BL21*(DE3)-pLysS. Two conditions: A) Purification following the addition of protease inhibitor to fresh cells. B) Purification following a short 3 h induction after IPTG addition.

From the SDS-PAGE analysis it was clear that the addition of protease inhibitor had no effect on the presence of two potential CobP bands, or the use of a shorter induction time. This could perhaps be due to the degradation occurring within the cells themselves, as opposed to during the purification of the protein, after cell lysis.

In order to demonstrate that both bands are in fact representative of CobP, SDS-PAGE was repeated and western blot analysis was performed using anti-His antibodies (capable of detecting the His-tag at the N-terminus of CobP). Samples from both the fresh purification, and the one with added protease inhibitors, were analysed (Figure 4.5).

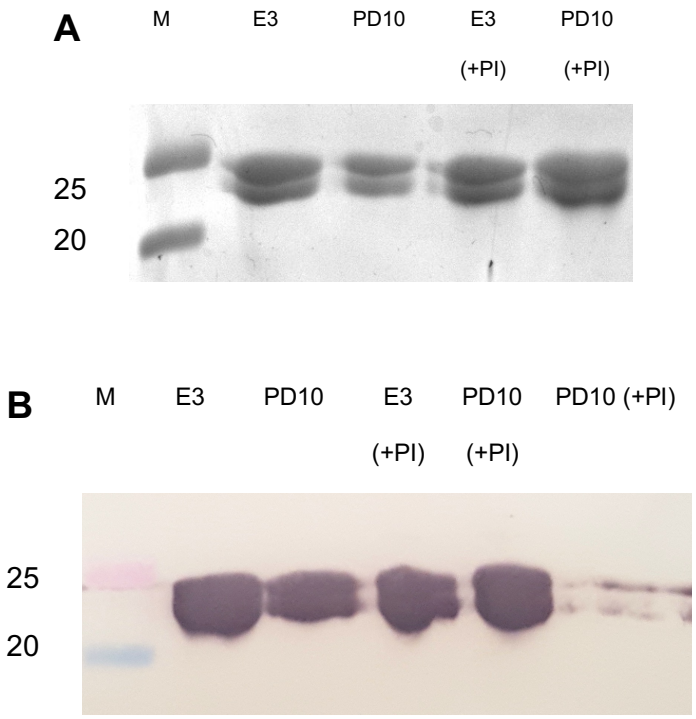


Figure 4.5. Analysis of the two protein bands for CobP purified with and without the addition of protease inhibitor (PI). Two methods of analysis: A.) SDS-PAGE B.) Western blot.

Results from the western blot analysis suggest that both bands must represent CobP (with the N-terminal His-tag) as both cross-reacted with the anti-His antibodies. Therefore, regardless of their exact size or sequence, both should contain at least the N-terminal region of the CobP protein, as this is where the His-tag (6 Histidine residues) is located. The results suggest that the N-terminal is intact and there is degradation of the protein at the C-terminal region. It is highly unlikely that the shorter band is a result of the protein being made at a different start codon since this would results in only one band on the western blot.

As the degradation of CobP could not be prevented, an alternative route to obtaining pure, full-sized CobP protein was investigated. This involved producing CobP with a His-tag at the C-terminus instead of the N-terminus. If degradation is in fact occuring at the C-terminus, then only full sized CobP (with the C-terminal His-tag) will bind the nickel column due to the cleaved fragment having lost its His-tag. A C-terminal His-tag can be incorporated by cloning the gene into a pET23b vector instead of a pET14b vector, and removing the natural STOP codon so that protein synthesis ends at the STOP codon found after the His-tag in pET23b.

In order to clone *cobP* into pET23b, both a 5' end NdeI restriction site and a 3' end XhoI restriction site were required. To obtain these new sites, a reverse primer containing the XhoI restriction site was designed. The gene was then re-amplified, using PCR, from the *A. vinosum* genome. The amplified fragment was cloned into pET23b using the corresponding restriction sites within the plasmid and the new plasmid sequence was verified by sequencing.

Once the *cobP*-pET23b plasmid was successfully constructed, it was used for the transformation of BL21*(DE3)-pLysS. Transformed bacterial cells were cultured, harvested and the purification of CobP was undertaken. Purification fractions were analysed using SDS-PAGE (Figure 4.6).

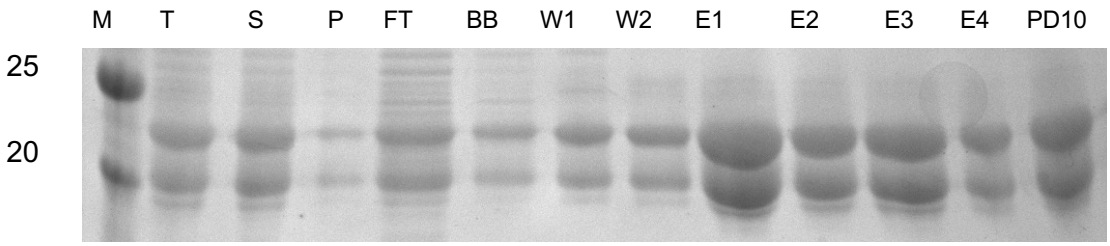


Figure 4.6. SDS-PAGE analysis of fractions from the purification of C-terminal His-tagged CobP from *E. coli* BL21*(DE3)-pLysS. CobP protein is 21.4 kDa (with the 1.2 kDa C-terminal His tag), and is visible in the elutions.

From the SDS-PAGE it appears there are still two bands for the CobP protein despite the change in location of the His-tag. Therefore, either there is no degradation at the C-terminus or the shorter, truncated protein is forming a complex with the full-sized protein which allows it to bind the nickel column and come through in the elutions. To differentiate between these possibilities, C-terminal His-tagged CobP was analysed via western blotting (Figure 4.7).

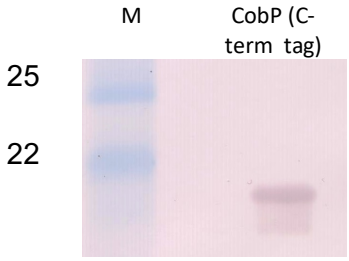


Figure 4.7 Western blot Analysis of CobP with a C-terminal His-tag. CobP with a C-terminal His tag is 21.4 kDa.

Results from the western blot show the C-terminal His-tagged CobP sample to have only one band, suggesting that some form of degradation at the C-terminus has taken place.

4.2.2 Crystallisation of CobP

Currently, there is no known crystal structure for CobP. A crystal structure would be beneficial in helping to understand how the protein works. Therefore, crystallisation was attempted in order to understand the mechanism of the enzyme, despite CobP purifying as two major forms.

CobP was purified using IMAC and PD10 buffer exchange before being concentrated to a concentration of 15.6 mg/mL and further purified through a Fast Protein Liquid Chromatography (FPLC) machine that uses size exclusion chromatography to separate proteins. A G200 Increase column was first calibrated in PD10 buffer and proteins of known mass run through the column to create a standard curve for calibration, before running the protein of interest. The FPLC trace for CobP can be seen in Figure 4.8.

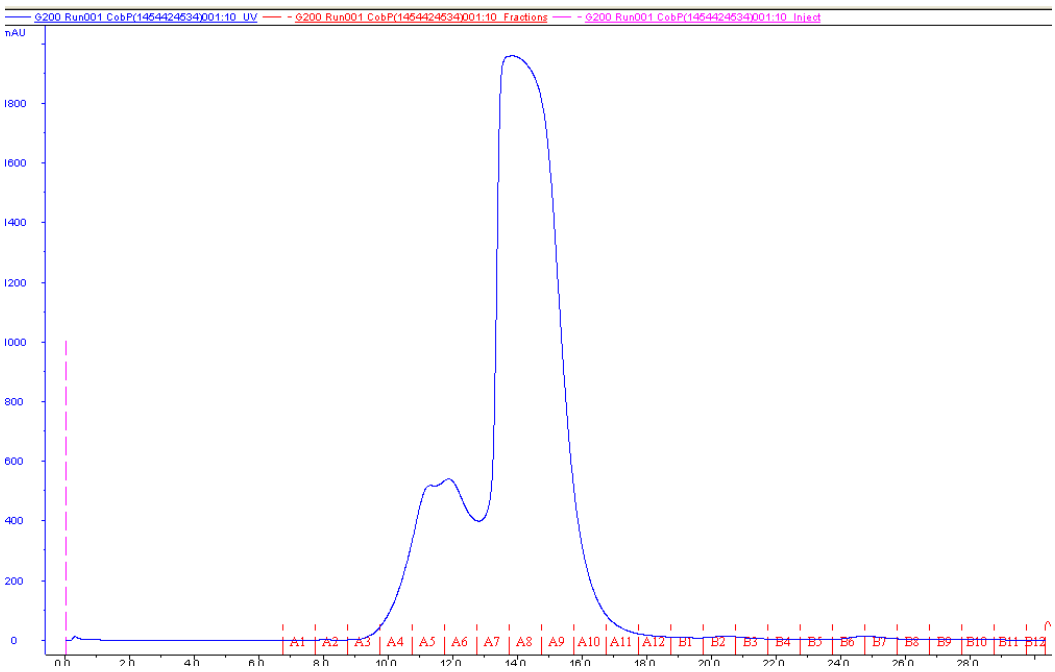


Figure 4.8. Purified CobP FPLC trace. CobP protein elution peak at 13.82 mL.

Using the calibration curve, the protein mass was calculated to be 94 kDa. By dividing this value by the CobP monomer mass, 22.5 kDa, the number of monomers making up a molecule of the active protein could be determined. This was found to be approximately four, suggesting that CobP is running on the column as a tetramer. However, previous research from the literature has suggested CobP to be a homodimer (Blanche, Debussche et al. 1991).

Fractions within and around the protein peak were run on an SDS gel to see which fractions had the most protein (Figure 4.9).

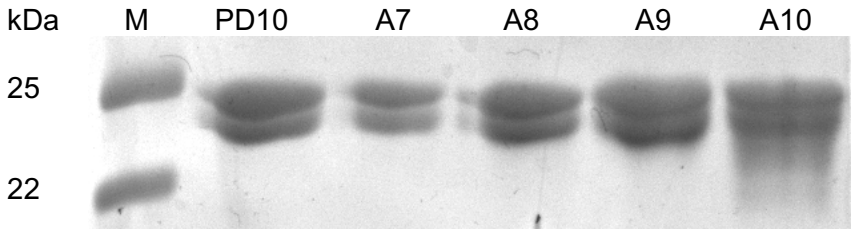


Figure 4.9. SDS-PAGE analysis of CobP (22.5 kDa) purified sample and fractions A7-A10 from within the FPLC protein peak region. The majority of the protein is seen in the A9 fraction.

The exact concentrations were verified using a nanodrop instrument. The most concentrated fraction (A9), with 9 mg/mL protein, was used to prepare four 24-well crystallisation trays, using conditions from structure screens 1 and 2 (Chapter 2, Section 2.4.21).

No crystals were detected and therefore the process was repeated using the addition of substrate. Condition 8 of structure screen 2, (0.2 M ammonium phosphate monobasic, 0.1 M Tris pH 8.5 and 50% v/v 2-methyl-2,4-pentanediol (MPD)), was found to have a potential protein crystal, (Figure 4.10).

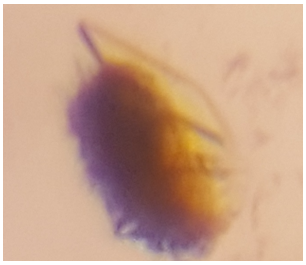


Figure 4.10. Potential CobP crystal. Condition 8 of molecular dimensions SS2 (0.2 M ammonium phosphate monobasic, 0.1 M Tris pH 8.5 and 50% v/v MPD).

The crystal, however, dissolved before it could be analysed further. The condition was therefore optimised (Chapter 2, Section 2.4.22, Table 2.13). Other promising conditions were also repeated simultaneously; 12, 13 and 44 of screen 1, along with 3, 19, 25, 42 and 43 of screen 2. Needles were observed in condition 25 of structure screen 2 (0.01 M cobalt (II) chloride hexahydrate, 1.8 M $(NH_4)_2SO_4$ pH 6.5, 0.1 M MES), Figure 4.11.

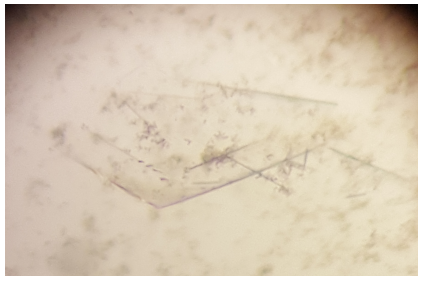


Figure 4.11. CobP needles. Detected in condition 25 of molecular dimensions structure screen 2.

Attempts were made to optimise the condition that had resulted in needle formation (Chapter 2, Section 2.4.22, Table 2.14). Trays were set up using a higher CobP concentration of 0.5 mM.

New trays were also prepared for the 100 conditions from structure screens 1 and 2, using 0.7 mM CobP. Small crystals were observed in condition 8 (structure screen 2), as well as a larger crystal in an optimised condition (0.2 M ammonium phosphate monobasic, 0.1 M Tris, pH 7.5, 40% v/v MPD), Figure 4.12.

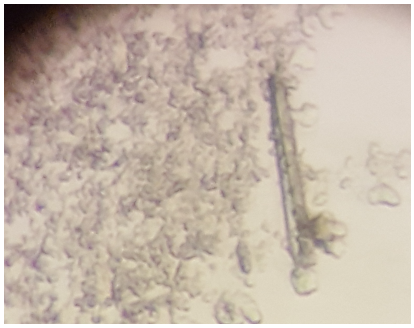


Figure 4.12. Large CobP crystal. Detected in optimised condition containing: 0.2 M ammonium phosphate monobasic, 0.1 M Tris, pH 7.5, 40% v/v MPD.

The condition producing a large CobP crystal was optimised further (Chapter 2, Section 2.4.22, Table 2.15). Crystals were observed in two of these newly optimised conditions (Figure 4.13).

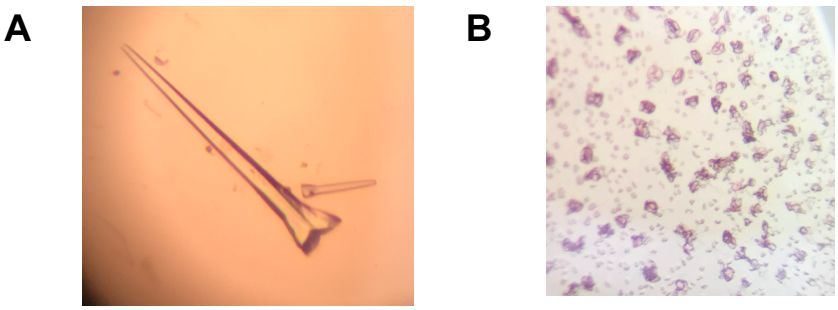


Figure 4.13. Potential CobP crystals. A.) Two identical shaped crystals in the condition containing; 0.1 M Tris pH 8.5, 50% v/v MPD, 0.2 M ammonium phosphate B.) Coloured microcrystals in the condition containing: 1 M Tris, pH 8.3, 50% v/v MPD, 0.2 M ammonium phosphate.

The larger crystal (Figure 4.13A) was analysed by X-ray diffraction, with results demonstrating the crystal to be salt instead of protein. The coloured nature of the second set of crystals, suggested protein had bound to the coloured Cbi substrate and crystallised. Alternatively, it could be that the Cbi had crystallised on its own. Unfortunately, the crystals dissolved before they could be analysed further.

4.2.3 Crystallisation of CobU protein

Currently there are a number of structures solved for the anaerobic version of the CobU enzyme, known as CobT. However, there are no structures for the B_{12} biosynthesis enzymes found in *A. vinosum* bacteria. Since this strain of bacteria has been reported to produce metal-free B_{12} analogues, it is of great interest to see how the enzymes from *A. vinosum* compare to those that are unable to produce metal-free B_{12} analogues. Therefore, crystallisation of *A. vinosum* CobU was attempted.

CobU was purified using IMAC and PD10 buffer exchange before being concentrated to a concentration of 2 mg/mL and further purified through an FPLC machine, using a G200 Increase column. The FPLC trace can be seen in Figure 4.14.

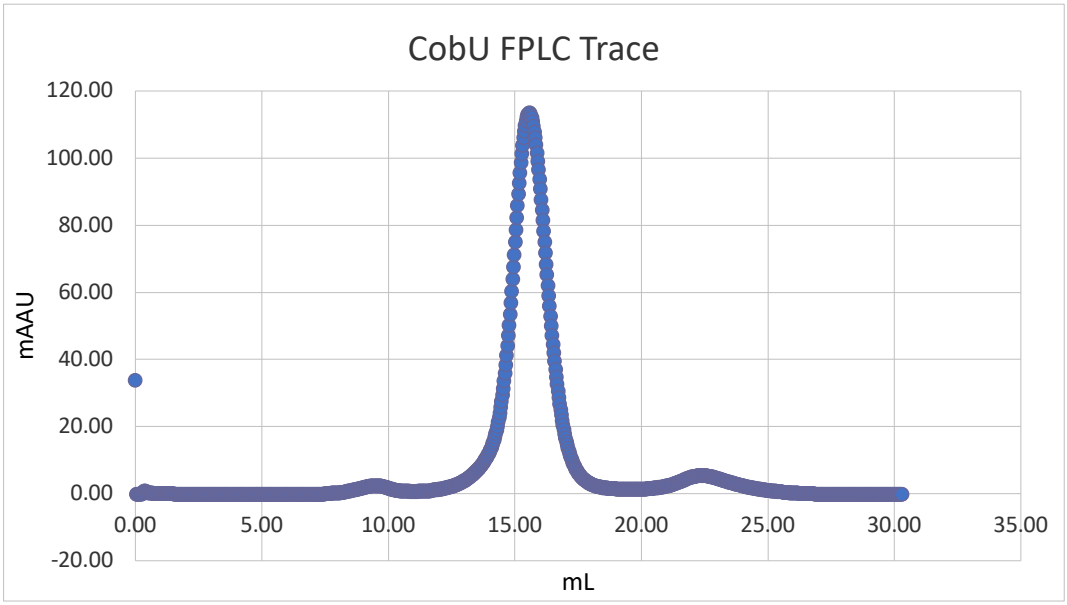


Figure 4.14. Purified CobU FPLC trace. CobU protein elution peak at 15.67 mL.

The concentration of eluted protein was calculated using an A_{280} reading and the CobU extinction coefficient of 12490 L/mol/cm. The concentration, 0.05 mM, was

verified using a nanodrop. The eluted protein was used to prepare four 24-well crystallisation trays, using conditions from structure screens 1 and 2 (Chapter 2, Section 2.4.21).

Small crystals were observed in conditions 14 (0.2 M $(NH_4)_2SO_4$, 30% w/v polyethylene glycol (PEG) 8000, 0.1 M sodium cacodylate pH 6.5), 15 (0.2 M magnesium acetate tetrahydrate, 20% w/v PEG 8000, 0.1 M sodium cacodylate pH 6.5) and 30 (2.0 M $(NH_4)_2SO_4$, 2% V/v PEG 400, 0.1 M sodium 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.5) of structure screen 1. Optimisation screens were prepared for each condition, in order to produce larger, more well-defined crystals (Chapter 2, Section 2.4.22; Tables 2.16, 2.17 and 2.18). Two conditions (0.1 M sodium cacodylate (Na/Caco), 0.15 M $(NH_4)_2SO_4$, 20% PEG and 0.1 M Na/Caco, 0.1 M magnesium acetate (Mg/Ace), 15%) grew distinct crystals (Figure 4.15) which were picked for X-ray diffraction (Chapter 2, Section 2.4.23).

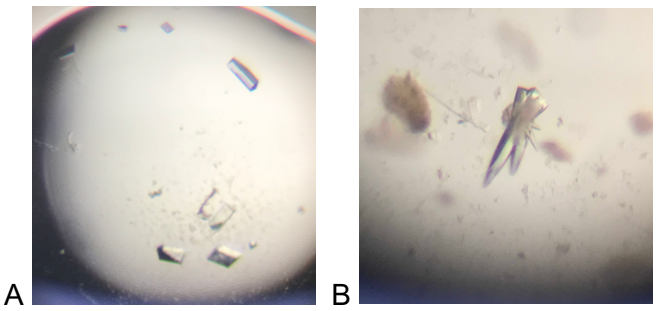


Figure 4.15. CobU crystals. Detected in the following conditions: A) 0.1 M Na/Caco, 0.15 M $(NH_4)_2SO_4$, 20% PEG from SS1 condition 14 screen, B) 0.1 M Na/Caco, 0.1 M Mg/Ace, 15% PEG from SS1 condition 15 screen.

Both samples were found to contain protein crystals and diffraction data were obtained for each. However, due to anisotropy and the formation of ice rings, a structure for CobU could not be generated.

4.2.4 The synthesis of adenosylcobinamide to investigate the activity of CobP

In order to determine if CobP is active, substrate for the enzyme needs to be prepared, specifically the synthesis of AdoCbi. This section describes the use of BtuR to produce AdoCbi from Cbi, following the decyanation of CNCbi into Cbi, for subsequent use in CobP activity assays. One way of synthesising the CobP substrate, AdoCbi, involves the use of the BtuR enzyme, which is the adenosyltransferase found in *E. coli* that has the ability to adenosylate Cbi into AdoCbi (Lundrigan et al, 1991), Figure 4.16. BtuR was overexpressed by culturing BL21*(DE3)-pLysS-*btuR*-pET14b (provided by Dr. Evelyne Deery), and purified using IMAC.

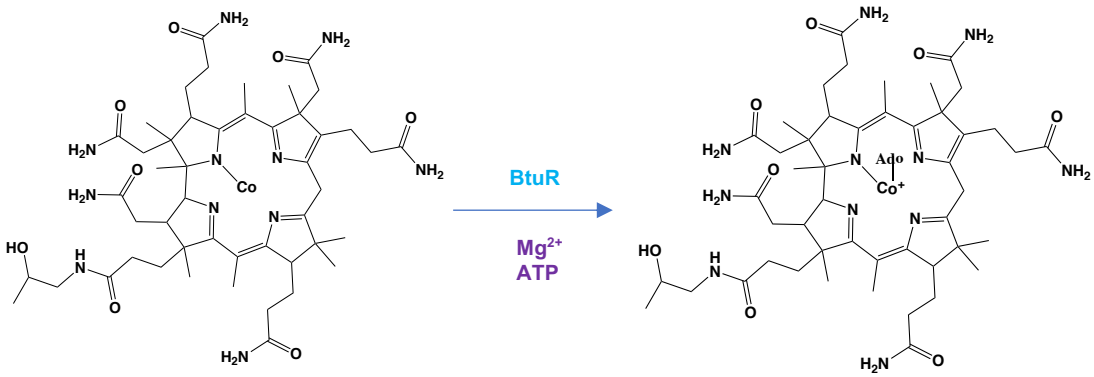


Figure 4.16. The conversion of cobinamide to adenosylcobinamide by BtuR. BtuR, an adenosyltransferase, catalyses the conversion of cobinamide into adenosylcobinamide.

The first stage involved the decyanation of CNCbi (obtained from Sigma) into Cbi, via a number of chemical reactions, under anaerobic conditions. This involved the reduction of Co^{3+} in CNCbi to Co^{1+} using sodium borohydride. Incubation with sodium borohydride was carried out for 2-3 hours, by which point completion of the reaction could be detected by a colour change from red/purple to grey/purple. Sep Pak Plus C18 RP18 cartridges were set up and activated with 30 mL methanol before equilibration with 10 mL de-gassed H_2O. The reduced Cbi was then passed through the column and eluted with 5 mL methanol before drying down in a vacuum centrifuge.

In the second stage, the adenosyl upper ligand was attached to Cbi through the action of BtuR, in order to produce AdoCbi. Purified BtuR protein was taken into a glove box and run through a PD10 column anaerobically so that an absorbance scan measurement could be taken. This was used to calculate the concentration of BtuR. Similarly, the concentration of Cbi was calculated after dissolving the dried powder in PD10 buffer. The assay (set up as described in Chapter 2, Section 2.3.17) was carried out in the dark due to the light sensitivity of AdoCbi, and to prevent removal of the adenosyl group, before incubating overnight in the glove box.

Once incubation was complete, the solution was centrifuged for 15 minutes at 4000 rpm in order to deposit the protein to the bottom of the tube. The supernatant was run through an RP18 column and AdoCbi eluted in 50% ethanol (Chapter 2, Section 2.4.18). A 50 μL sample was also removed for analysis using HPLC-MS (Chapter 2, Sections 2.4.19 and 2.4.20) to confirm that the sample was indeed AdoCbi. Results confirmed that AdoCbi had been successfully made (Figure 4.17).

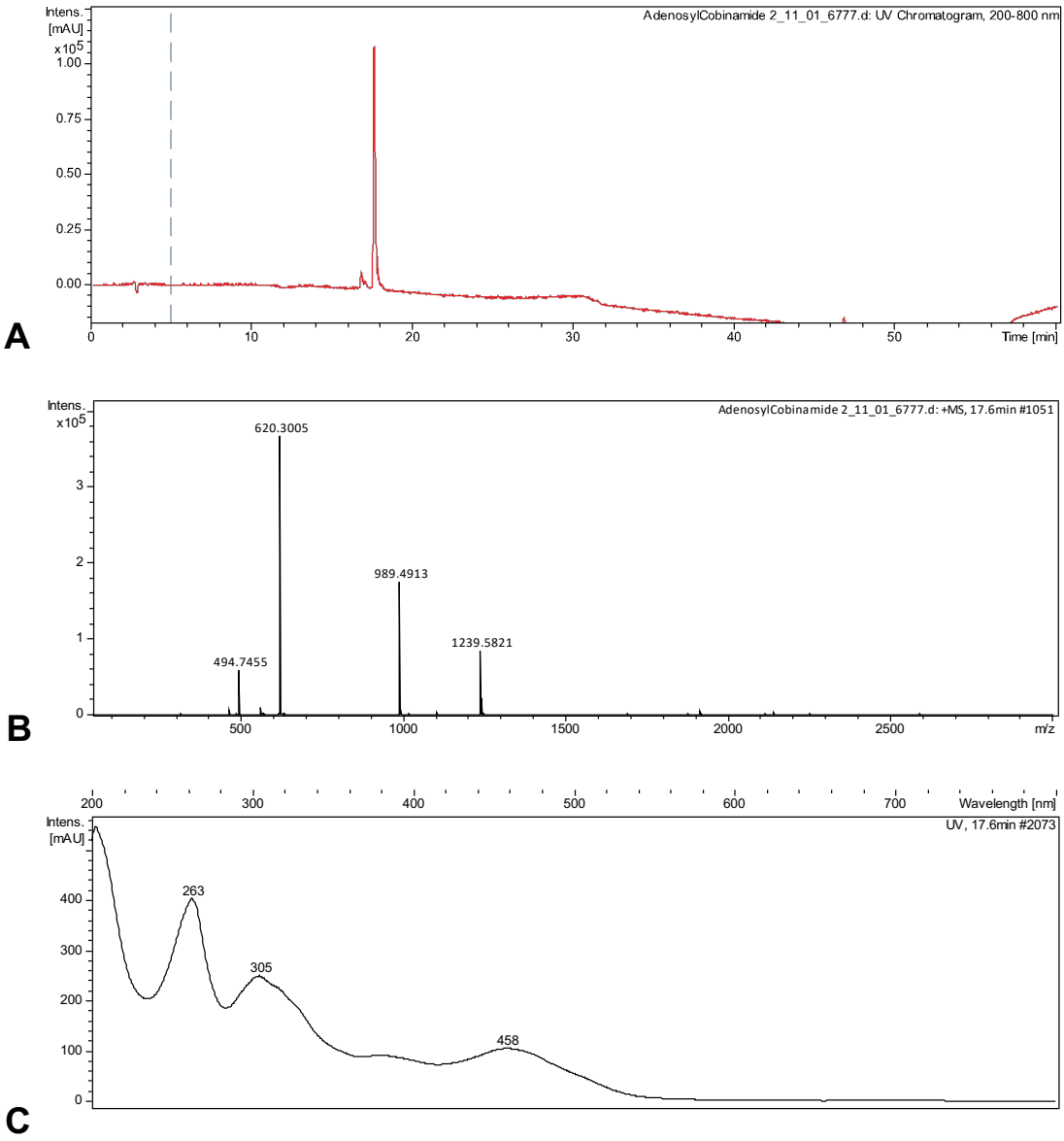


Figure 4.17. Mass spectrometry results showing the production of AdoCbi. A.) UV chromatogram showing the different peaks coming off at different timepoints. B.) Mass spectrometry trace of the large peak seen in (A) coming off at 17.6 minutes. The 1239.58 peak represents the mass for AdoCbi, suggesting that it has been correctly made. C.) UV spectrometry displaying the correct trace for AdoCbi.

4.2.5 *In vivo* CobP activity assays

In Chapter 3, it was seen that the enzymes encoded by the cob*CDPVU*-pET3a plasmid had no effect on HBAH as they would be expected to. Therefore, to see if any CobP, CobV or CobU activity could be detected, the same growth experiments were carried out using a later intermediate in the pathway, Cbi (the substrate for CobP).

E. coli BL21*(DE3)-pLysS-cob*CDPVU-pET3a* was grown with the addition of 100 pM CNCbi, and samples prepared as described in Chapter 2, Section 2.3.18. Corrinoids were purified using a BtuF-bound nickel column and a RP18 column. Elutions from the supernatant and pellet samples were analysed using HPLC-MS. The results indicated no production of Cbl or any CobP enzyme products (Cbi-GDP or Cbi-P) with only the starting material, CNCbi, detected.

Next, *E. coli* BL21*(DE3)-pLysS-cob*CDPVU-pET3a* and *E. coli* BL21*(DE3)-pLysS were individually grown, with the addition of 80 nM newly synthesised AdoCbi. In the control culture, lacking any *A. vinosum* enzymes, no molecules could be detected with masses that correspond to Ado-Cbi-P, Ado-Cbi-GDP or AdoCbl. In the culture containing cob*CDPVU-pET3a*, AdoCbi appears to have been modified to a molecule with a mass unit value of 1557.63 Da, ~22 Da less than that of AdoCbl and ~318 Da greater than AdoCbi (Figure 4.18). This was not detected in the control experiment.

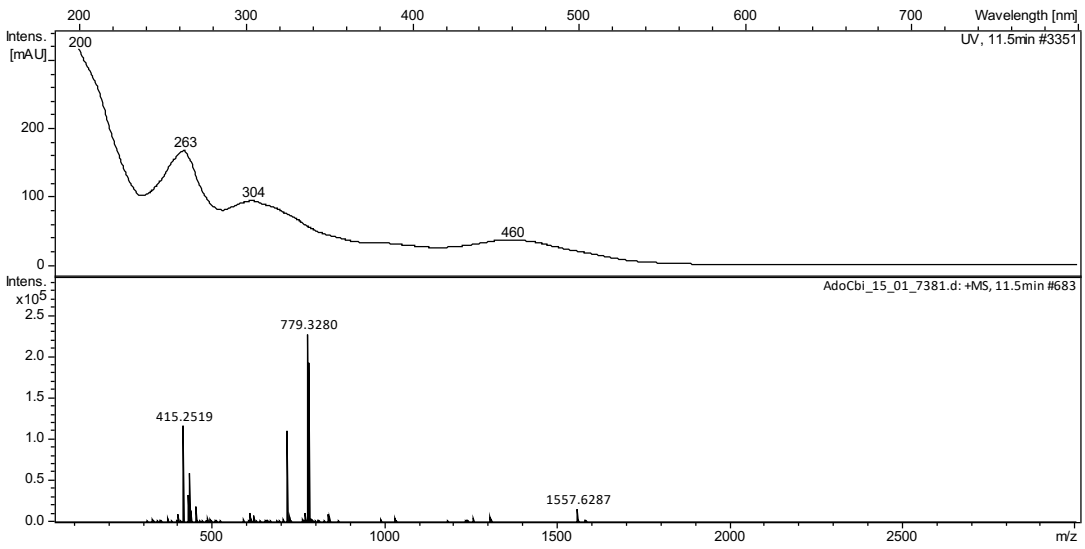


Figure 4.18. Mass spectrometry results showing the production of an unrecognisable molecule from AdoCbi. A.) UV chromatogram. B.) Mass spectrometry trace showing the 1557.63 peak representing a molecule synthesised from AdoCbi.

In order to confirm whether the adenosyl group was still attached in the molecule of 1557.63 Da mass, the sample had cyanide added (to displace any adenosyl at the upper ligand position) before analysing again. Results show a peak of smaller mass suggesting that the adenosyl group was successfully replaced with the cyano-group (Figure 4.19). The most likely suggestion for this is the addition of a lower ligand that is not DMB (found in cobalamin) due to the absence of DMB in the culture.

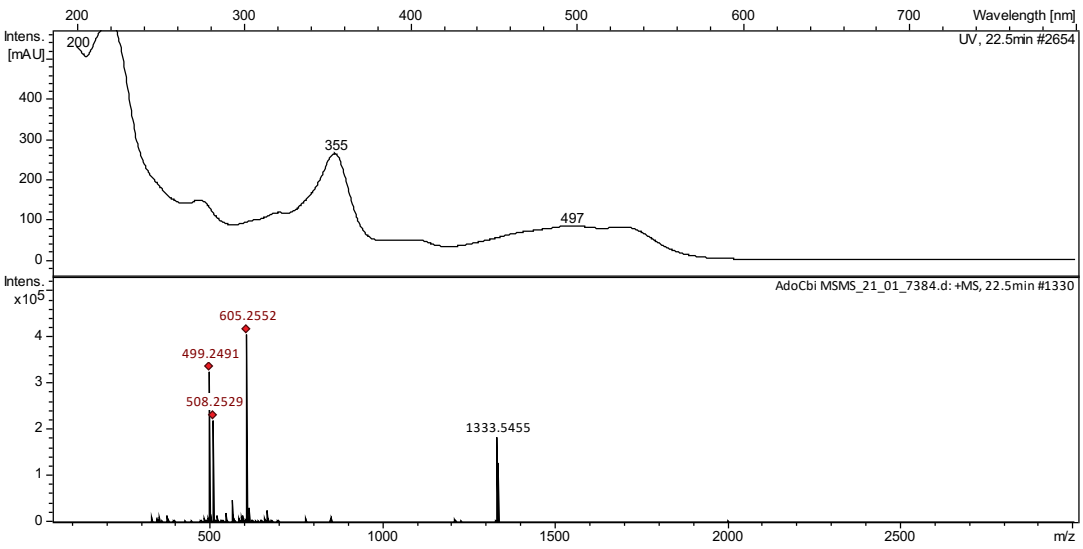


Figure 4.19. Mass spectrometry results showing the change in the assay product following cyanide addition. A.) UV trace of the 1333.55 mass unit molecule. B.) Mass spectrometry trace showing the 1333.55 peak.

Although the structure of the compound remains unknown, the fact that this was not seen in the control experiment suggests that the *A. vinosum* enzymes had some role in its production. A more in-depth MS-MS and NMR study would be needed to identify the lower ligand that was being attached. Due to this ambiguity of results when working *in vivo*, attempts were made at an *in vitro* approach using purified *A. vinosum* CobP.

4.2.6 *In vitro* CobP activity assays

In order to detect *A. vinosum* CobP enzyme activity, an *in vitro* approach using purified enzyme, various different Cbi substrates, and additional cofactors ($MgCl_2$ and GTP/ATP) was trialled. These experiments are described in Chapter 2, Section 2.4.25.

Initially, 0.6 μM of purified CobP was used in the assays (Table 2.19) and tested alongside controls lacking protein and controls lacking GTP/ATP. Results showed no difference in the assays that did or did not have protein (Figure 4.20). Therefore, a higher concentration of protein, 6 μM, was used and a different buffer pH also tested. Each reaction was carried out with 20 μM of each Cbi at both the original pH of 8 and a new higher pH of 9, to see if conditions could be optimised (Table 2.20). Results showed a peak for AdoCbi-P in the sample provided with AdoCbi, in pH 9 buffer. This suggested that the enzyme is active and a pH of 9 is more optimal for the enzyme (Figure 4.21). Previously carried out research also suggests an optimal pH of around 9 (Blanche et al, 1991). Additionally, AdoCbi was seen to be the optimal substrate as opposed to OHCbi, and therefore only AdoCbi was used for the following assays.

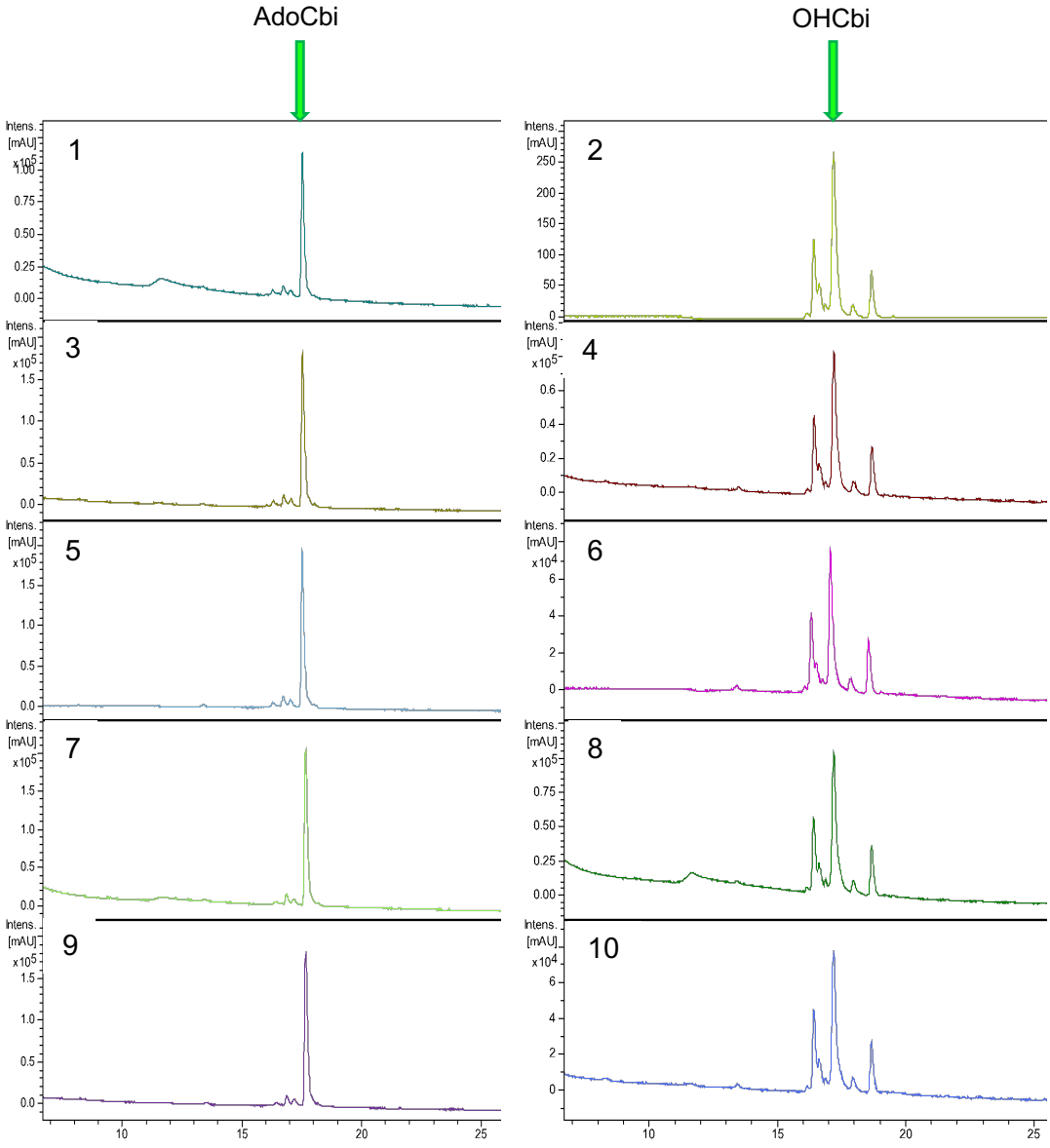


Figure 4.20. Mass Spectrometry traces from CobP activity assays. Assays 1-4 were controls containing no CobP protein, in which no activity should be detected. Assays 5-6 were no ATP/GTP controls, in which no activity should be detected. Assays 7-10 all contain CobP, either ATP or GTP (to determine which works better in this assay) and either OHCbi or AdoCbi as the substrate. The exact components in each assay are outlined as follows: 1.) GTP, AdoCbi. 2.) GTP, OHCbi. 3.) ATP, AdoCbi. 4.) ATP, OHCbi. 5.) CobP, AdoCbi. 6.) CobP, OHCbi. 7.) CobP, AdoCbi, GTP. 8.) CobP, Cbi, GTP. 9.) CobP, AdoCbi, ATP. 10.) CobP, Cbi, ATP. In all ten assays, no changes were seen to be made to either AdoCbi or OHCbi, therefore CobP activity could not be detected in any of them.

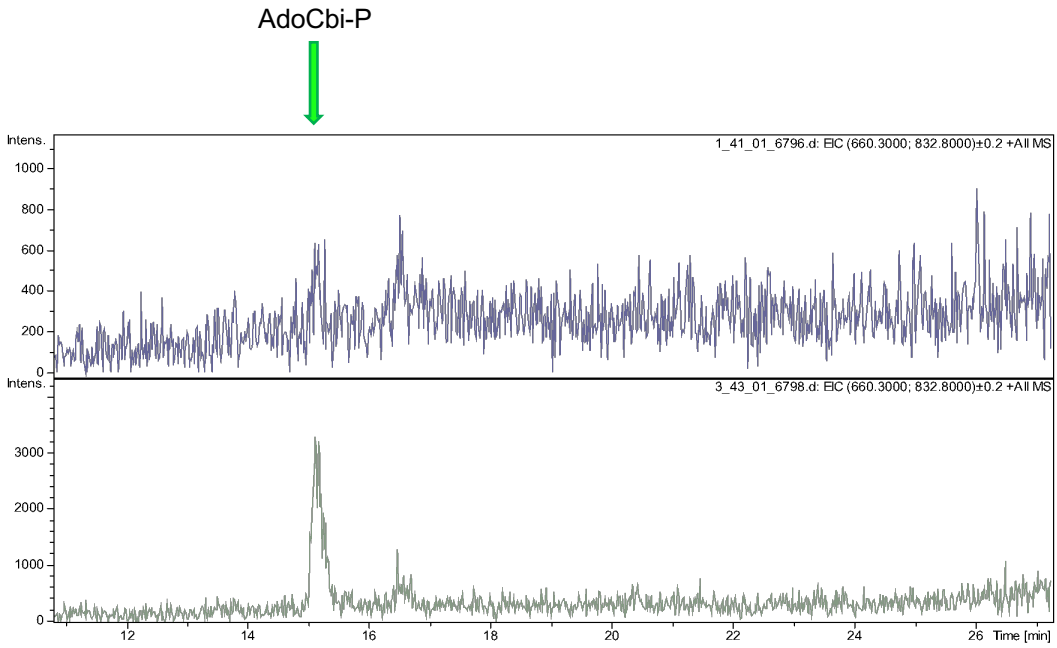


Figure 4.21. Mass spectrometry traces from CobP activity assays in optimised buffer conditions. Assay carried out at pH 8 (above) shows only background peaks and no indication of AdoCbi-P presence. Assay carried out at pH 9 (below) shows a peak with a mass corresponding to that of AdoCbi-P (~1334 Da).

Following the initial detection of CobP activity, an even higher concentration of CobP (60 μM) was investigated for improved assay performance (Table 2.21). Reactions were tested in both the presence and absence of DTT (Table 2.21). This is because DTT prevents cysteine oxidation (Yang et al, 2012), and cysteine oxidation within the active site could potentially be a factor hindering CobP activity. Additionally, control reactions were carried out using lysate from cultures overproducing *A. vinosum* CobP, lysate from cultures overproducing *R. capsulatus* CobP and purified *R. capsulatus* CobP, in addition to the standard reaction using only purified *A. vinosum* CobP.

Results showed that the presence of DTT had no effect on CobP activity (Figure 4.22A and 4.22B), suggesting that cysteine oxidation at the CobP active site was not hindering its activity. *A. vinosum* CobP was found to be active, since purified protein generated the desired product, Ado-GDP-Cbi. This signal was stronger when lysate from the same culture was used instead of purified protein directly (Figure 4.22A and 4.22C), suggesting that perhaps there may be a component present in the lysate that is required for optimal CobP activity. This could potentially be a cofactor or metal ion. *R. capsulatus* CobP was found to be more active than *A. vinosum* CobP since the Ado-GDP-Cbi signal was stronger than that generated by *A. vinosum* CobP (Figure 4.22A and 4.22D). However, without quantifying the substrate to product turnover, it cannot be said how much more active the *R. capsulatus* version is.

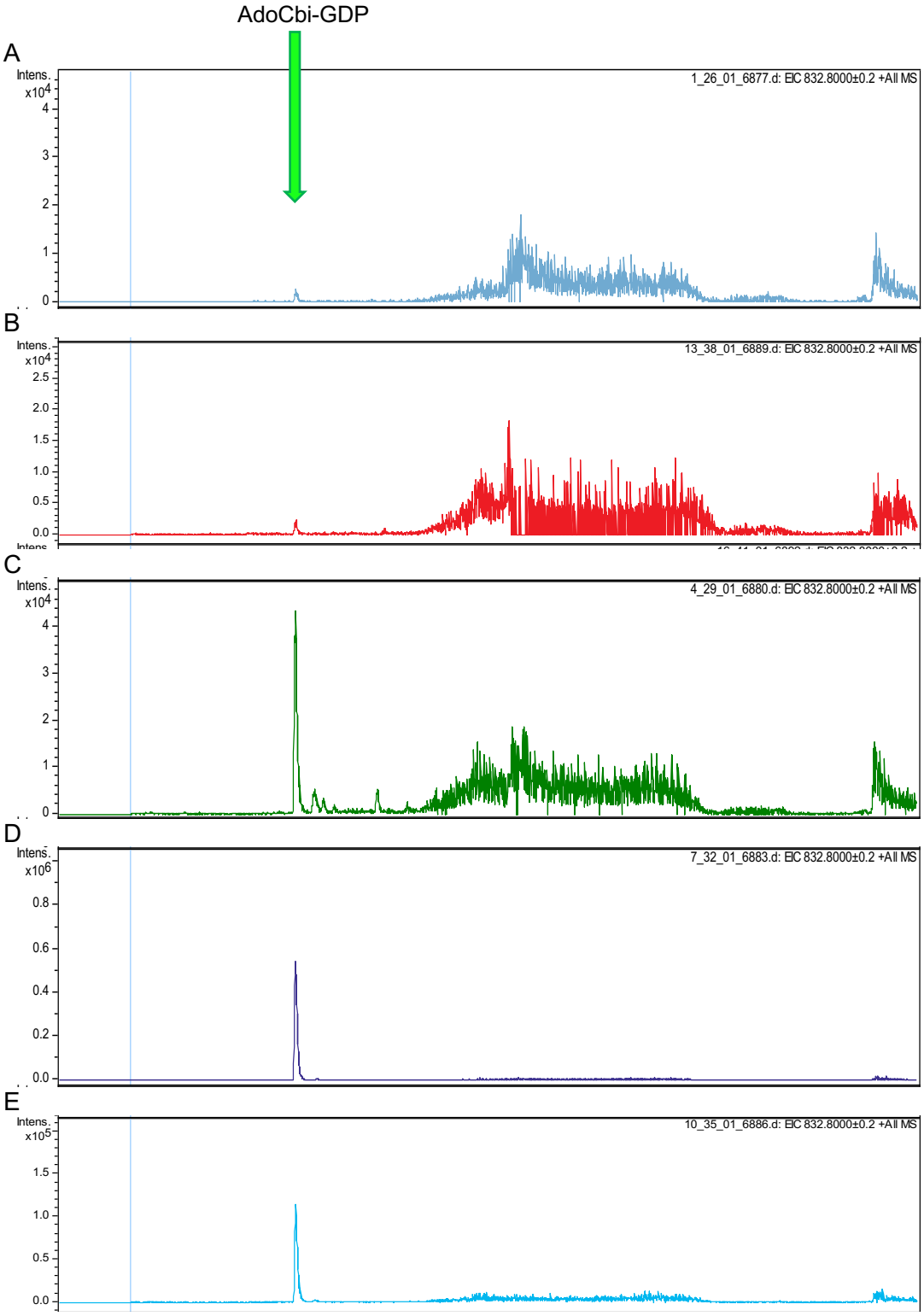


Figure 4.22. Mass spectrometry results for the production of AdoCbi-GDP from AdoCbi. A) Assay using purified *A. vinosum* CobP, B) Assay using purified *A. vinosum* CobP in the presence of 10 mM DTT, C) Assay using lysate from *A. vinosum* CobP overproducing culture, D) Assay using purified *R. capsulatus* CobP, E) Assay using lysate from *R. capsulatus* CobP overproducing culture.

In a final assay set, the lysate overproducing *A. vinosum* CobP was compared to that of standard BL21*(DE3)-pET14b, in order to confirm that the activity was indeed coming from the *A. vinosum* CobP and not solely from something found in the lysate. Purified *A. vinosum* CobP at a 10X higher concentration of 600 μM was also tested in the presence and absence of standard BL21*(DE3)-pET14b lysate (Table 2.22). Additionally, a final control containing no lysate and no purified protein was also prepared. Of the five reactions, the *A. vinosum* CobP overproducing lysate was seen to generate the most Ado-GDP-Cbi, with the greatest turnover of substrate to product (Figure 4.23), demonstrating that BL21*(DE3)-pET14b lysate alone has some background activity, but since the substrate to product turnover seen in the *A. vinosum* CobP overproducing lysate is over 4 times greater, the CobP enzyme must be responsible for this greater activity. In the BL21*(DE3)-pET14b lysate samples, turnover was 1-2%, regardless of the presence of purified CobP protein. This final set of results shows that while *A. vinosum* CobP does have activity, this activity is very low and there must be some underlying problem causing this, such as the substrate not being optimal for the enzyme. Additionally, the activity seen here appears to be aided by one or more unknown components in the lysate.

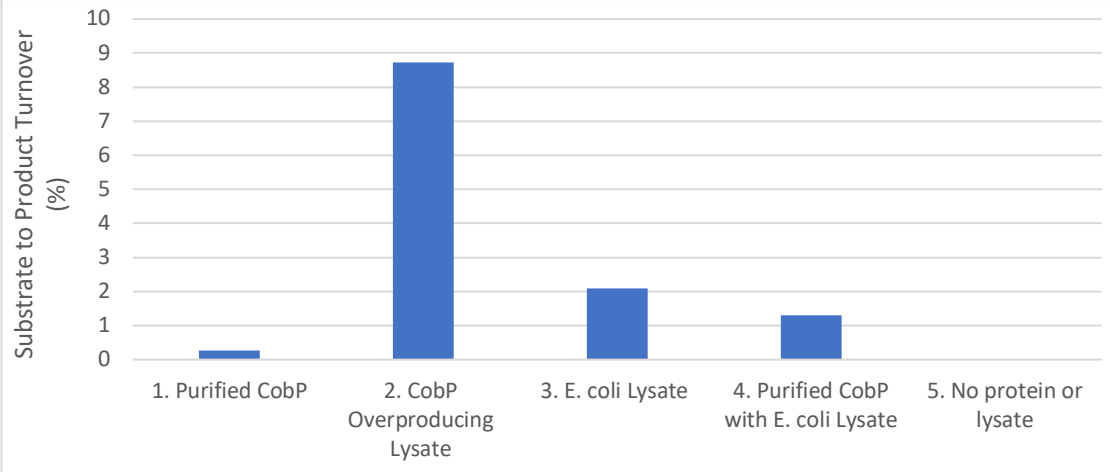


Figure 4.23. Graph to show the rate of AdoCbi turnover, to form AdoCbi-GDP. Highest turnover is seen from the *A. vinosum* CobP overproducing lysate.

4.3 Discussion

The work carried out in this chapter is heavily focused on working out the structure of two B_{12} biosynthesis enzymes from *A. vinosum*, CobP and CobU, and determining whether the recombinantly produced proteins are active and suitable for the '*ex vivo'* synthesis of corrins. Currently, very little is known in the literature about the structure of CobP due to the lack of a crystal structure. The structure of CobU, from the aerobic pathway, has not yet been solved either. However, the anaerobic version of the enzyme, known as CobT, has had its structure solved. For example, CobT from *Salmonella enterica* has a crystal structure which shows the enzyme to be a dimer, with each subunit consisting of two domains (Cheong et al, 1999). Similarly, CobP is also thought to be a homodimer due to previous research carried out where the protein was found to have double the molecular weight on native-PAGE compared to when denatured on SDS-PAGE (Blanche et al, 1991).

In terms of its function as an enzyme, CobP is thought to be a bifunctional protein with the ability to act both as a cobinamide kinase and cobinamide phosphate guanylyltransferase (Blanche et al, 1991). This putative bifunctional nature of CobP opens up an interesting question in relation to its structure and whether the two reactions that are carried out occur at the same active site or at different sites, whereby an intermediate is transferred between them. Additionally, the *A. vinosum* enzymes are interesting in general, due to the bacterium's ability to synthesise metal-free corrinoids (as discussed in Chapter 3).

Unfortunately, despite many attempts, crystallisation of CobP proved unsuccessful. Initial experiments on the protein determined that CobP was subject to some form of C-terminal cleavage. It is perhaps for this reason that the protein could not be

crystallised, since crystallisation requires homogenous solutions with all protein components being identical, allowing them to form identical, tightly packed lattices.

In contrast, CobU crystallised much more readily, producing a range of crystal shapes depending on the condition of growth. Unfortunately, due to problems during the freezing of crystals, difficulties arose during X-ray diffraction which prohibited the collection of appropriate structural data. One crystal provided a diffraction data set that was Space Group P21 and which diffracted with an average overall resolution of 2.7Å. However, as the data were anisotropic, it was not possible to generate a structure using molecular replacement with the structural information from the *S. enterica* CobT crystal structure (Cheong et al, 1999). Future work would require further refining of the crystal-producing conditions, followed by improved cryo-protectants such that there is no anisotropy or formation of ice rings.

Subsequent to the structural analysis work carried out on CobP and CobU, the activity of CobP was investigated, following on from *in vivo* growth assays carried out in Chapter 3, where all five genes in the *cobCDPVU*-3a plasmid were trialled together. Here the focus was more specifically on the activity of CobP. Initial *in vivo* experiments showed some form of modification to the AdoCbi substrate, which suggested that *A. vinosum* CobP may be active (since the same result did not occur in the control lacking the *cobCDPVU*-3a plasmid). Additionally, CobV appeared to incorporate some form of lower ligand to the cobalt ion, suggesting that CobV is also being recombinantly produced. However, the resulting product could not be deduced, as the attached lower ligand was not of a size that was recognised.

In order to investigate CobP activity further, *in vitro* activity assays were carried out. Results showed that optimal activity was observed when taking protein along with the lysate from which it originates, suggesting that there may be an unknown

component in the lysate aiding the reaction. Alternatively, it may be that the correct substrate is not present and that the enzyme requires a metal-free form of Cbi.

Chapter 5: An *in vitro* characterisation of the *E. coli btuB* riboswitch

5.1 Introduction

This study focuses on the control of the *btuB* riboswitch, and in turn B_{12} uptake, via the *btuB* gene product, BtuB. Riboswitches are short mRNA sequences that are bound by a ligand/metabolite to terminate transcription/translation of the associated gene(s). One of the first discovered riboswitches was the B_{12} riboswitch. When B_{12} is bound to this mRNA sequence, expression of B_{12}-related genes is switched off. These riboswitches are discussed in more detail in Chapter 1, where structures of B_{12} riboswitches are also shown.

Initial work on the *btuB* riboswitch began when the level of *btuB* expression within *E. coli* was found to be controlled by an mRNA leader sequence (Lundrigan et al, 1991). This mRNA leader sequence was later discovered to be an AdoCbl-responsive riboswitch, the *btuB* riboswitch, controlling *btuB* gene expression in response to AdoCbl binding. AdoCbl was found to bind directly to the mRNA and induce a structural reorganisation, which results in a change in expression levels (Nahvi et al, 2002). Through the use of primer extension-inhibition assays, which measure the ability of bound ribosomes to block extension by reverse transcriptase of a primer hybridized to the *btuB* RNA, the concentration of AdoCbl required to cause half-maximal inhibition has been estimated to be 0.3 μM (Nou and Kadner, 2000). Following this, in-line probing studies were carried out which yielded apparent K_D values of 0.3 μM for AdoCbl binding to the *btuB* riboswitch (Nahvi et al, 2002). Thus, a concentration of 0.3 μM of ligand is required to alter the structure of half of the RNA molecules present.

The following investigations were aimed at determining the effectiveness of the *E. coli btuB* riboswitch using an *in vitro* transcription/translation kit, a method allowing

changes in levels of the gene product to be established in response to increasing B_{12} concentrations. This approach has the potential to allow a more accurate and direct determination of the riboswitch response to B_{12} concentrations and examine the ability of the riboswitch to respond to B_{12} variants. In order to gain greater insight into aspects of molecular recognition, various B_{12} analogues were investigated for their ability to bind and structurally alter the *btuB* riboswitch.

The natural variants investigated in this chapter include AdoCbl, CNCbl, AdoCbi and CNCbi, all of which are compounds with either a 5'-deoxyadenosyl or cyano group upper ligand. The remaining unnatural variants were synthesised by Dr. Florian Widner and Christoph Kieninger, and provided by Professor Bernhard Kräutler at the University of Innsbruck. The first of these is the unnatural AdoCbl analogue, adenosylrhodibalamin (AdoRbl), which differs to AdoCbl only in the nature of the central metal ion where cobalt is replaced with rhodium (Widner et al, 2016). The next two variants differ to AdoCbl in their upper ligand where the 5'-deoxyadenosyl group is replaced with difluorophenylethinyl and ethylphenyl upper ligands in order to produce difluorophenylethinylcobalamin (F_2PhEtyCbl) and ethylphenylcobalamin (EtPhCbl), respectively (Ruetz and Salchner et al, 2013; Ruetz and Gherasim et al, 2013). The final variants consist of altered central metal ions where cobalt is replaced with zinc and nickel to produce zincobalamin (Znbl) and nibalamin (Nibl), respectively. Due to a different coordination chemistry to both cobalt and rhodium, neither zinc nor nickel are able to coordinate upper ligands and therefore Znbl and Nibl lack an upper ligand completely. They do however contain DMB at the lower ligand position. The various combinations of ligands and the general B_{12} structure can be seen in Figure 5.1.

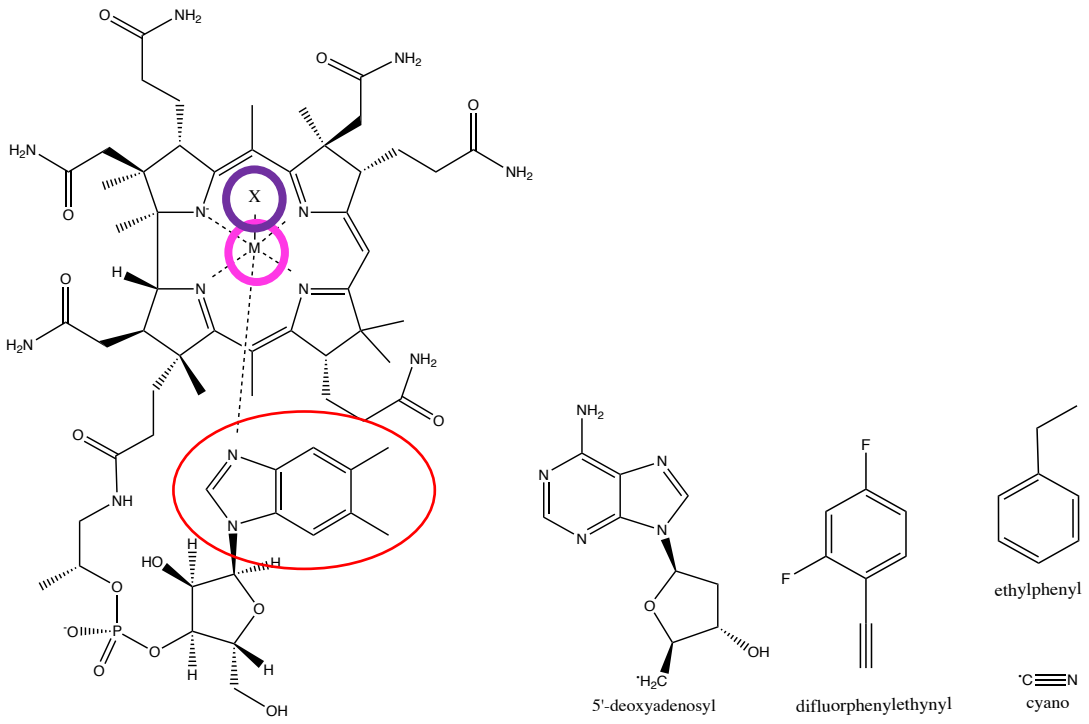


Figure 5.1 General structure of B_{12} alongside the different upper ligand variations found in variants of the molecule. DMB (the lower ligand, shown inside red circle) is not present in the cobinamides (AdoCbi and CNCbi). X (upper ligand) = 5'deoxyadenosyl (AdoCbl, AdoRbl, AdoCbi), CN (CNCbl, CNCbi), ethylphenyl (EtPhCbl), diflurophenylethinyl (F_2PhEtyCbl). M (central metal ion) = cobalt (AdoCbl, CNCbl, AdoCbi, CNCbi, EtPhCbl, F_2PhEtyCbl), rhodium (AdoRbl), zinc (Znbl), nickel (Nibl).

In this study, three key areas of B_{12} are investigated for their importance in riboswitch recognition, the central metal ion, the upper ligand and the lower ligand. Firstly, the central metal ion, which coordinates the upper and lower ligand, was studied. This was done through the use of various different unnatural B_{12} variants, whose antivitamin potentials (whether they bind the same sites as B_{12}, but without carrying out the same function) were also looked into. These include antivitamins with rhodium (AdoRbl), zinc (Znbl), & nickel (Nibl) metal ions, instead of the cobalt

ion found naturally in B_{12}. The second aspect that was looked into was the upper ligand itself. The current research shows a preference for the adenosyl group at the upper ligand position (Nahvi et al, 2002; Gallo et al, 2008). Various upper ligands are tested here (Figure 5.1). Finally, the importance of the lower ligand and the B_{12} base on/ base off coordination was investigated.

Another aim was to develop a switch with the ability to sense B_{12} levels, linked to a fluorescent reporter. This was done using synthetic biology techniques, using an approach similar to that recently reported by Cai *et al*, (Cai et al, 2018). If successful, such a switch could be used in various different biological systems to determine a correlation between B_{12} concentrations and the concentration of a reporter gene. Furthermore, the system can then be trialled in high B_{12}-producing bacterial strains to compare B_{12} production levels.

5.2 Results

5.2.1 Preparation of a riboswitch-controlled reporter gene construct, 'Rib70_*eGFP*_pET14b'

A number of studies on the *btuB* riboswitch have all found that the putative 240 bp *btuB* riboswitch (Figure 5.2) on its own is not enough to control the expression of a reporter gene, suggesting that the riboswitch is non-functional without the beginning of the *btuB* coding region (Franklund and Kadner, 1997; Nou and Kadner, 2000; Lundrigan et al, 1991). These studies show that a large fragment from the coding region of the *btuB* gene is needed along with the non-coding riboswitch.

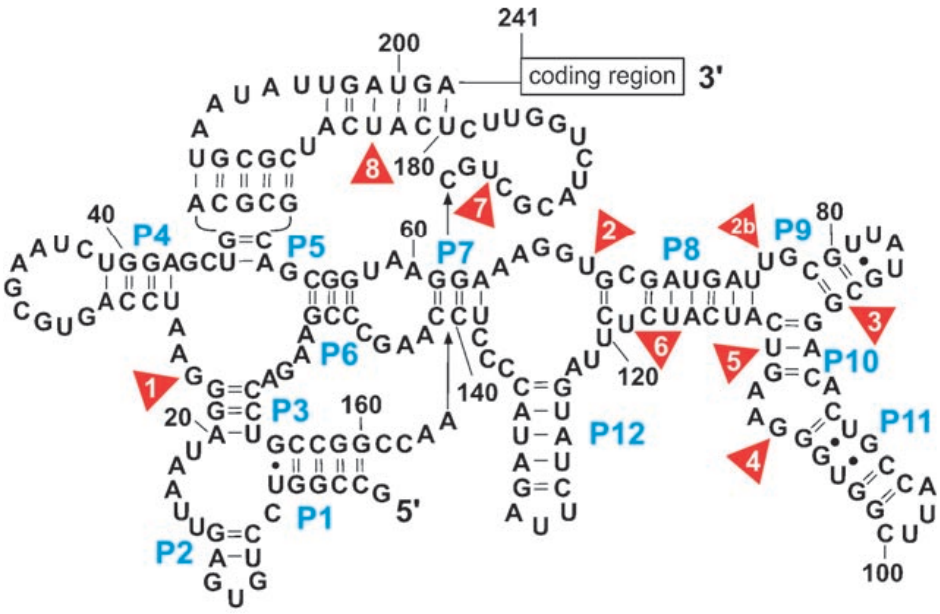


Figure 5.2. The *E. coli btuB* riboswitch sequence. Image adapted from Gallo *et al*, 2008. Cleavage sites, in the absence of AdoCbl, are indicated by red triangles (1-8).

Initially, the *btuB* riboswitch containing the first 70 codons of the *btuB* gene was amplified from the *E. coli* BL21*(DE3) genome. However, upon closer analysis, it appeared that the BL21*(DE3) genome has a natural, in frame, stop codon (TAG), within the first 70 codons of *btuB* (Marisch et al, 2013) due to a C to T mutation at the +38 position. Therefore, *E. coli* K12 MG1655 genomic DNA (without the mutation in the *btuB* gene) was used instead.

PCR was carried out (Chapter 2, Section 2.3.7), using the Rib_For and Rib70_Rev primers (Table 2.3). The PCR-amplified sample was run on a 1% agarose gel, the appropriate ~ 460 bp band was extracted and the DNA was purified from the agarose (Chapter 2, Section 2.3.10), before digesting with XbaI/NdeI restriction enzymes. The use of these particular restriction enzymes allows the fragment to be cloned downstream of the T7 promoter in the pET14b vector, before the start codon of the *eGFP* reporter gene. Therefore, this combination of sites allows the riboswitch to sit correctly in frame between the T7 promoter and the *eGFP* reporter gene. A ligation between the XbaI/NdeI digested riboswitch (insert) and XbaI/NdeI digested *eGFP*-pET14b (vector) DNA was performed. The resultant plasmid (Figure 5.3), called Rib70_*eGFP*_pET14b, was used for the transformation of *E. coli* JM109.

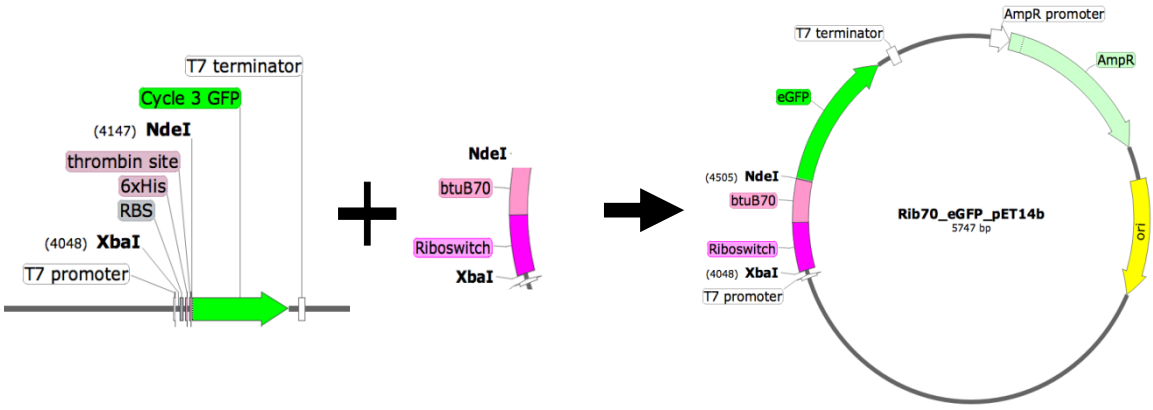


Figure 5.3. Cloning strategy for the construction of the Rib70_*eGFP*_pET14b plasmid. Insertion of the amplified riboswitch fragment (Riboswitch + btuB70) into *eGFP*-pET14b was carried out using XbaI and NdeI restriction sites. Rib70_*eGFP*_pET14b contains Riboswitch-btuB70 upstream of the *eGFP* reporter gene under the control of a T7 promoter and terminator. The plasmid also contains an ampicillin resistance gene, *AmpR* and its promoter, AmpR promoter, for selection purposes.

Overnight LB starter cultures were prepared from the transformed cells. After overnight growth the starter cultures were centrifuged and plasmid DNA was extracted. Samples were sent off to GeneWiz for sequencing, which verified that the riboswitch had been amplified correctly and the Rib70_*eGFP*_pET14b plasmid sequence is as expected.

In order to check whether Rib70_*eGFP*_pET14b would allow production of eGFP, the plasmid was used for the transformation of BL21*(DE3)-pLysS and transformants were cultured in M9 minimal medium. The pellets from the test cultures were found to be green, indicating that eGFP was being produced successfully.

5.2.2 Investigation of riboswitch-controlled *eGFP* expression, in the presence of AdoCbl, using *in vitro* protein synthesis experiments

In order to investigate the control and activity of the riboswitch, in response to increasing concentrations of AdoCbl, *in vitro* protein synthesis experiments were carried out using the NEB PURExpress *in vitro* protein synthesis kit. PURExpress is a cell-free transcription/translation system that is reconstituted from the purified components that are necessary for *E. coli* gene expression. Transcription and translation are carried out in a one-step reaction, and require the mixing of only two tubes (Figure 5.4).

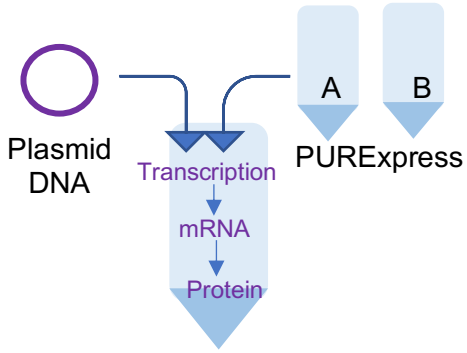


Figure 5.4. Diagram representing the PURExpress one-step reaction required for transcription/translation. Tubes A and B contain the components required for transcription/translation.

Initially, only AdoCbl (obtained from Sigma-Aldrich) was investigated at three different concentrations using the *in vitro* system, (Chapter 2, Section 2.3.15), as a proof of principle to show that the effect of B_{12} on riboswitch-controlled *eGFP* gene expression can be investigated in this manner. Resultant protein production was analysed using SDS-PAGE (Figure 5.5).

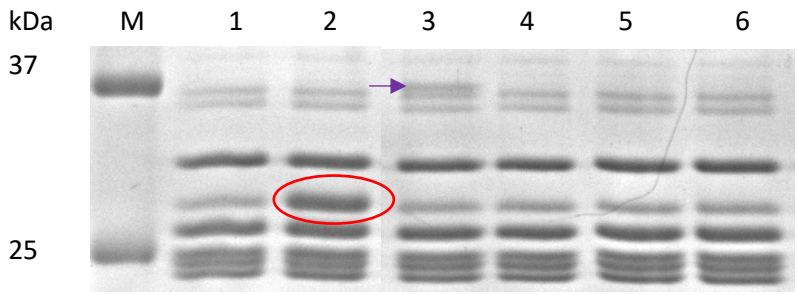


Figure 5.5. SDS-PAGE analysis of *in vitro* reaction samples. All reaction samples were run on a 12.5% SDS gel in order to compare protein bands in each sample. Lanes 1-6 contain different plasmids and plasmid conditions: 1.) pET14b 2.) *eGFP*-pET14b 3.) Rib70_*eGFP*_pET14b 4.) Rib70_*eGFP*_pET14b + 10 μM AdoCbl 5.) Rib70_*eGFP*_pET14b + 25 μM AdoCbl 6.) Rib70_*eGFP*_pET14b + 50 μM AdoCbl. Alongside all the transcription/translation components, two additional proteins can be seen. These are the 28.9 kDa eGFP protein (circled in red) and the 36.2 kDa eGFP fusion protein containing the first 70 codons of btuB (indicated by the purple arrow).

The eGFP protein (28.9 kDa with the His-tag) appears to be present in the *eGFP*-pET14b control (lane 2). This was expected, as this plasmid is the only one capable of making His-tagged eGFP of this size. Lane 3 appears to have an extra band sized around 36 kDa (indicated by the purple arrow). This is the BtuB-eGFP hybrid, consisting of the first 70 amino acids of BtuB attached to eGFP. Its absence in lanes 4-6 shows that the added AdoCbl is behind this reduced protein production. This demonstrates that AdoCbl is successfully switching off the riboswitch in the Rib70-*eGFP*-pET14b plasmid.

In order to verify that the suspected eGFP bands were in fact eGFP, a western blot was carried out using anti-eGFP antibodies (Figure 5.6).

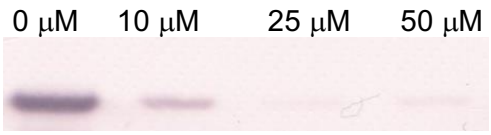


Figure 5.6. Western blot analysis showing the effect of AdoCbl on the *E. coli btuB* riboswitch. Bands represent the eGFP present in the *in vitro* protein synthesis reactions, using the Rib70_*eGFP*_pET14b plasmid in the presence of increasing concentrations of AdoCbl. Decreased levels of eGFP can be seen as AdoCbl concentration is increased from 0 to 50 μM.

The western blot confirms that the riboswitch encoded within the Rib70-*eGFP*-pET14b plasmid is switched off in the presence of AdoCbl, as expected, and that the decrease in production of eGFP is related to the concentration of AdoCbl within the reaction mix.

In order to get an estimate of the levels of eGFP being produced in each reaction, the western blot procedure was repeated with the reactions run alongside a set of eGFP standards of known concentration. eGFP standards, from 0 to 4 μM, were used to generate a standard curve for estimating relative eGFP levels in each of the *in vitro* reaction samples, (Figure 5.7). This was carried out each time a set of reactions was run on a western blot.

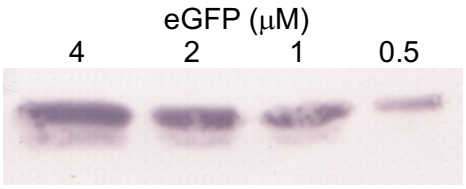


Figure 5.7. Western blot showing eGFP standards of known concentration. Four different concentrations of eGFP were run: 4 μM, 2 μM, 1 μM and 0.5 μM.

Densitometry analysis was carried out on the blot using ImageJ software. The density of the four visible eGFP standard bands was used to prepare a calibration curve (Figure 5.8).

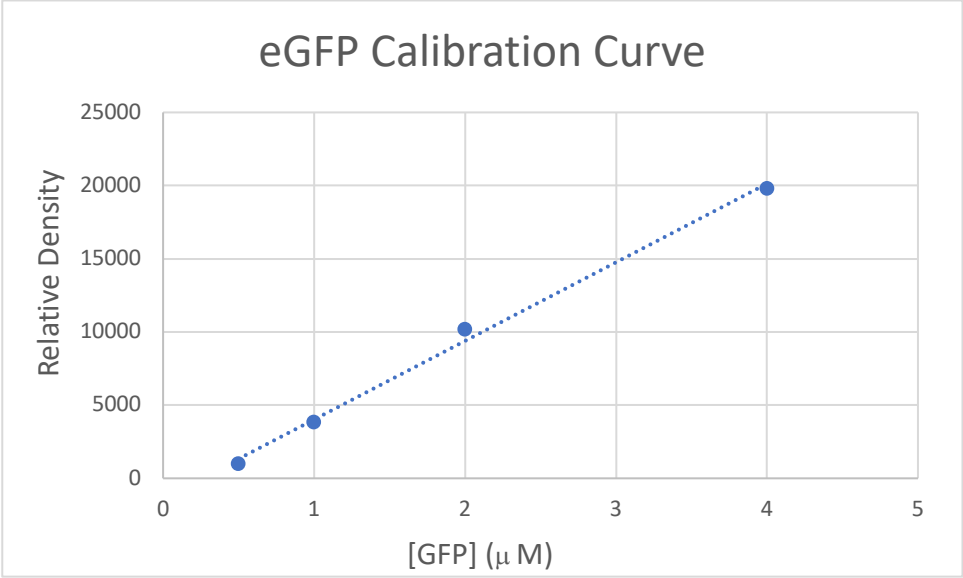


Figure 5.8. eGFP calibration curve. Four different concentrations of eGFP (4 μM, 2 μM, 1 μM and 0.5 μM) were analysed on a western blot and densitometry was carried out for each of the bands, using ImageJ. The densitometry values were used to generate a calibration curve.

Using the equation from the calibration curve, relative eGFP concentrations could be calculated for the relative density values from each of the reaction sample bands. The densitometry results allow a much clearer distinction between the eGFP in the samples, as opposed to comparing band intensity by eye.

Following the initial proof of principle experiments, the effect of AdoCbl on the riboswitch was investigated over a wider concentration range and experiments were carried out in duplicate or triplicate. eGFP concentrations were not only analysed by densitometry but also by measuring fluorescence intensity. However, fluorescence analysis was possible only with certain variants due to fluorescence quenching by some of the B_{12} analogues, as seen later on.

5.2.3 Investigation of the effect of AdoCbl and AdoRbl on the *btuB* riboswitch

In order to further investigate the abilities of AdoCbl binding to and switching off the riboswitch, reactions were performed in duplicate using a larger range of concentrations (0-60 μM). *In vitro* protein synthesis reactions were set up as described in Chapter 2 (Section 2.3.15) and analysed using western blotting, (Figure 5.9).

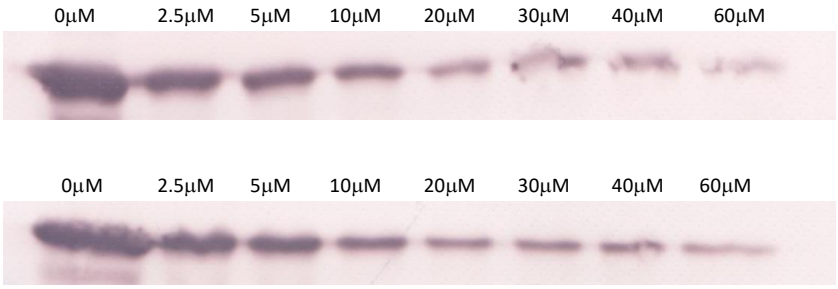


Figure 5.9. Western blot analysis showing the effect of AdoCbl on the *E. coli btuB* riboswitch over a 0-60 μM concentration range. Bands represent the eGFP present in the *in vitro* protein synthesis reactions, using the Rib70_*eGFP*_pET14b plasmid in the presence of 0-60 μM AdoCbl. Decreased levels of eGFP can be seen as AdoCbl concentration is increased from 0 to 60 μM in both repeats.

Using densitometry analysis, band intensities from the *in vitro* reaction samples were compared to those from the eGFP standards of known concentration. This allowed the relative eGFP concentration in each *in vitro* reaction sample to be calculated. eGFP levels were also analysed by measuring the fluorescence

intensities emitted by each *in vitro* reaction sample at 510 nm, after excitation at 488 nm. Since AdoCbl is coloured at high concentrations, it was checked that it did not interfere with eGFP fluorescence. Fluorescence intensities were measured for eGFP samples of known concentration, in the presence of various different AdoCbl concentrations, in order to prepare eGFP calibration curves at each different AdoCbl concentration (Figure 5.10). Very similar readings are seen for fluorescence intensity, with or without AdoCbl, suggesting that AdoCbl has no significant effect on eGFP fluorescence. Calibration curves were used, corresponding to AdoCbl concentration, in order to calculate relative eGFP concentrations in each *in vitro* reaction sample from the fluorescence intensities.

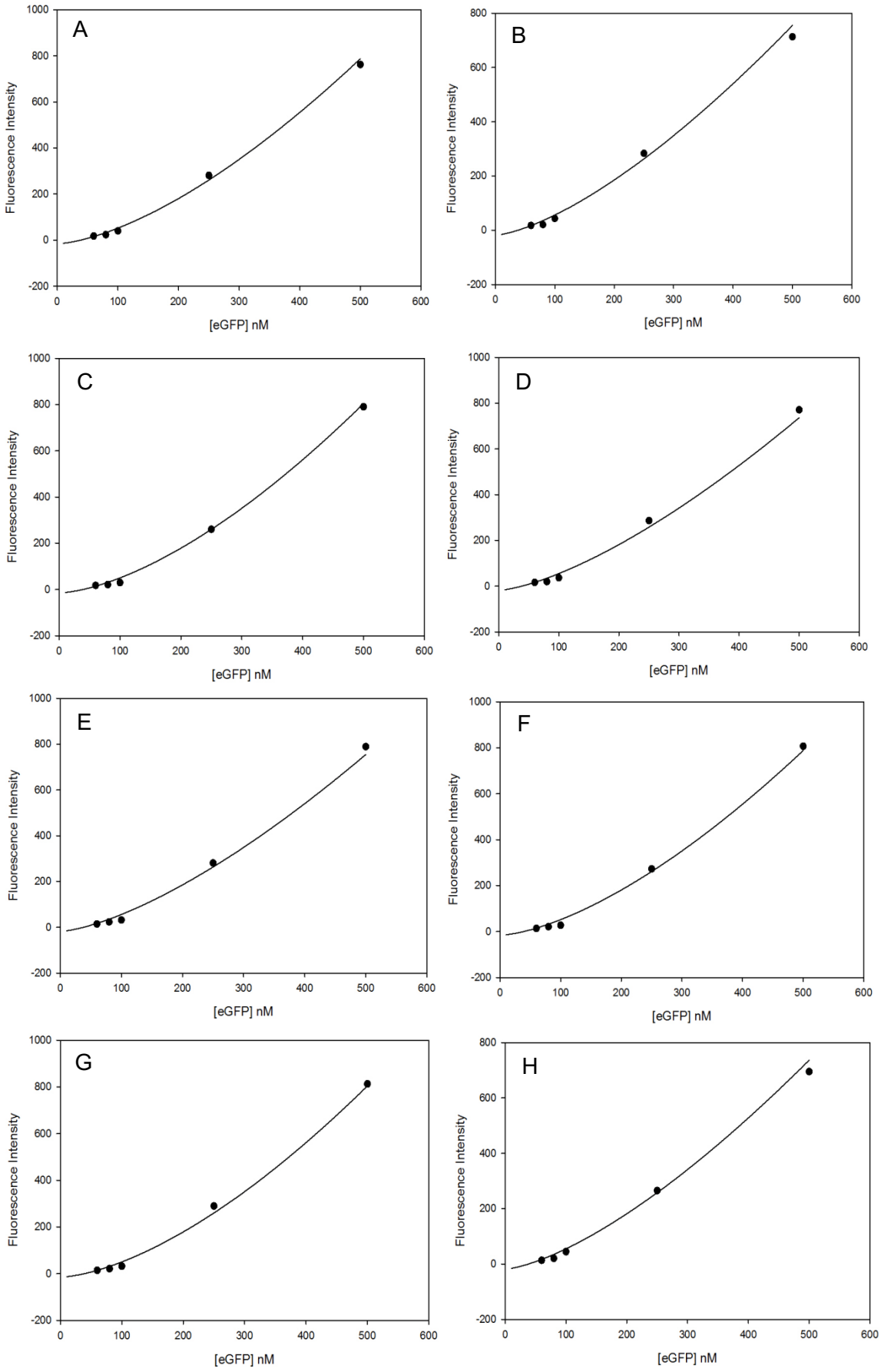

A
B
C
D
E
F
G
H
Fluorescence Intensity
[eGFP] nM

Figure 5.10. Calibration curves for eGFP fluorescence intensity. The fluorescence intensity from eGFP samples of 5 different concentrations were each measured in the presence of 8 different concentrations of AdoCbl, in order to generate calibration curves for each of the different AdoCbl conditions: A) 0 μM B) 0.25 μM C) 0.5 μM D) 1.0 μM E) 2.0 μM F) 3.0 μM G) 4.0 μM H) 6.0 μM

Both densitometry and fluorescence datasets were used to generate graphs showing the effect of increasing AdoCbl concentration on the levels of eGFP produced under *btuB* riboswitch control (Figure 5.11).

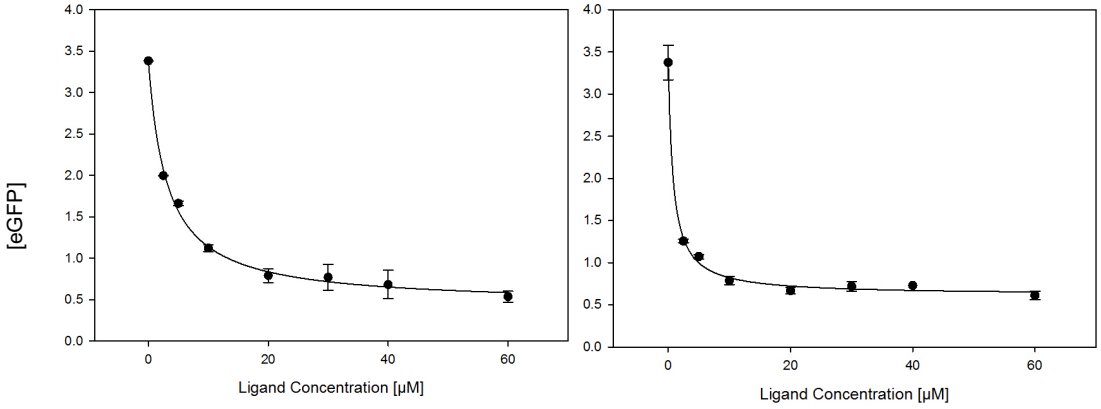


Figure 5.11. Graphs to show the effect of increasing AdoCbl concentrations on riboswitch-controlled eGFP production. Densitometry analysis (left) and fluorescence analysis (right) with estimated K_i values of 3.1 +/- 0.3 μM and 0.8 +/- 0.1 μM, respectively. Note that these K_i values represent 50% of the response, since AdoCbl does not cause 100% activation of the riboswitch.

The results confirm the ability of AdoCbl to bind and switch off the riboswitch, as there is a decline in eGFP production with the increase of AdoCbl concentration, from 2.5 μM to 60 μM.

With these non-linear regression curves (Figure 5.11) an estimated K_i value is generated that is determined from the response curve. However, as it is seen that AdoCbl does not cause 100% activation (switching off) of the riboswitch, then this K_i value does not represent the point at which the switch is 50% off. Since the ligand does not cause 100% activation, the maximal activation that it does allow will be

used to compare its effectiveness as an activator. This maximal activation will be referred to as the A_M value. The best activators are those with the highest A_M values. The fluorescence and densitometry curves suggest AdoCbl activates to a maximal level of 82% (A_M of 82%) and the concentration where half the maximal activation occurs is between 0.8 – 3.1 μM. In this respect, the K_i value (a term that is generally defined as the concentration required for half maximal inhibition) represents the concentration where 50% of the A_M occurs and will henceforth be termed as K_{off} (the switching off constant). Nevertheless, the general trend can still be seen, allowing for a comparison between the different B_{12} variants to see how effective these molecules are at activating the riboswitch.

Subsequent to investigating the binding capabilities of AdoCbl, the effect of the unnatural AdoCbl analogue, AdoRbl, was explored in the same way. Both densitometry and fluorescence analysis were carried out in order to generate graphs representing the effect of AdoRbl on riboswitch-controlled eGFP expression. Results can be seen in Figures 5.12 and 5.13.

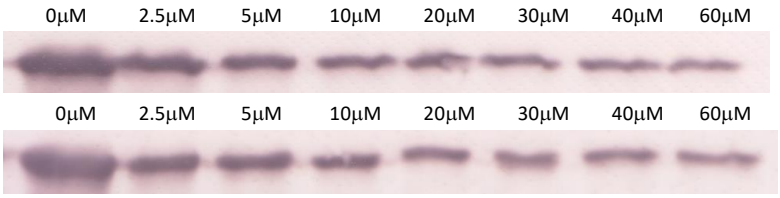


Figure 5.12. Western blot analysis showing the effect of AdoRbl on the *E. coli btuB* riboswitch. Bands represent the eGFP present in the *in vitro* protein synthesis reactions, using the Rib70_*eGFP*_pET14b plasmid in the presence of 0-60 μM AdoRbl. Decreased levels of eGFP can be seen as AdoRbl concentration is increased from 0 to 60 μM in both repeats.

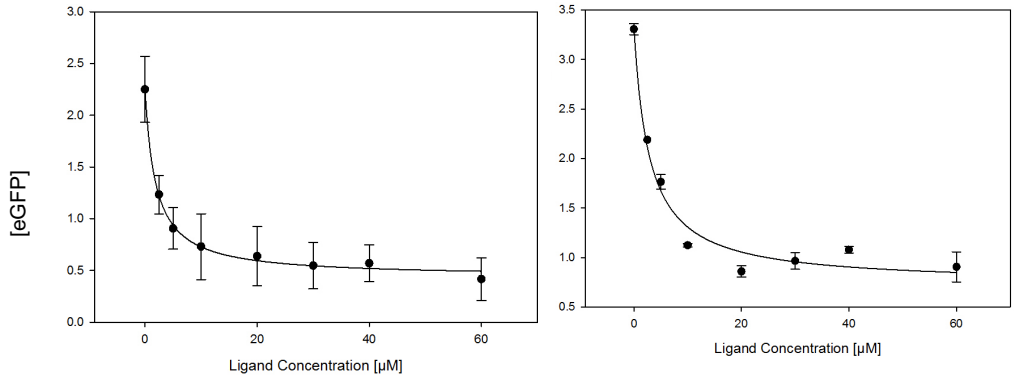


Figure 5.13. Graphs showing the effect of increasing AdoRbl concentration on riboswitch-controlled eGFP production. Densitometry analysis (left) and fluorescence analysis (right) with estimated K_{off} values of 1.9 +/- 0.2 μM and 2.8 +/- 0.7 μM, respectively. Densitometry analysis and fluorescence analysis suggest A_M values of 78% and 73%, respectively.

The results indicate that AdoRbl also binds and switches off the riboswitch, displaying similar effects to that of AdoCbl. The estimated K_{off} values from both methods of analysis are very similar, 2.8 μM and 1.9 μM from the fluorescence and densitometry, respectively. Both suggest tight binding of AdoRbl to the *btuB* riboswitch. This is not surprising as the structures of AdoCbl and AdoRbl are very similar (Widner et al, 2016) .

In order to confirm that the reduction in eGFP production is due to AdoCbl and AdoRbl binding the riboswitch, as opposed to them having an effect on the *in vitro* protein synthesis machinery itself, appropriate controls were carried out. This involved use of the 'positive control' plasmid, *eGFP*-pET14b, which contains no riboswitch and should make eGFP regardless of the presence or absence of B_{12}. The plasmid was expressed using the *in vitro* protein synthesis kit, in the presence of no B_{12}, 30 μM AdoCbl, and 30 μM AdoRbl. eGFP levels were compared using

western blot analysis (Figure 5.14). The results show no visible change in eGFP protein band intensity, in the presence of AdoCbl or AdoRbl, suggesting that neither have an effect on the *in vitro* machinery itself and only change eGFP levels when its expression is under the control of the *btuB* riboswitch.

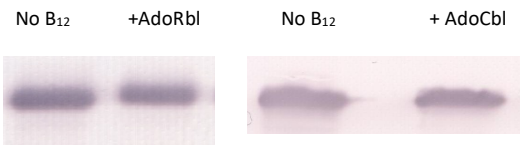


Figure 5.14. Western blot analysis of the eGFP present in the AdoCbl and AdoRbl control reactions. *eGFP*-pET14b plasmid tested using the *in vitro* protein synthesis kit, in the presence of no B_{12}**,** 30 μM AdoRbl and 30 μM AdoCbl. AdoCbl and AdoRbl caused no noticeable change to eGFP levels, in the absence of a riboswitch, therefore having no visible effect on the *in vitro*

5.2.4 Investigation of the effect of CNCbl, CNCbi and AdoCbi on the *btuB* riboswitch

The investigation of two almost identical analogues, AdoCbl and AdoRbl, provided the opportunity to investigate a larger range of corrinoids to learn more about molecular recognition by the riboswitch. Therefore, it was decided to look more specifically into which regions of AdoCbl are required for binding and changing the structure of the riboswitch. Consequently, the next step required looking at a corrinoid with a different upper ligand. CNCbl, with a cyano group at the upper ligand in place of the 5'-deoxyadenosyl group (found in AdoCbl), was ideal.

Before investigating the effect of CNCbl on the riboswitch, a control experiment was performed using *eGFP*-pET14b to check that CNCbl does not have any effect on the *in vitro* machinery. Western blot analysis showed that CNCbl causes no noticeable change to eGFP levels, in the absence of a riboswitch, indicating that CNCbl does not affect the *in vitro* machinery (Figure 5.15).

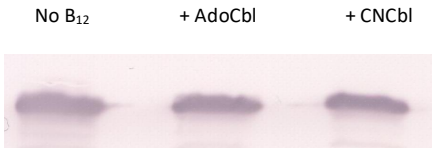


Figure 5.15. Western blot analysis of the eGFP present in the AdoCbl and CNCbl control reactions. The *eGFP*-pET14b plasmid was tested using the *in vitro* protein synthesis kit, in the presence of no B_{12}, 30 μM AdoCbl and 30 μM CNCbl. AdoCbl and CNCbl caused no noticeable change to eGFP levels, in the absence of a riboswitch, therefore having no visible effect on the *in vitro* machinery.

Using the *in vitro* protein synthesis kit, eGFP expression within the Rib70_*eGFP*_pET14b plasmid was carried out in the presence of increasing concentrations of CNCbl (0 to 1000 μM). Experiments were done in triplicate before analysing eGFP production on western blots (Figure 5.16) and by measuring fluorescence intensities from each of the reaction samples. Densitometry analysis was carried out on the western blots and the results from both methods of analysis were compared (Figure 5.17).

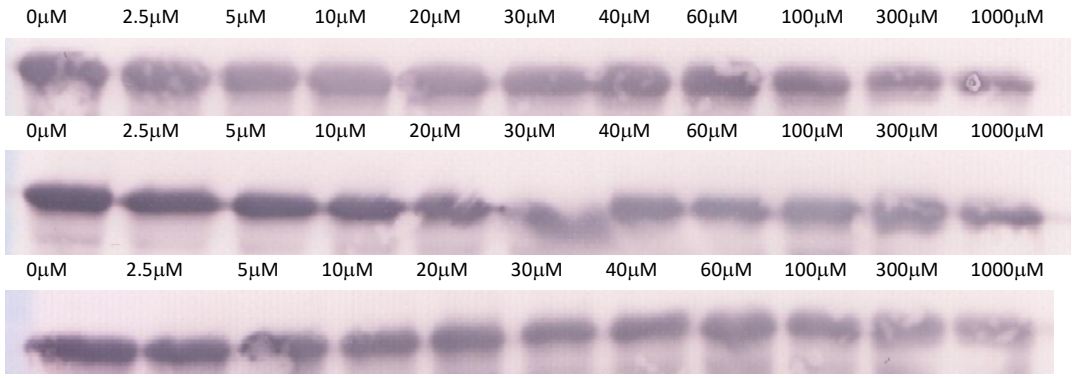


Figure 5.16. Western blot analysis showing the effect of CNCbl on the *E. coli btuB* riboswitch. Bands represent the eGFP present in the *in vitro* protein synthesis reactions, using the Rib70_*eGFP*_pET14b plasmid in the presence of 0-1000 μM CNCbl. Decreased levels of eGFP can be seen as CNCbl concentration is increased in the three repeats.

Prior to fluorescence analysis of the reaction samples, it was important to verify that CNCbl has no effect on the change in fluorescence intensity. Therefore, 25 μM CNCbl was combined with 1 μM eGFP to see if this had any effect on the fluorescence intensity. This was shown to have no visible effect on the fluorescence intensity, suggesting that any eGFP fluorescence analyses, in the presence of CNCbl, can be reliable.

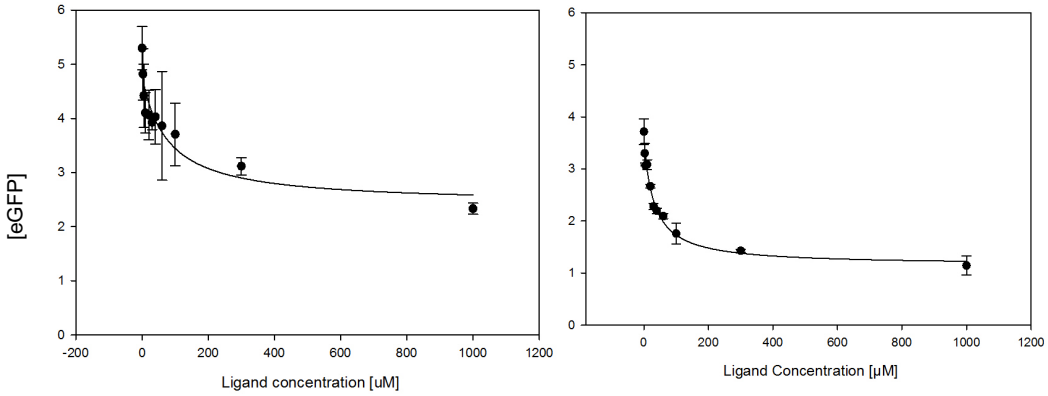


Figure 5.17. Graphs to show the effect of increasing CNCbl concentration on riboswitch-controlled eGFP production. Densitometry analysis (left) and fluorescence analysis (right) with K_{off} values of 77.3 +/- 39.2 μM and 31.4 +/- 5.2 μM respectively. Densitometry analysis and fluorescence analysis suggest A_M values of 54% and 63%, respectively.

Both datasets suggest binding of CNCbl to the *btuB* riboswitch as well as the ability to switch it off, as seen by the gradual decrease in eGFP production with an increase in CNCbl concentration. Each method of analysis differs slightly, with less visible error coming from the fluorescence analysis. This suggests that, in the absence of any fluorescence quenching substrates, fluorescence analysis is the more reliable and preferred method of estimating eGFP levels.

The fluorescence curve suggests CNCbl activates to a maximal level of 63% (A_M of 63%) and that the concentration where half the maximal activation occurs is 31.4 μM. Similarly, the densitometry curve suggests an A_M value of 54% and a concentration of 77.3 μM required for half maximal activation. Therefore, the results indicate that CNCbl is not as strong or productive a ligand as AdoCbl or AdoRbl, due to significantly higher K_{off} values.

A key question to address here is why do ligands such as CNCbl not cause a greater activation of the riboswitch? Why, in this case, is there still a significant amount of protein translation taking place even at very high concentrations of CNCbl? This is likely due to the way in which the molecule binds to the riboswitch, binding largely in a non-productive manner that does not allow complete activation of the switch.

Next, the effects of CNCbi and AdoCbi on the riboswitch, both lacking the cobalamin lower loop, were studied. First, control experiments were carried out in order to confirm that they do not have an effect on the *in vitro* protein synthesis machinery itself. Again, *eGFP*-pET14b plasmid was expressed using the *in vitro* protein synthesis kit, in the presence of 0 μM B_{12}, 30 μM AdoCbi and 30 μM CNCbi. eGFP production was analysed using western blots (Figure 5.18). The results show no visible difference in eGFP protein band intensity in the presence or absence of either Cbi (Figure 5.18).

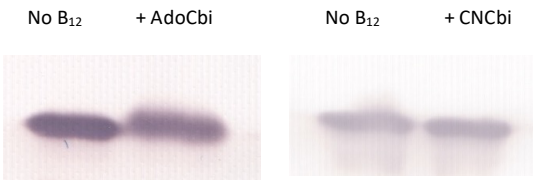


Figure 5.18. Western blot analysis of the eGFP present in the AdoCbi and CNCbi control reactions. The *eGFP*-pET14b plasmid was tested using the *in vitro* protein synthesis kit, in the presence of no B_{12}**,** 30 μM AdoCbi and 30 μM CNCbi. AdoCbi and CNCbi caused no noticeable change to eGFP levels, in the absence of a riboswitch, therefore having no visible effect on the *in vitro* machinery.

When investigating the effects of AdoCbi and CNCbi on the riboswitch, fluorescence analysis could not be used. This was due to quenching of eGFP fluorescence by AdoCbi and CNCbi. Tests were carried out in which a 25 μM sample of each Cbi was combined with 1 μM and 4 μM eGFP to see the effect on fluorescence intensity. Results showed a significantly diminished eGFP fluorescence reading in the presence of each Cbi. Therefore, eGFP production within each *in vitro* reaction sample was analysed using densitometry of bands on western blots (Figures 5.19 and 5.20).

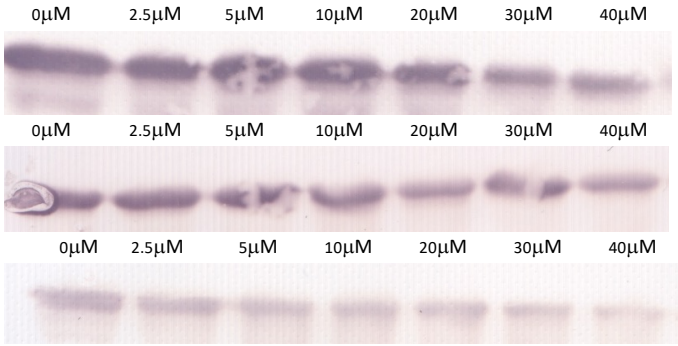


Figure 5.19. Western blot analysis showing the effect of AdoCbi on the *E. coli btuB* riboswitch. Bands represent the eGFP present in the *in vitro* protein synthesis reactions, using the Rib70_*eGFP*_pET14b plasmid in the presence of 0-40 μM AdoCbi. Decreased levels of eGFP can be seen as AdoCbi concentration is increased in all three repeats.

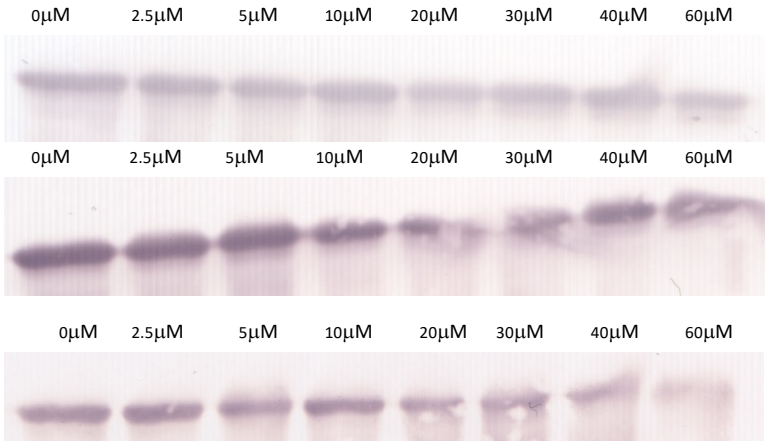


Figure 5.20. Western blot analysis showing the effect of CNCbi on the *E. coli btuB* riboswitch. Bands represent the eGFP present in the *in vitro* protein synthesis reactions, using the Rib70_*eGFP*_pET14b plasmid in the presence of 0-60 μM CNCbi. Decreased levels of eGFP can be seen as CNCbi concentration is increased in all three repeats.

Densitometry analysis was carried out in order to generate curves representing the effect of AdoCbi and CNCbi on riboswitch-controlled *eGFP* expression (Figure 5.21).

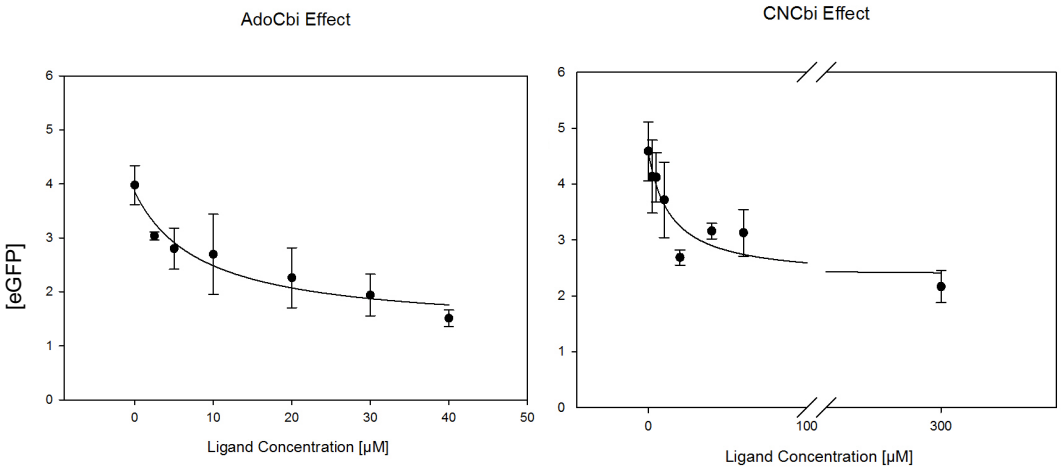


Figure 5.21. Graphs to show the effect of increasing AdoCbi and CNCbi concentrations on riboswitch-controlled eGFP production. eGFP production was analysed using densitometry analysis of eGFP bands on western blots. K_{off} values are 8.6 +/- 4.7 µM and 14.2 +/- 9.6 µM for AdoCbi and CNCbi, respectively. A_M values for AdoCbi and CNCbi are 55% and 48%, respectively.

The fluorescence analysis graphs show that both AdoCbi and CNCbi are capable of binding to the *btuB* riboswitch and causing a change in its structure to hinder gene expression. However, AdoCbi appears to have greater binding capabilities than CNCbi, since it has an A_M value of 55% compared to that of CNCbi (48%). It seems that while AdoCbl and AdoRbl are very good at switching off the riboswitch, these other variants are not so good. The most these other variants appear to be able to do is bind and switch off expression to get approximately 50% of production. As with CNCbl, it could be that AdoCbi and CNCbi are binding to the riboswitch but not binding in a productive way that is able to switch off the riboswitch.

5.2.5 Investigation of the effect of further unnatural cobalamin variants on the *btuB* riboswitch

Other unnatural cobalamin variants (produced by the research group of Professor Bernhard Kräutler) which differ in their upper ligand or completely lack the upper ligand (due to a change to the central metal ion), were tested for their effect on the riboswitch. This was done in order to see whether they could be potential antivitamins that display similar activity to AdoRbl. These variants include ethylphenylcobalamin (EtPhCbl), difluorophenylethinylcobalamin (F_2PhEtyCbl), with ethylphenyl and difluorophenylethinyl upper ligands, respectively, and zincobalamin (Znbl) and nibalamin (Nibl), with zinc and nickel as their central metal ions, respectively. Beforehand, it had been seen that a change from cobalt to rhodium had very little effect, if any, on riboswitch recognition and binding. This time, however, the two new metal ions are unable to coordinate an upper ligand as done so by cobalt and rhodium.

Experiments were carried out in duplicate using the PURExpress *in vitro* protein synthesis kit, followed by eGFP detection on western blots and analysis using densitometry. Due to the quenching of fluorescence by all four unnatural variants, fluorescence analysis could not be used as a method of eGFP quantification.

The bands from the western blots indicate that increasing concentrations of EtPhCbl have no effect on eGFP levels (Figure 5.22), and therefore on the riboswitch. This was confirmed using densitometry analysis (Figure 5.23).

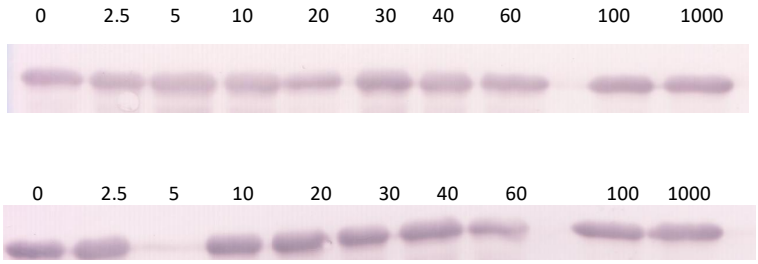


Figure 5.22. Western blot analysis showing the effect of EtPhCbl on the *E. coli btuB* riboswitch. Bands represent the eGFP present in the *in vitro* protein synthesis reactions, using the Rib70_*eGFP*_pET14b plasmid in the presence of 0-1000 μM EtPhCbl. EtPhCbl has no effect on eGFP levels in both repeats.

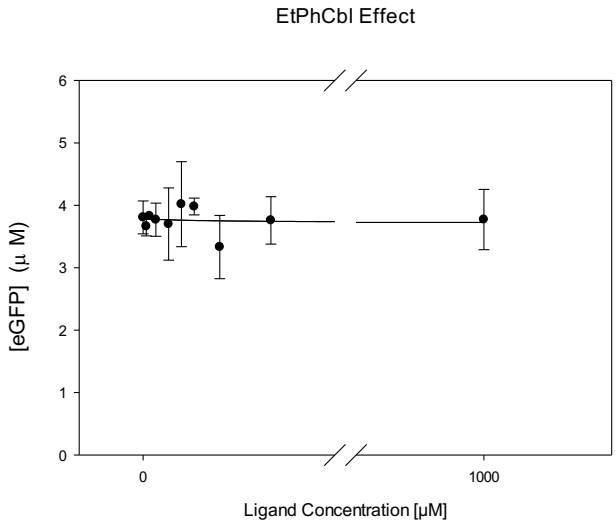


Figure 5.23. Graphs to show the effect of increasing EtPhCbl concentrations on riboswitch-controlled eGFP production. eGFP production was analysed using densitometry of eGFP bands on western blots. EtPhCbl was seen to have no

Next, the effect of F_2PhEtyCbl on the riboswitch was investigated. The western blots showed no clear indication of what effect F_2PhEtyCbl was having on *eGFP* levels (Figure 5.24). Densitometry analysis was therefore carried out (Figure 5.25).

Figure 5.24. Western blot analysis showing the effect of F_2PhEtyCbl on the *E. coli btuB* riboswitch. Bands represent the eGFP present in the *in vitro* protein synthesis reactions, using the Rib70_*eGFP*_pET14b plasmid in the presence of 0 - 60 μM F_2PhEtyCbl as well as at two higher concentrations of 100 μM and 1000 μM.

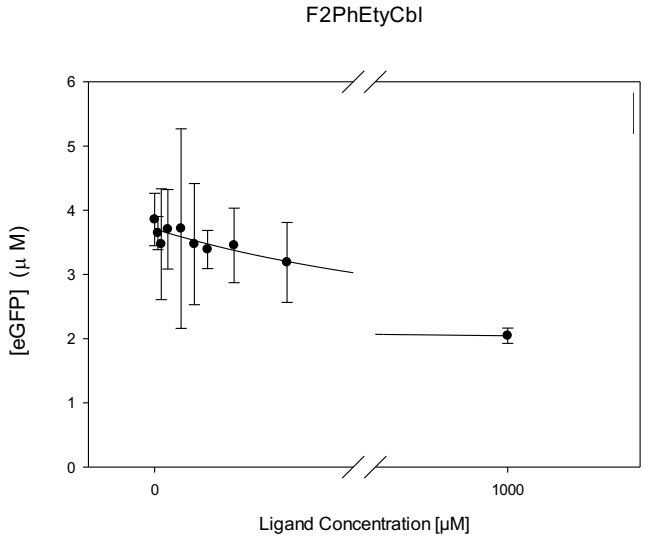


Figure 5.25. Graphs to show the effect of increasing F_2PhEtyCbl concentrations on riboswitch-controlled eGFP production. eGFP production was analysed using densitometry of eGFP bands on western blots. No effect was seen for F_2PhEtyCbl between 0 and 60 μM. However, reduced levels are seen at 1000 μM.

Results from the densitometry analysis suggest that F_2PhEtyCbl could have an effect on *eGFP* expression at higher concentrations, such as 1 mM. However, a broader range of concentrations would need to be tested within the 100 μM to 1 mM range in order to verify this. Additionally, to confirm F_2PhEtyCbl was not having an effect on the *in vitro* protein synthesis machinery, control experiments were carried out (Figure 5.26). These confirmed that F_2PhEtyCbl did not have an effect on the *in vitro* protein synthesis machinery.

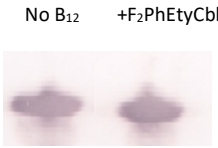


Figure 5.26. Western blot analysis of the eGFP present in the F_2PhEtyCbl control reactions. The *eGFP*-pET14b plasmid was tested using the *in vitro* protein synthesis kit, in the presence of no B_{12} and 30 μM F_2PhEtyCbl. F_2PhEtyCbl caused no noticeable change to eGFP levels in the absence of a riboswitch, therefore having no visible effect on the *in vitro* machinery**.**

Finally, the effects of Znbl and Nibl were investigated. Results from the western blots (Figures 5.27 and 5.28) and densitometry analyses (Figure 5.29) showed that the variants have no effect on the riboswitch.

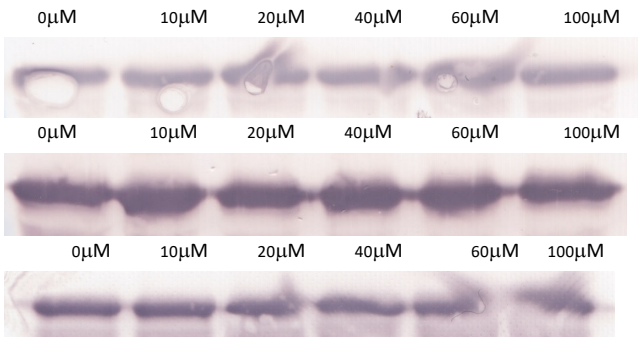


Figure 5.27. Western blot analysis showing the effect of ZnbI on the *E. coli btuB* riboswitch. Bands represent the eGFP present in the *in vitro* protein synthesis reactions, using the Rib70_*eGFP*_pET14b plasmid in the presence of 0-100 μM ZnbI. ZnbI has no effect on eGFP levels in all three repeats.

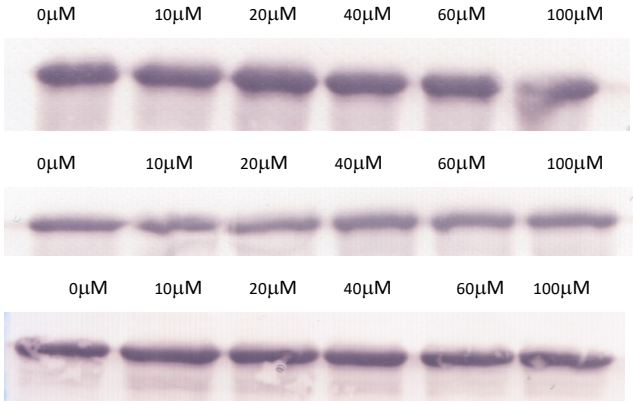


Figure 5.28. Western blot analysis showing the effect of NibI on the *E. coli btuB* riboswitch. Bands represent the eGFP present in the *in vitro* protein synthesis reactions, using the Rib70_*eGFP*_pET14b plasmid in the presence of 0-100 μM NibI. NibI has no effect on eGFP levels in all three repeats.

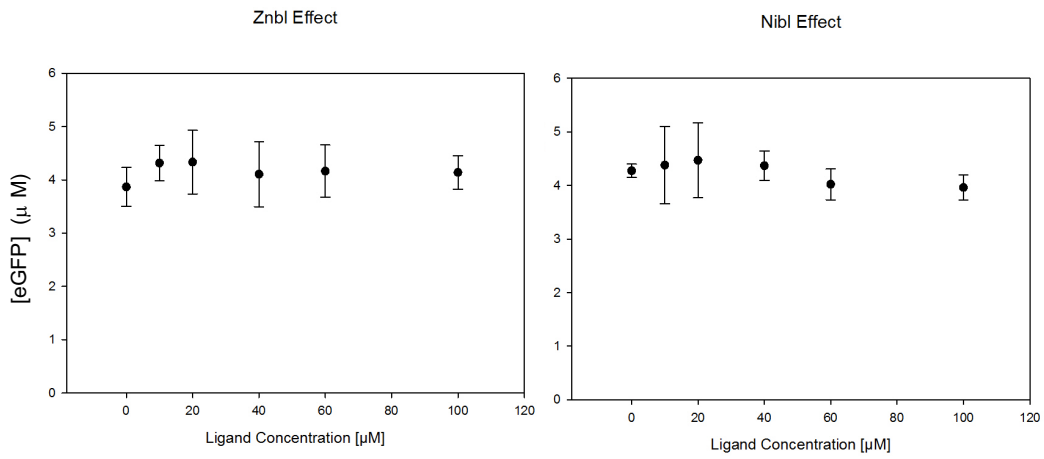


Figure 5.29. Graphs to show the effect of increasing Znbl and Nibl concentrations on riboswitch-controlled eGFP production. eGFP production was analysed using densitometry analysis of eGFP bands on western blots. No effect can be seen from either ligand on eGFP production.

In both datasets (Figure 5.29) there appears to be no direct effect of increasing ligand concentration on the eGFP concentration detected at the end of each reaction. Results show that neither of the analogues have an effect on the riboswitch. This is most likely due to the fact that both of these B_{12} variants lack an upper ligand. Unlike rhodium, which has a very similar coordination chemistry to cobalt, the zinc and nickel ions in these two compounds are unable to coordinate an upper ligand. Therefore, the most likely explanation for their inability to induce riboswitch recognition and binding is their lack of the crucial upper ligand group.

5.2.6 Investigating the potential of reversing the riboswitch control over *eGFP* reporter gene expression

Following on from the work done on the *E. coli btuB* riboswitch, where a system was developed that allows a decline in eGFP production to be seen with the addition of known concentrations of $B_{12,}$ a new system was developed here that involves reversing the current workings similar to that done by Cai *et al*. In this system (the reverse riboswitch system), the *E. coli btuB* riboswitch (including the first 70 codons of *btuB*) is placed directly in front of the *tetR* gene such that B_{12} is able to switch off *tetR* expression and control production of TetR. The riboswitch-tetR fragment is under the control of the P_{J23110} promoter (medium strength constitutive promoter designed by John Anderson at the iGem group in Berkeley, 2006). Meanwhile, *eGFP* is placed under the control of the P_{TetA} promoter which is controlled by TetR due to its ability to switch off any expression reliant on the P_{TetA} promoter. Thus, B_{12} indirectly controls the production of eGFP through control of levels of TetR (Figure 5.30).

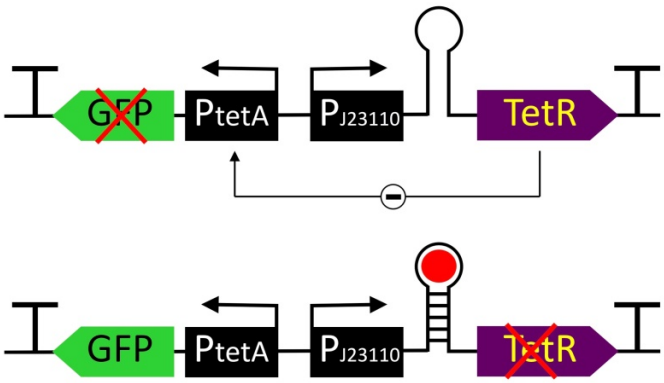


Figure 5.30. Diagrammatic representation of the reverse riboswitch system in the presence and absence of B_{12} (red circle). In the absence of B_{12} (above), the riboswitch is switched on and *tetR* is expressed. The TetR protein is then able to negatively control *eGFP* expression by binding to the P_{TetA} promoter. This should lead to a lack of eGFP production. In the presence of B_{12} (below), TetR is not produced and therefore eGFP production is unhindered. Image Source: Dr.

The system was designed with the addition of strategically placed restriction sites around each component (P_{J23110}, *btuB* riboswitch, *tetR* gene, P_{TetA} promoter, *eGFP* gene), such that the components can be modified at any point. The sequence was then sent to Eurofins so that the 2235 bp DNA fragment (NK DNA) could be synthesised. The synthesised DNA was received in the form of a plasmid, NK_DNA_pEX-A258, containing an ampicillin resistance gene for selection purposes.

In order to test the functionality of the plasmid, transformation of JM109 was carried out with NK_DNA_pEX-A258 and used for the preparation of overnight starter cultures, with and without B_{12}, to see whether this caused a difference in eGFP production. From the presence of green-coloured bacterial pellets, it was seen that all four cultures (no B_{12}, 30 μM AdoCbl, 30 μM CNCbl and 30 μM OHCbl) were

producing eGFP. The test was repeated using minimal medium with the addition of 1 mM thiamine, instead of LB, in order to prevent contamination of B_{12} in any samples (particularly the sample that should have no B_{12}). However, the results were the same and the use of minimal medium instead of LB made no difference. This suggests that, even in the absence of B_{12}, the levels of TetR are too low, for which reason TetR is unable to switch off eGFP production. In order to confirm this, TetR production would need to be looked into further in each of the various conditions.

Before investigating TetR production, a new plasmid (called NK_DNA2_pEX-A258) was constructed in which the promoter upstream of the riboswitch was changed from P_{J23110} to P_{TetR} (also a constitutive promoter). This allows TetR production levels, under the control of two different promoters, to be compared to see which is preferred in this particular application, and to see if there is a difference. To change the promoter, the NK_DNA_pEX-A258 plasmid was digested using the StuI restriction enzyme to cut out the fragment containing P_{J23110} and replace it with P_{TetR} (Figure 5.31).

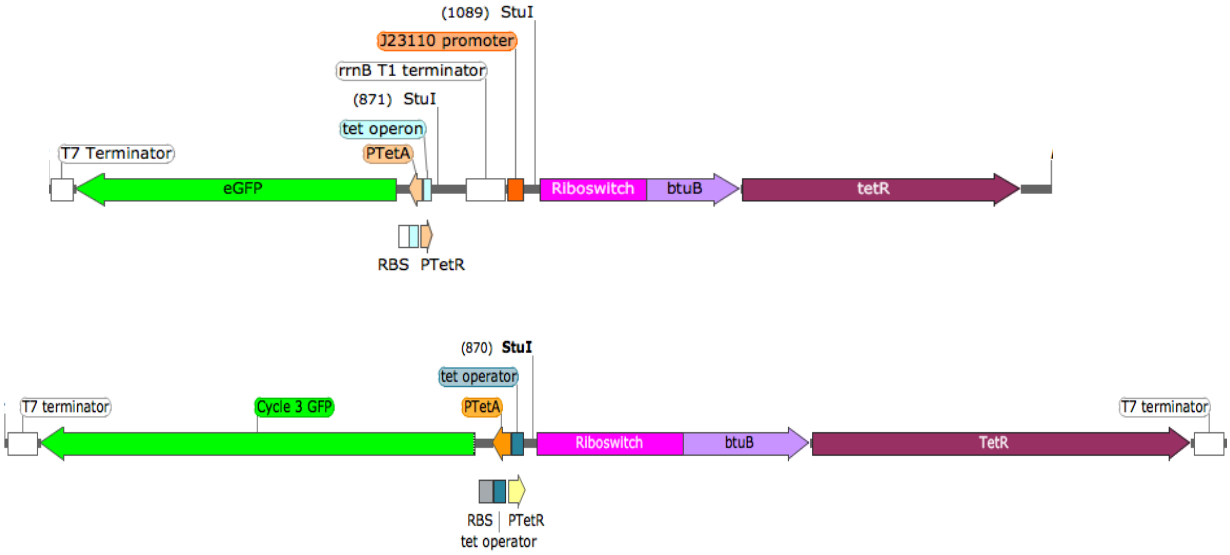


Figure 5.31. Diagram to compare the reverse riboswitch system in NK_DNA_pEX-A258 (above) and in NK_DNA2_pEX-A258 (below). NK_DNA_pEX-A258 has *tetR* under the control of P_{J23110}, whereas NK_DNA2_pEX-A258 has *tetR* under the control of P_{TetR}.

New assays were set up in minimal medium using transformants carrying one of each of the two plasmid variants. Each transformant was grown up in two different conditions, no B_{12} and 30 μM AdoCbl, with each reaction carried out in duplicate. Following overnight growth, samples were prepared as described in Chapter 2 (Section 2.3.16) and eGFP levels analysed using SDS-PAGE (Figure 5.32).

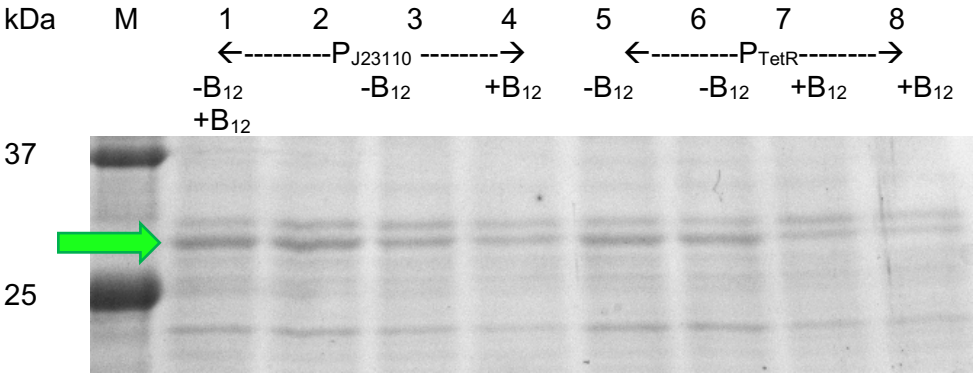


Figure 5.32. SDS-PAGE analysis comparing eGFP levels within JM109-NK_DNA_pEX-A258 and JM109-NK_DNA2_pEX-A258 in the presence and absence of B_{12}. Lanes 1-4; JM109-NK_DNA_pEX-A258 with the P_{J23110} promoter controlling *tetR*, lanes 5-8; JM109-NK_DNA2_pEX-A258 with the P_{TetR} promoter controlling *tetR*. Lanes 3, 4, 7 and 8 contain samples from cultures prepared with the addition of 30 μM AdoCbl. eGFP (26.8 kDa) appears to be present in all samples with no visible difference in results between the two bacterial strains.

The SDS-PAGE analysis (Figure 5.32) suggests that eGFP (26.8 kDa) is present in all the samples. In order to confirm that these bands represent eGFP, a western blot was carried out using antibodies against eGFP (Figure 5.33).

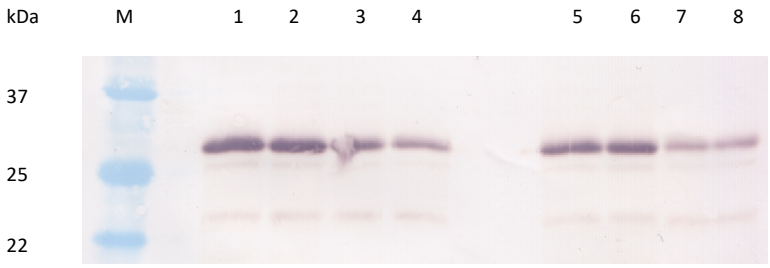


Figure 5.33. Western blot analysis comparing eGFP levels within JM109-NK_DNA_pEX-A258 and JM109-NK_DNA2_pEX-A258 in the presence and absence of B_{12}. Lanes 1-4; JM109-NK_DNA_pEX-A258 with the P_{J23110} promoter controlling *tetR*, lanes 5-8; JM109-NK_DNA2_pEX-A258 with the P_{TetR} promoter controlling *tetR*. Lanes 3, 4, 7 and 8 contain samples from cultures prepared with the addition of 30 μM AdoCbl. eGFP (26.8 kDa) appears to be present in all samples with no visible difference in results between the two bacterial strains.

The western blot confirms the presence of eGFP in each of the samples. However, the addition of B_{12} does not appear to increase the levels of eGFP as expected in the reverse riboswitch system. The presence of B_{12} is ideally expected to switch off *tetR* expression and in turn lead to maximum levels of eGFP production (Figure 5.30). It may be that, even in the absence of B_{12}, TetR levels are too low to have any effect on the P_{TetA} promoter (controlling *eGFP* expression). Alternatively, it may be that the hybrid form of TetR (TetR fused with the first 70 amino acids of BtuB) is unable to dimerise like regular TetR (Orth et al, 2000). This hybrid form of TetR may not then be able to bind the P_{TetA} promoter. Additionally, in the case that TetR-BtuB does dimerise, the 70 amino acids of BtuB at the N-terminus of TetR could be interfering with TetR binding to DNA, since it is the N-terminus of TetR that is involved in DNA binding (Orth et al, 2000). In order to better understand what is happening, TetR levels in each condition would need to be analysed closely and

compared to eGFP levels. Future work could also involve moving on to the use of an alternate repressor system, such as the lac repressor system.

5.3 Discussion

Numerous studies on the *E. coli btuB* riboswitch have determined that there is a requirement for the 5'-deoxyadenosyl group, at the upper ligand position of B_{12}, in order to bind the B_{12}-binding pocket of the riboswitch and allow structural modulation of the RNA. Many of these studies have used in-line probing to show that AdoCbl is the preferred ligand for the *btuB* riboswitch (Nahvi et al, 2002; Nou and Kadner, 2000).

In this thesis, the effectiveness of various different B_{12} analogues, both natural and unnatural, in binding to and switching off the *E. coli btuB* riboswitch has been investigated using *in vitro* assays. This involved the use of an *in vitro* transcription/translation kit, since it was thought that this might allow more accurate determination of the ability of the riboswitch to respond to B_{12} concentrations and different B_{12} types. eGFP reporter concentrations were measured using both fluorescence analysis and densitometry of western blot bands. However, fluorescence analysis could not be done for all the experiments since certain B_{12} variants caused quenching of eGFP fluorescence.

Fluorescence analysis showed decreasing levels of eGFP production in the presence of increasing concentrations of AdoCbl, AdoRbl, CNCbl, AdoCbi and CNCbi. Densitometry analysis provided very similar results, but with slightly larger scope for error. Hence it can be assumed that each of these B_{12} variants are able to bind to the *btuB* riboswitch and cause a change in its structure, such that the RBS gets sequestered and protein translation is hindered.

As seen from the response curves, the ligands do not appear to have the ability to cause 100% activation (switching off) of the riboswitch. Therefore, A_M (maximal activation) and K_{off} values (the concentration where 50% of the A_M occurs) were used to compare how effective the different B_{12} variants are at activating the riboswitch. A summary of these values can be found in Table 5.1. The most effective ligands are those with the highest A_M values and lowest K_{off} values.

Table 5.1. Table to compare K_{off} and A_M values for each ligand. K_{off} and A_M values were obtained from fluorescence and densitometry analysis for AdoCbl, AdoRbl and CNCbl, but only densitometry analysis was used for AdoCbi and

A comparison between different B_{12} variants on their effect on the *E. coli btuB* riboswitch				
Cobalamin type	**Method of Analysis**	**K_{off} (μM)**	**Standard Error for K_{off} value**	**A_M (%)**
AdoCbl	Fluorescence	0.8	+/- 0.1	82
AdoCbl	Densitometry	3.1	+/- 0.3	82
AdoRbl	Fluorescence	2.8	+/- 0.7	73
AdoRbl	Densitometry	1.9	+/- 0.2	78
CNCbl	Fluorescence	31.4	+/- 5.2	63
CNCbl	Densitometry	77.3	+/- 39.2	54
AdoCbi	Densitometry	8.6	+/- 4.7	55
CNCbi	Densitometry	14.2	+/- 9.6	48

The A_M and K_{off} values for each riboswitch binding ligand, as summarised in Table 5.1, show the most effective at binding and switching off the riboswitch to be AdoCbl, with an A_M value of 82%, followed by its analogue AdoRbl, with an A_M value of 73-78%. This means that both AdoCbl and AdoRbl have the ability to cause

around 80% activation of the riboswitch, thereby causing a reduction in protein synthesis by 80%.

This supports the well-known theory that AdoCbl is the preferred ligand for the *E. coli btuB* riboswitch. Previous studies, using primer extension-inhibition assays, have even estimated a K_i of AdoCbl at 0.3 μM (Nou and Kadner, 2000). Following this, further studies using in-line probing yielded apparent K_D values of 0.3 μM for AdoCbl binding to the *btuB* riboswitch (Nahvi et al, 2002). Although K_i values in this study were not representative of half the concentration required for full inhibition (since full switching off could not be observed), K_{off} values were used instead to represent half the maximal activation of the riboswitch. These allowed an effective comparison between B_{12} variants and meant that the activating potential of analogues or antivitamins could be compared. For example, the antivitamin AdoRbl had a K_{off} value of between 1.9 and 2.8 μM which was very close to that of the preferred ligand AdoCbl, with a K_{off} value of 0.8 to 3.1 μM. This suggests that AdoRbl has high potential to be used an AdoCbl inhibitor as opposed to the remaining antivitamins, which were seen to have no effect on the riboswitch.

The natural cobalt ion has an important role in coordination of the upper and lower ligand. In AdoRbl, cobalt is replaced with rhodium. Since cobalt and rhodium share a similar coordination chemistry, rhodium is able to hold the same groups as cobalt to form AdoRbl (Widner et al, 2016). The work presented in this thesis has shown that the change from cobalt to rhodium has little effect on riboswitch recognition and binding. However, when testing two other analogues with unnatural central metal ions, the difference in results was dramatic. Nibl and Znbl were both seen to have no effect on the riboswitch despite the use of larger concentrations. The reason for this is most likely to be the differing coordination chemistry to cobalt, for which reason nickel and zinc are only able to form five bonds and cannot coordinate an

upper ligand. This again highlights the importance of the upper ligand, as opposed to the central metal ion itself, in riboswitch recognition and binding.

In summary, the work described in this report confirms the requirement for the 5'-deoxyadenosyl group in the control of the *E. coli btuB* riboswitch. The results also confirm the ability of the unnatural analogue, AdoRbl, to mimic AdoCbl and cause a similar effect on the riboswitch, suggesting that it has the potential to control BtuB protein production and therefore cobalamin uptake. In terms of mimicking the functions of AdoCbl, AdoRbl would be unable to do so due to its lack of the crucial hydrolysable C-Co bond that allows AdoCbl to function as a coenzyme. This means that AdoRbl would be recognised by cells as AdoCbl, but B_{12}-dependent enzymes would remain inactive, thus giving AdoRbl the potential to put cells in a B_{12}-starved state and inhibit further growth. Thus, AdoRbl can be said to have antivitamin potential (binding the same sites as B_{12}, but without carrying out the same function). Through further research, such antivitamins could potentially be used for the development of various medicinal applications or other applications in the field of synthetic biology.

Chapter 6: Discussion

6.1 General discussion

The fascinating world of B_{12} research began in the early 1920's, through the work of three researchers, Whipple, Minot and Murphy, who demonstrated that they were able to cure pernicious anaemia with the addition of whole liver in the diet (Whipple et al, 1920). This then led to the discovery of the "anti-pernicious anaemia factor", vitamin B_{12}, whose three-dimensional structure was later deduced by Dorothy Hodgkin in 1955 (Hodgkin et al, 1955). Following on from this, the work of several different research groups identified multiple B_{12}-dependent enzymes, in all kingdoms of life. Such studies have helped generate a greater understanding of the role of the molecule in metabolism and the consequences of its absence.

Subsequently, years of research has built up an extensive idea of how this structure is synthesised from uroporphyrinogen III, through a series of around 30 complex enzyme-mediated steps, in bacteria, to form nature's most complex vitamin (Warren et al, 2002). Not only has the biological synthesis been elucidated, but the total chemical synthesis of B_{12} has also been achieved, an outstanding achievement that was orchestrated by Woodward and Eschenmoser. The chemical synthesis involved the dedicated work of over 100 scientists, during a period of 11 years between the 1960s and 1970s. Utilising the knowledge attained from both the biosynthesis and chemical synthesis of B_{12} has allowed current scientists to create a range of different B_{12} variants, both natural and unnatural.

The work carried out here investigates the possibility of engineering the production of B_{12} analogues, through the construction of an empty, metal-free B_{12} shell, and exploring what effect such compounds might have on cellular control elements, such as riboswitches. In 1965, metal-free corrinoids were first discovered and isolated from the photosynthetic bacterium, *A. vinosum* (Toohey, 1965), and hence, enzymes from this bacterium were used in this project to construct an empty, metal-

free B_{12} shell to be used as the starting point for creating new B_{12} analogues. Unfortunately, this approach did not work, most likely due to the inactivity of one or more of the *A. vinosum* enzymes. Therefore, analogues provided by the Krautler group were used in investigating the effect of such compounds on a B_{12}-dependent riboswitch.

In order to study this, a method was designed and developed that utilises an *in vitro* transcription/translation system to determine the effect of both natural and unnatural B_{12} analogues on the well-characterised B_{12}-dependent riboswitch, known as the *btuB* riboswitch. This method allowed a quick and easy way of testing the ability of B_{12} variants to directly bind to and deactivate the *btuB* riboswitch, without relying on indirect cellular effects. eGFP production was monitored in the presence of AdoCbl, AdoRbl, CNCbl, AdoCbi, CNCbi, Nibl, Znbl, EtPhCbl and F_2PhEtyCbl. The results showed that AdoCbl and AdoRbl had the greatest effect on the riboswitch, with AdoRbl closely mimicking the levels of inhibition caused by AdoCbl. CNCbl, AdoCbi and CNCbi were seen to have much lesser effects and required higher concentrations to cause similar effects to that of AdoCbl or AdoRbl. The remaining four unnatural B_{12} variants were seen to have little effect, if any, on the *E. coli btuB* riboswitch.

The work described in this thesis confirms that an adenosyl group in place of the upper ligand is crucial to control of the *E. coli btuB* riboswitch. The demonstration that AdoRbl is able to mimic AdoCbl and bind to the riboswitch highlights the potential for such analogues, with the required components for recognition, to interfere with and inhibit cellular processes. Since AdoRbl binds to the riboswitch in the same way as AdoCbl, it is also likely to bind any B_{12}-dependent enzymes that recognise AdoCbl. In terms of mimicking the functions of AdoCbl as a coenzyme, AdoRbl would be unable to do so due to the lack of the crucial hydrolysable C-Co

bond that allows AdoCbl to participate in chemical reactions. This has been shown in recent work by Widner *et al* where the provision of AdoRbl to B_{12}-dependent cells caused thinner, more spread out growth, suggesting that AdoRbl is binding the *btuB* riboswitch and causing reduced levels of *btuB* gene expression and therefore reduced B_{12} uptake (Widner et al, 2016). However, since the AdoRbl was unable to work with B_{12}-dependent enzymes, growth was less dense compared to when using CNCbl (Widner et al, 2016). The work presented in this thesis has confirmed that AdoRbl does indeed bind to the *btuB* riboswitch to reduce the levels of *btuB* gene expression. Together, this all demonstrates that analogues such as AdoRbl have the potential to disrupt biological systems, by inactivating B_{12}-dependent enzymes and putting cells in a B_{12}-starved state, opening up a whole new area of research into the use of B_{12} analogues as antibacterial agents.

Researchers have already begun exploiting the fundamental knowledge acquired on riboswitches by designing riboswitches with medicinal and biotechnological applications. One such example involves the recent synthesis of functional riboswitches that are able to detect explosives, due to the design of an aptamer domain that is able to bind a component known as 2,4-dinitrotoluene (2,4-DNT), commonly found in explosives (Borujeni et al, 2015).

In other applications, riboswitches are being utilised as powerful RNA imaging tools to elucidate mechanisms such as those underlying the dynamics of the mRNA and ncRNA life cycle. Here, an RNA tag has been developed containing a B_{12}-riboswitch as well as Cbl-fluorophore probes, in which the fluorescence of a synthetic fluorophore is quenched by Cbl until Cbl binds to the B_{12}-riboswitch, found in the RNA tag, and subsequently leads to the production of fluorescence (Braselmann et al, 2018).

In addition to this, riboswitches controlling key metabolic systems in pathogenic organisms can be exploited through the use of metabolite/ligand analogues for the development of antibacterial compounds (Blount and Breaker, 2006). These advances suggest that the continued discovery and analysis of riboswitches will allow progress in creating more useful molecular tools as well as help in the identification of drug targets. The work presented in this thesis has shown the ability of AdoRbl to work as an effective B_{12} inhibitor in the case of one particular B_{12}-dependent riboswitch, the *btuB* riboswitch. However, by investigating a whole other range of riboswitches, it could be that other riboswitches require different components for recognition, instead of the upper ligand, as in the case of the *btuB* riboswitch. It may be that different riboswitches prefer analogues such as Znbl or Nibl, without upper ligands, over AdoRbl. Thus, the research carried out in this thesis opens up a fascinating new field of research into the use of B_{12} analogues as antibacterial agents.

References

ABEND, A., BANDARIAN, V., NITSCHE, R., STUPPERICH, E., RÉTEY, J. and REED, G.H., 1999. Ethanolamine ammonia-lyase has a “base-on” binding mode for coenzyme B12. *Archives of Biochemistry and Biophysics,* **370**(1), p. 138-141.

ANDRÈS, E., LOUKILI, N.H., NOEL, E., KALTENBACH, G., ABDELGHENI, M.B., PERRIN, A.E., NOBLET-DICK, M., MALOISEL, F., SCHLIENGER, J. and BLICKLÉ, J., 2004. Vitamin B12 (cobalamin) deficiency in elderly patients. *Canadian Medical Association journal,* **171**(3), p. 251-259.

BANERJEE, R., 1999. *Chemistry and Biochemistry of B12.* John Wiley & Sons.

BANERJEE, R., GHERASIM, C. and PADOVANI, D., 2009. The tinker, tailor, soldier in intracellular B12 trafficking. *Current opinion in Chemical Biology,* **13**(4), p. 484-491.

BANERJEE, R. and RAGSDALE, S.W., 2003. The many faces of vitamin B12: catalysis by cobalamin-dependent enzymes. *Annual Review of Biochemistry,* **72**(1), p. 209-247.

BATEY, R.T., 2012. Structure and mechanism of purine-binding riboswitches. *Quarterly Reviews of Biophysics,* **45**(3), p. 345-381.

BITO, T., MATSUNAGA, Y., YABUTA, Y., KAWANO, T. and WATANABE, F., 2013. Vitamin B12 deficiency in Caenorhabditis elegans results in loss of fertility, extended life cycle, and reduced lifespan. *FEBS Open Bio,* **3**, p. 112-117.

BITO, T. and WATANABE, F., 2016. Biochemistry, function, and deficiency of vitamin B12 in Caenorhabditis elegans. *Experimental Biology and Medicine,* **241**(15), p. 1663-1668.

BLANCHE, F., DEBUSSCHE, L., FAMECHON, A., THIBAUT, D., CAMERON, B. and CROUZET, J., 1991. A bifunctional protein from Pseudomonas denitrificans carries cobinamide kinase and cobinamide phosphate guanylyltransferase activities. *Journal of Bacteriology,* **173**(19), p. 6052-6057.

BLANCHE, F., CAMERON, B., CROUZET, J., DEBUSSCHE, L., THIBAUT, D., VUILHORGNE, M., LEEPER, F.J. and BATTERSBY, A.R., 1995a. Vitamin B12: how the problem of its biosynthesis was solved. *Angewandte Chemie International Edition in English,* **34**(4), p. 383-411.

BLANCHE, F., CAMERON, B., CROUZET, J., DEBUSSCHE, L., THIBAUT, D., VUILHORGNE, M., LEEPER, F.J. and BATTERSBY, A.R., 1995b. Vitamin B12: wie das Problem seiner Biosynthese gelöst wurde. *Angewandte Chemie,* **107**(4), p. 421-452.

BLOUNT, K.F. and BREAKER, R.R., 2006. Riboswitches as antibacterial drug targets. *Nature Biotechnology,* **24**(12), p. 1558.

BRASELMANN, E., WIERZBA, A.J., POLASKI, J.T., CHROMIŃSKI, M., HOLMES, Z.E., HUNG, S., BATAN, D., WHEELER, J.R., PARKER, R. and JIMENEZ, R., 2018. A multicolor riboswitch-based platform for imaging of RNA in live mammalian cells. *Nature Chemical Biology,* **14**(10), p. 964.

BRAUN, V., 2006. Energy transfer between biological membranes. ACS Chem. Biol. P352-354

BRAUN, V. and ENDRISS, F., 2007. Energy-coupled outer membrane transport proteins and regulatory proteins. *Biometals,* **20**(3-4), p. 219-231.

BREAKER, R.R., 2012. Riboswitches and the RNA world. *Cold Spring Harbor Perspectives in Biology,* **4**(2), p. a003566.

BRIDWELL-RABB, J. and DRENNAN, C.L., 2017. Vitamin B12 in the spotlight again. *Current Opinion in Chemical Biology,* **37**, p. 63-70.

BRODERICK, J.B., DUFFUS, B.R., DUSCHENE, K.S. and SHEPARD, E.M., 2014. Radical S-adenosylmethionine enzymes. *Chemical Reviews,* **114**(8), p. 4229-4317.

BROWN, K.L., 2005. Chemistry and enzymology of vitamin B12. *Chemical Reviews,* **105**(6), p. 2075-2150.

CADIEUX, N., BRADBEER, C., REEGER-SCHNEIDER, E., KÖSTER, W., MOHANTY, A.K., WIENER, M.C. and KADNER, R.J., 2002. Identification of the periplasmic cobalamin-binding protein BtuF of Escherichia coli. *Journal of Bacteriology,* **184**(3), p. 706-717.

CAI, Y., XIA, M., DONG, H., QIAN, Y., ZHANG, T., ZHU, B., WU, J. and ZHANG, D., 2018. Engineering a vitamin B 12 high-throughput screening system by riboswitch sensor in Sinorhizobium meliloti. *BMC Biotechnology,* **18**(1), p. 27.

CHAN, C.H. and ESCALANTE-SEMERENA, J.C., 2011. ArsAB, a novel enzyme from Sporomusa ovata activates phenolic bases for adenosylcobamide biosynthesis. *Molecular Microbiology,* **81**(4), p. 952-967.

CHAN, C.H., NEWMISTER, S.A., TALYOR, K., CLAAS, K.R., RAYMENT, I. and ESCALANTE-SEMERENA, J.C., 2014. Dissecting cobamide diversity through structural and functional analyses of the base-activating CobT enzyme of Salmonella enterica. *Biochimica et Biophysica Acta (BBA)-General Subjects,* **1840**(1), p. 464-475.

CHEAH, M.T., WACHTER, A., SUDARSAN, N. and BREAKER, R.R., 2007. Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. *Nature,* **447**(7143), p. 497.

CHEONG, C., ESCALANTE-SEMERENA, J.C. and RAYMENT, I., 2002. Structural Studies of the l-Threonine-O-3-phosphate Decarboxylase (CobD) Enzyme from Salmonella enterica: The Apo, Substrate, and Product− Aldimine Complexes. *Biochemistry,* **41**(29), p. 9079-9089.

CHEONG, C., ESCALANTE-SEMERENA, J.C. and RAYMENT, I., 1999. The three-dimensional structures of nicotinate mononucleotide: 5, 6-dimethylbenzimidazole phosphoribosyltransferase (CobT) from Salmonella typhimurium complexed with 5, 6-dimethybenzimidazole and its reaction products determined to 1.9 Å resolution. *Biochemistry,* **38**(49), p. 16125-16135.

CLARDY, S.M., ALLIS, D.G., FAIRCHILD, T.J. and DOYLE, R.P., 2011. Vitamin B12 in drug delivery: breaking through the barriers to a B12 bioconjugate pharmaceutical. *Expert Opinion on Drug Delivery,* **8**(1), p. 127-140.

COHEN, G.N., 2014. Biosynthesis of cobalamins including vitamin B12. *Microbial Biochemistry.* Springer, p. 555-565.

COPLEY, S.D., 1998. Microbial dehalogenases: enzymes recruited to convert xenobiotic substrates. *Current Opinion in Chemical Biology,* **2**(5), p. 613-617.

CROFT, M.T., LAWRENCE, A.D., RAUX-DEERY, E., WARREN, M.J. and SMITH, A.G., 2005. Algae acquire vitamin B 12 through a symbiotic relationship with bacteria. Nature, 438 (7064), p90-3.

CROFTS, T., SETH, E., HAZRA, A. and TAGA, M, 2013. Cobamide Structure Depends on Both Lower Ligand Availability and CobT Substrate Specificity. Chem Biol. p1265-74.

CROUZET, J., LEVY-SCHIL, S., CAMERON, B., CAUCHOIS, L., RIGAULT, S., ROUYEZ, M.C., BLANCHE, F., DEBUSSCHE, L. and THIBAUT, D., 1991a. 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, cobyric acid synthase, and bifunctional cobinamide kinase-cobinamide phosphate guanylyltransferase. *Journal of Bacteriology,* **173**(19), p. 6074-6087.

DRENNAN, C.L., HUANG, S., DRUMMOND, J.T., MATTHEWS, R.G. and LIDWIG, M.L., 1994. How a protein binds B12: A 3.0 A X-ray structure of B12-binding domains of methionine synthase. *Science,* **266**(5191), p. 1669-1674.

ESPAH BORUJENI, A., MISHLER, D.M., WANG, J., HUSO, W. and SALIS, H.M., 2015. Automated physics-based design of synthetic riboswitches from diverse RNA aptamers. *Nucleic Acids Research,* **44**(1), p. 1-13.

FAN, C. and BOBIK, T.A., 2008. The PduX enzyme of Salmonella enterica is an L-threonine kinase used for coenzyme B12 synthesis. *Journal of Biological Chemistry,* **283**(17), p. 11322-11329.

FRANKLUND, C.V. and KADNER, R.J., 1997. Multiple transcribed elements control expression of the Escherichia coli btuB gene. *Journal of Bacteriology,* **179**(12), p. 4039-4042.

FREY, P.A., HEGEMAN, A.D. and RUZICKA, F.J., 2008a. The radical SAM superfamily. *Critical Reviews in Biochemistry and Molecular Biology,* **43**(1), p. 63-88.

GALLO, S., MUNDWILER, S., ALBERTO, R. and SIGEL, R.K., 2011. The change of corrin-amides to carboxylates leads to altered structures of the B 12-responding btuB riboswitch. *Chemical Communications,* **47**(1), p. 403-405.

GALLO, S., OBERHUBER, M., SIGEL, R.K. and KRÄUTLER, B., 2008. The corrin moiety of coenzyme B12 is the determinant for switching the btuB riboswitch of E. coli. *ChemBioChem,* **9**(9), p. 1408-1414.

GONZALEZ, J.C., BANERJEE, R.V., HUANG, S., SUMNER, J.S. and MATTHEWS, R.G., 1992. Comparison of cobalamin-independent and cobalamin-dependent methionine synthases from Escherichia coli: two solutions to the same chemical problem. *Biochemistry,* **31**(26), p. 6045-6056.

GRUBER, K., PUFFER, B. and KRAEUTLER, B., 2011a. Vitamin B 12-derivatives—enzyme cofactors and ligands of proteins and nucleic acids. *Chemical Society Reviews,* **40**(8), p. 4346-4363.

GUSAROV, I. and NUDLER, E., 1999. The mechanism of intrinsic transcription termination. *Molecular Cell,* **3**(4), p. 495-504.

HAMZA, M.S., ZOU, X., BANKA, R., BROWN, K.L. and VAN ELDIK, R., 2005. Kinetic and thermodynamic studies on ligand substitution reactions and base-

on/base-off equilibria of cyanoimidazolylcobamide, a vitamin B 12 analog with an imidazole axial nucleoside. *Dalton Transactions,* (4), p. 782-787.

HAZRA, A.B., HAN, A.W., MEHTA, A.P., MOK, K.C., OSADCHIY, V., BEGLEY, T.P. and TAGA, M.E., 2015. Anaerobic biosynthesis of the lower ligand of vitamin B12. *Proceedings of the National Academy of Sciences,* **112**(34), p. 10792-10797.

HAZRA, A.B., TRAN, J.L., CROFTS, T.S. and TAGA, M.E., 2013. Analysis of substrate specificity in CobT homologs reveals widespread preference for DMB, the lower axial ligand of vitamin B12. *Chemistry & biology,* **20**(10), p. 1275-1285.

HELDT, D., LAWRENCE, A.D., LINDENMEYER, M., DEERY, E., HEATHCOTE, P., RIGBY, S.E. and WARREN, M.J., 2005. Aerobic synthesis of vitamin B12: ring contraction and cobalt chelation. Biochem Soc Trans. p815-9.

HELLIWELL, K.E., LAWRENCE, A.D., HOLZER, A., KUDAHL, U.J., SASSO, S., KRÄUTLER, B., SCANLAN, D.J., WARREN, M.J. and SMITH, A.G., 2016. Cyanobacteria and eukaryotic algae use different chemical variants of vitamin B12. *Current Biology,* **26**(8), p. 999-1008.

HODGKIN, D.C., PICKWORTH, J., ROBERTSON, J.H., TRUEBLOOD, K.N., PROSEN, R.J. and WHITE, J.G., 1955. The crystal structure of the hexacarboxylic acid derived from B12 and the molecular structure of the vitamin. *Nature,* **176**, pp. 325-328.

JANSSEN, D.B., OPPENTOCHT, J.E. and POELARENDS, G.J., 2001. Microbial dehalogenation. *Current Opinion in Biotechnology,* **12**(3), p. 254-258.

JETER, R.M., OLIVERA, B.M. and ROTH, J.R., 1984. Salmonella typhimurium synthesizes cobalamin (vitamin B12) de novo under anaerobic growth conditions. *Journal of Bacteriology,* **159**(1), p. 206-213.

JOHNSON JR, J.E., REYES, F.E., POLASKI, J.T. and BATEY, R.T., 2012. B 12 cofactors directly stabilize an mRNA regulatory switch. *Nature,* **492**(7427), pp. 133.

JOST, M., FERNÁNDEZ-ZAPATA, J., POLANCO, M.C., ORTIZ-GUERRERO, J.M., CHEN, P.Y., KANG, G., PADMANABHAN, S., ELÍAS-ARNANZ, M. and DRENNAN, C.L., 2015. Structural basis for gene regulation by a B 12-dependent photoreceptor. *Nature,* **526**(7574), p. 536.

KOPPENHAGEN, V.B. and PFIFFNER, J.J., 1970. Currins and zirrins, two new classes of corrin analogues. *Journal of Biological Chemistry,* **245**(21), p. 5865-5867.

KOURY, M.J. and PONKA, P., 2004. New insights into erythropoiesis: the roles of folate, vitamin B12, and iron. *Annual Review of Nutrition,* **24**, p. 105-131.

KRÄUTLER, B., FIEBER, W., OSTERMANN, S., FASCHING, M., ONGANIA, K., GRUBER, K., KRATKY, C., MIKL, C., SIEBERT, A. and DIEKERT, G., 2003. The cofactor of tetrachloroethene reductive dehalogenase of Dehalospirillum multivorans is norpseudo-B12, a new type of a natural corrinoid. *Helvetica chimica acta,* **86**(11), p. 3698-3716.

KUBODERA, T., WATANABE, M., YOSHIUCHI, K., YAMASHITA, N., NISHIMURA, A., NAKAI, S., GOMI, K. and HANAMOTO, H., 2003. Thiamine-regulated gene expression of Aspergillus oryzae thiA requires splicing of the intron containing a riboswitch-like domain in the 5′-UTR. *FEBS letters,* **555**(3), p. 516-520.

LAWRENCE, A.D., NEMOTO-SMITH, E., DEERY, E., BAKER, J.A., SCHROEDER, S., BROWN, D.G., TULLET, J.M., HOWARD, M.J., BROWN, I.R. and SMITH, A.G.,

2018. Construction of Fluorescent Analogs to Follow the Uptake and Distribution of Cobalamin (Vitamin B12) in Bacteria, Worms, and Plants. Cell Chemical Biology. 16;25(8): p 941-951

LAWRENCE, C.C., GERFEN, G.J., SAMANO, V., NITSCHE, R., ROBINS, M.J., RÉTEY, J. and STUBBE, J., 1999. Binding of Cob(II)alamin to the Adenosylcobalamin-dependent Ribonucleotide Reductase from Lactobacillus leichmannii IDENTIFICATION OF DIMETHYLBENZIMIDAZOLE AS THE AXIAL LIGAND. *Journal of Biological Chemistry,* **274**(11), p. 7039-7042.

LAWRENCE, J.G. and ROTH, J.R., 1996. Evolution of coenzyme B12 synthesis among enteric bacteria: evidence for loss and reacquisition of a multigene complex. *Genetics,* **142**(1), p. 11-24.

LEWINSON, O., LEE, A.T., LOCHER, K.P. and REES, D.C., 2010. A distinct mechanism for the ABC transporter BtuCD–BtuF revealed by the dynamics of complex formation. *Nature Structural & Molecular Biology,* **17**(3), p. 332.

LUNDRIGAN, M.D. and KADNER, R.J., 1989. Altered cobalamin metabolism in Escherichia coli btuR mutants affects btuB gene regulation. *Journal of Bacteriology,* **171**(1), p. 154-161.

LUNDRIGAN, M.D., KÖSTER, W. and KADNER, R.J., 1991. Transcribed sequences of the Escherichia coli btuB gene control its expression and regulation by vitamin B12. *Proceedings of the National Academy of Sciences,* **88**(4), p. 1479-1483.

MANCIA, F., KEEP, N.H., NAKAGAWA, A., LEADLAY, P.F., MCSWEENEY, S., RASMUSSEN, B., DIAT, O. and EVANS, P.R., 1996. How coenzyme B12 radicals are generated: the crystal structure of methylmalonyl-coenzyme A mutase at 2 Å resolution. *Structure,* **4**(3), p. 339-350.

MANDER, L. and LIU, H., 2010. *Comprehensive natural products II: Chemistry and Biology.* Elsevier.

MARISCH, K., BAYER, K., SCHARL, T., MAIRHOFER, J., KREMPL, P.M., HUMMEL, K., RAZZAZI-FAZELI, E. and STRIEDNER, G., 2013. A Comparative Analysis of Industrial Escherichia coli K–12 and B Strains in High-Glucose Batch Cultivations on Process-, Transcriptome-and Proteome Level. *PloS one,* **8**(8), p. e70516.

MARSH, E.N., 1999. Coenzyme B12 (cobalamin)-dependent enzymes. *Essays in Biochemistry,* **34**, p. 139-154.

MATHEWS, F.S., GORDON, M.M., CHEN, Z., RAJASHANKAR, K.R., EALICK, S.E., ALPERS, D.H. and SUKUMAR, N., 2007. Crystal structure of human intrinsic factor: cobalamin complex at 2.6-Å resolution. *Proceedings of the National Academy of Sciences,* **104**(44), p. 17311-17316.

MATTHEWS, R.G., KOUTMOS, M. and DATTA, S., 2008a. Cobalamin-dependent and cobamide-dependent methyltransferases. *Current Opinion in Structural Biology,* **18**(6), p. 658-666.

MCGOLDRICK, H.M., ROESSNER, C.A., RAUX, E., LAWRENCE, A.D., MCLEAN, K.J., MUNRO, A.W., SANTABARBARA, S., RIGBY, S.E., HEATHCOTE, P. and SCOTT, A.I., 2005a. Identification and characterization of a novel vitamin B12 (cobalamin) biosynthetic enzyme (CobZ) from Rhodobacter capsulatus, containing

flavin, heme, and Fe-S cofactors. *Journal of Biological Chemistry,* **280**(2), p. 1086-1094.

MINOT, G.R. and MURPHY, W.P., 1926. Treatment of pernicious anemia by a special diet. *Journal of the American Medical* Association, 87(7), p. 470-476.

MOORE, S.J. and WARREN, M.J., 2012. The anaerobic biosynthesis of vitamin B12. Biochem Soc Trans. 2012 Jun 1;40(3):581-6. doi: 10.1042/BST20120066.

NAHVI, A., BARRICK, J.E. and BREAKER, R.R., 2004. Coenzyme B12 riboswitches are widespread genetic control elements in prokaryotes. *Nucleic Acids Research,* **32**(1), p. 143-150.

NAHVI, A., SUDARSAN, N., EBERT, M.S., ZOU, X., BROWN, K.L. and BREAKER, R.R., 2002. Genetic control by a metabolite binding mRNA. *Chemistry & Biology,* **9**(9), p. 1043-1049.

NORDLUND, P. and EKLUND, H., 1993. Structure and function of the Escherichia coli ribonucleotide reductase protein R2. *Journal of Molecular Biology,* **232**(1), p. 123-164.

NOU, X. and KADNER, R.J., 2000. Adenosylcobalamin inhibits ribosome binding to btuB RNA. *Proceedings of the National Academy of Sciences,* **97**(13), p. 7190-7195.

NOU, X. and KADNER, R.J., 1998. Coupled Changes in Translation and Transcription during Cobalamin-Dependent Regulation of btuB Expression inEscherichia coli. *Journal of Bacteriology,* **180**(24), p. 6719-6728.

ORTH, P., SCHNAPPINGER, D., HILLEN, W., SAENGER, W. and HINRICHS, W., 2000. Structural basis of gene regulation by the tetracycline inducible Tet repressor–operator system. *Nature Structural and Molecular Biology,* **7**(3), p. 215.

ORTIZ-GUERRERO, J.M., POLANCO, M.C., MURILLO, F.J., PADMANABHAN, S. and ELÍAS-ARNANZ, M., 2011. Light-dependent gene regulation by a coenzyme B12-based photoreceptor. *Proceedings of the National Academy of Sciences,* **108**(18), p. 7565-7570.

PERLMAN, D. and TOOHEY, J.I., 1966. Cobalt-free corrinoids as vitamin B12 antagonists. *Nature,* **212**(5059), p. 300.

PESELIS, A. and SERGANOV, A., 2012. Structural insights into ligand binding and gene expression control by an adenosylcobalamin riboswitch. *Nature Structural & Molecular Biology,* **19**(11), p. 1182.

POSTLE, K. and LARSEN, R.A., 2007. TonB-dependent energy transduction between outer and cytoplasmic membranes. *Biometals,* **20**(3-4), p. 453.

REITZER, R., GRUBER, K., JOGL, G., WAGNER, U.G., BOTHE, H., BUCKEL, W. and KRATKY, C., 1999. Glutamate mutase from Clostridium cochlearium: the structure of a coenzyme B12-dependent enzyme provides new mechanistic insights. *Structure,* **7**(8), p. 891-902.

RODIONOV, D.A., VITRESCHAK, A.G., MIRONOV, A.A. and GELFAND, M.S., 2003. Comparative genomics of the vitamin B12 metabolism and regulation in prokaryotes. *Journal of Biological Chemistry,* **278**(42), p. 41148-41159.

ROTH, J.R., LAWRENCE, J.G. and BOBIK, T.A., 1996. Cobalamin (coenzyme B12): synthesis and biological significance. *Annual Reviews in Microbiology,* **50**(1), p. 137-181.

RUETZ, M., GHERASIM, C., GRUBER, K., FEDOSOV, S., BANERJEE, R. and KRÄUTLER, B., 2013. Access to Organometallic Arylcobaltcorrins through Radical Synthesis: 4-Ethylphenylcobalamin, a Potential "Antivitamin B12". *Angewandte Chemie International Edition,* **52**(9), p. 2606-2610.

RUETZ, M., SALCHNER, R., WURST, K., FEDOSOV, S. and KRÄUTLER, B., 2013. Phenylethynylcobalamin: A Light-Stable and Thermolysis-Resistant Organometallic Vitamin B12 Derivative Prepared by Radical Synthesis. *Angewandte Chemie International Edition,* **52**(43), p. 11406-11409.

SANTOS, F., VERA, J.L., LAMOSA, P., DE VALDEZ, G.F., DE VOS, W.M., SANTOS, H., SESMA, F. and HUGENHOLTZ, J., 2007. Pseudovitamin is the corrinoid produced by Lactobacillus reuteri CRL1098 under anaerobic conditions. *FEBS letters,* **581**(25), p. 4865-4870.

SERGANOV, A. and NUDLER, E., 2013. A decade of riboswitches. *Cell,* **152**(1-2), p. 17-24.

SHELTON, A.N., SETH, E.C., MOK, K.C., HAN, A.W., JACKSON, S.N., HAFT, D.R. and TAGA, M.E., 2018. Uneven distribution of cobamide biosynthesis and dependence in bacteria predicted by comparative genomics. *bioRxiv,* p. 342006.

STABLER, S.P. and ALLEN, R.H., 2004. Vitamin B12 deficiency as a worldwide problem. *Annual Review of Nutrition,* **24**, p. 299-326.

STUBBE, J., 1994. Binding site revealed of nature's most beautiful cofactor. *Science,* **266**(5191), p. 1663.

STUDIER, F.W., DAEGELEN, P., LENSKI, R.E., MASLOV, S. and KIM, J.F., 2009a. Understanding the differences between genome sequences of Escherichia coli B strains REL606 and BL21 (DE3) and comparison of the E. coli B and K-12 genomes. *Journal of Molecular Biology,* **394**(4), p. 653-680.

TOOHEY, J.I., 1965. A vitamin B12 compound containing no cobalt. *Proceedings of the National Academy of Sciences,* **54**(3), p. 934-942.

TRZEBIATOWSKI, J.R., O'TOOLE, G.A. and ESCALANTE-SEMERENA, J.C., 1994. The cobT gene of Salmonella typhimurium encodes the NaMN: 5, 6-dimethylbenzimidazole phosphoribosyltransferase responsible for the synthesis of N1-(5-phospho-alpha-D-ribosyl)-5, 6-dimethylbenzimidazole, an intermediate in the synthesis of the nucleotide loop of cobalamin. *Journal of Bacteriology,* **176**(12), p. 3568-3575.

WARNER, D.F., SAVVI, S., MIZRAHI, V. and DAWES, S.S., 2007. A riboswitch regulates expression of the coenzyme B12-independent methionine synthase in Mycobacterium tuberculosis: implications for differential methionine synthase function in strains H37Rv and CDC1551. *Journal of Bacteriology,* **189**(9), p. 3655-3659.

WARREN, M.J., RAUX, E., SCHUBERT, H.L. and ESCALANTE-SEMERENA, J.C., 2002. The biosynthesis of adenosylcobalamin (vitamin B12). *Natural Product Reports,* **19**(4), p. 390-412.

WARREN, M.J. and SCOTT, A.I., 1990. Tetrapyrrole assembly and modification into the ligands of biologically functional cofactors. *Trends in Biochemical Sciences,* **15**(12), p. 486-491.

WEIR, D.G. and SCOTT, J.M., 1999. Brain function in the elderly: role of vitamin B12 and folate. *British Medical Bulletin,* **55**(3), p. 669-682.

WHIPPLE, G.H., ROBSCHEIT, F.S. and HOOPER, C.W., 1920. Blood Regeneration Following Simple Anemia: IV. Influence of Meat, Liver and Various Extractives, Alone or Combined with Standard Diets. *American Journal of Physiology-Legacy Content,* **53**(2), p. 236-262.

WIDNER, F.J., LAWRENCE, A.D., DEERY, E., HELDT, D., FRANK, S., GRUBER, K., WURST, K., WARREN, M.J. and KRÄUTLER, B., 2016. Total Synthesis, Structure, and Biological Activity of Adenosylrhodibalamin, the Non-Natural Rhodium Homologue of Coenzyme B12. *Angewandte Chemie International Edition,* **55**(37), p. 11281-11286.

WUERGES, J., GARAU, G., GEREMIA, S., FEDOSOV, S.N., PETERSEN, T.E. and RANDACCIO, L., 2006. Structural basis for mammalian vitamin B12 transport by transcobalamin. *Proceedings of the National Academy of Sciences,* **103**(12), p. 4386-4391.

YANG, H., LUNDBÄCK, P., OTTOSSON, L., ERLANDSSON-HARRIS, H., VENEREAU, E., BIANCHI, M.E., AL-ABED, Y., ANDERSSON, U., TRACEY, K.J. and ANTOINE, D.J., 2012. Redox modification of cysteine residues regulates the cytokine activity of high mobility group box-1 (HMGB1). *Molecular Medicine,* **18**(1), p. 250-9.

YARNELL, W.S. and ROBERTS, J.W., 1999. Mechanism of intrinsic transcription termination and antitermination. *Science,* **284**(5414), p. 611-615.

Appendix

Publications resulting from this thesis:

Understanding the control of a vitamin B_{12} riboswitch

Mohammad Abbadi[1], Sarah Spurgeon[2], Naziyat Khan[3] and Martin Warren[4]

***Abstract*— Within the life sciences switching mechanisms are pervasive at all levels, from molecules to cells and tissues. Their operation can be a key determinant of health or disease. Whilst the existence and importance of switches is widely acknowledged within the biological literature, many life scientists do not deal explicitly with the switching behaviour. Frequently, steady-state behaviour before and after switching is the primary focus. Here methods for analysis of switched systems from control engineering are applied to the modelling and analysis of a riboswitch. The model has been developed by studying the dynamics of the vitamin B_{12} riboswitch. The simulation results have been validated using *in vivo* experiments by checking the bacterial growth when using *Escherichia coli* and *Salmonella enterica* where the action of the vitamin B_{12} riboswitch is known to be a determinant of system behaviour. The paper first describes a simple model for the B_{12}-riboswitch regulatory network in *E. coli* and applies the same analysis when changing the strain to *S. enterica*. Validation of the simulation results has been undertaken by linking the dynamics of the riboswitch to bacterial growth.**

I. INTRODUCTION

Vitamin B_{12}, the cyano-derivative of cobalamin, is a water-soluble vitamin whose biological forms play key roles in metabolism that affect the normal functioning of the brain and nervous system in humans. Herein, we use the term B_{12} generically to refer to cobalamin. B_{12} is unique among the vitamins in that it is made exclusively by only certain prokaryotes and there is significant interest in how B_{12} production is controlled within microbial communities and how the nutrient makes its way through the food chain into animals. For humans B_{12} deficiency is most often associated with an autoimmune disorder that prevents the body from absorbing the nutrient but people on a strictly vegetarian diet are also prone to deficiency as plants neither make nor require B_{12}. In fact, dietary B_{12} deficiency is a severe problem in many developing countries including the Indian subcontinent, Mexico, central and South America and selected areas of Africa [1]. Moreover, B_{12} deficiency is also a common problem for the 115000 sufferers of Crohn's disease in the UK and millions more worldwide as the inflammation caused by the condition can affect the end of the ileum, which is the main area where B_{12} is absorbed [2]. To understand the control of B_{12} availability it is important to study the gene regulatory network (GRN), which is a collection of molecular regulators that interact with each other and with other substances in the cell to govern the gene expression [3].

Riboswitches are naturally occurring RNA-based regulatory components that essentially function by first sensing specific metabolites such as cofactors, amino acids and nucleotides and then regulating the expression level of proteins in the corresponding metabolic pathways [4]. The switches undergo a conformational change in response to a small-molecule ligand whereby increasing ligand concentrations either increase or decrease gene expression. The steadily increasing number of examples of such natural RNA regulators that control gene expression through diverse mechanisms in different organisms has fostered growing interest in using RNA to build synthetic controllers. Both naturally occurring and synthetic riboswitches are seen to be highly tunable components capable of regulating gene expression.

Some work to model riboswitch function has taken place and useful guidelines have been developed for tuning, but this work has been performed under steady-state operating conditions. The literature on the tuning of riboswitches also focusses primarily on the riboswitch as an individual element rather than as a component within a wider biological pathway [5]. The mathematical model which is used in this paper is based on the riboswitch regulatory pathway [3]. This riboswitch is capable, in principle, of regulating BtuB, which is an outer membrane porin that mediates high affinity binding and TonB- dependent active transport of vitamin B_{12} across the outer membrane. Regulation at the transcriptional level occurs when a B_{12} molecule binds the riboswitch aptamer domain and causes the formation of a terminator structure in the riboswitch expression platform. After estimating the parameters of the model, the system has been validated by creating an equation for the bacterial growth and it is proved experimentally that the model reproduces the system behaviour and captures the system dynamics within suitable bounds. For validation purposes, the growth levels in vitamin B_{12}-dependent bacteria were determined when grown in environments containing different concentrations of vitamin B_{12}. *E. coli* and *S. enterica* strains were used in the experiments and the growth levels were measured using optical density data recorded on triplicates at 600 nm. The verification process was performed by comparing simulation and experimental results. In addition, the assessment of the saturation effect when the concentration is high or low has been studied. The organisation of the paper is as follows: first, a preliminary analysis for the model of vitamin B_{12} is conducted. Second, the model validation is undertaken and finally, a comparison between the analytical analysis and the experimental results is made.

[1,2] Mohammad Abbadi and Sarah Spurgeon are from the Department of Electronic & Electrical Engineering, University College London, Gower St, Bloomsbury, London WC1E 6BT, United Kingdom uceemdo@ucl.ac.uk, s.spurgeon@ucl.ac.uk

[3,4] Naziyat Khan and Martin Warren are from the School of Biosciences, University of Kent, Canterbury CT2 7NZ, United Kingdom nk344@kent.ac.uk, M.J.Warren@kent.ac.uk

II. ANALYSIS OF A SIMPLE MODEL OF THE VITAMIN B_{12} RIBOSWITCH

The mathematical model which is used in this paper is based on the riboswitch regulatory pathway. Gene expression can be regulated at several levels in a cell. Differential gene transcription regulates which genes are allowed to be transcribed into RNA while selective nuclear RNA processing regulates which transcribed RNA can enter the cytoplasm and become messenger RNA. Genes can be regulated any time before, after or during the processes of translation and transcription. *E. coli* is incapable of synthesizing vitamin B_{12}. Instead, these bacteria actively transport the vitamin from the environment. BtuB, which is encoded by gene btuB, is an important component of the vitamin B_{12} transporter. Gene btuB is negatively regulated by vitamin B_{12} via a riboswitch. To develop a mathematical model for the B_{12}-riboswitch regulatory network in *E. coli*, Santillan et al. (2005) [3] identified three normalised model variables m, e and p which respectively represent the concentration of btuB mRNA , the concentration of BtuB and the concentration of vitamin B_{12}. The model itself is described by:

$$\dot{m} = \gamma\,[\phi(p) - m] \tag{1}$$

$$\dot{e} = \xi\,[m\,\theta(p) - e] \tag{2}$$

$$\dot{p} = \delta\,[\epsilon(P_{ext})\,e - p] \tag{3}$$

The biological meaning of the parameters in (1) - (3) are shown in Table I. The functions $\phi(p)$ and $\theta(p)$ denote the B_{12}-governed regulation at the transcriptional and translational levels respectively. These functions are shown to be given by the following Michaelis-Menten equations:

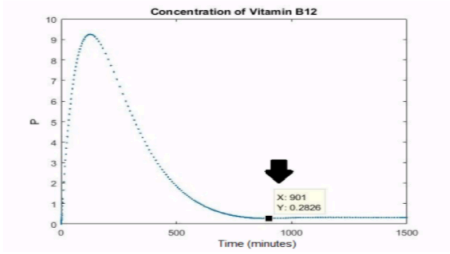
$$\phi(p) = \frac{K_\phi}{K_\phi + p}$$

$$\theta(p) = \frac{K_\theta}{K_\theta + p}$$

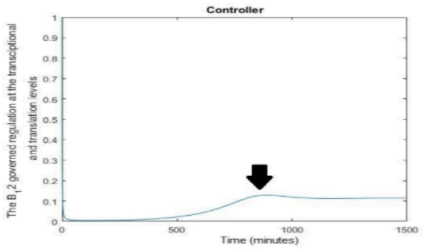
Santillan et al. [3] presents simulation results that have been undertaken for only 600 minutes, and the steady state value for the concentration of the vitamin B_{12} is 1. When the simulation is run for a longer time (1500 minutes), the concentration of vitamin B_{12} increases again after 900 minutes as shown in Fig. 1. This is an unexpected response and does not match the expected experimental results because the system should switch off when the concentration goes to zero. The controls $\Phi(P)$ and $\Theta(P)$ in the Michaelis-Menten equations depend on the value of the concentration of vitamin B_{12}, which is the third variable in the mathematical model (1) - (3) and thus the control system response is as shown in Fig. 1. As this does not match the expected experimental results, the parameters of the mathematical model (1) - (3) need to be modified. In the original paper [3], the dissociation constants at the transcription (K_ϕ) and translation (K_θ) levels are considered to be 1/4. Using this value, the steady-state value for the concentration will increase after 900 minutes. Based on the literature [8], [9] the dissociation constant for the reaction has a smaller magnitude. By changing the dissociation constant to 2.25×10^{-6}, the concentration output is correct and matches the results expected from experiments. The new parameter values are shown in Table I. The simulations using the modified parameters were carried out by solving (1) - (3) with MATLAB as shown in Fig. 2. Initial conditions of the normalised concentration of $btuB$ mRNA, the concentration of BtuB and the concentration of vitamin B_{12} were chosen as $m(0) = 1$, $e(0) = 1$ and $p(0) = 0$. All the simulations shown in this study used the same initial conditions. It is recognised from the simulation results of model (1) - (3) that the concentration of btuB mRNA has a rapid transient and thus it is assumed that

TABLE I: Pre-determined parameters in the mathematical model (1) - (3)

Symbol	Biological meaning	Original / New Experimental Value
γ	The mRNA degradation rate	7.1×10^{-2}
ξ	The btuB degradation rate	0.8×10^{-2}
δ	The B_{12} degradation rate	0.8×10^{-2}
$\epsilon(P_{ext})$	Represent the type of the strain	25
K_ϕ	The dissociation constant at the transcriptional level	1/4 / 2.25×10^{-6}
K_θ	The dissociation constant at the translation level	1/4 / 2.25×10^{-6}



(a) Concentration of vitamin B_{12}



(b) Controller

Fig. 1: The simulation results for the concentration of vitamin B_{12} and the controller in the original paper [3]

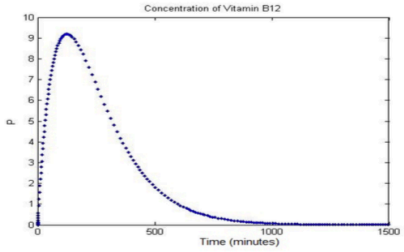


Fig. 2: Concentration of Vitamin B_{12} following simulation for 1500 minutes

this has reached steady-state in the subsequent analysis. The system's order is thus reduced to yield:

$$\dot{e} = \xi\left[\phi(p)\,\theta(p) - e\right] \tag{4}$$

$$\dot{p} = \delta\left[\epsilon(P_{ext})\,e - p\right] \tag{5}$$

using the steady-state value of m given by $\phi(p)$. There are several parameters that affect the reduced order mathematical model of the regulation equations for the B_{12} riboswitch as shown in (4) - (5). These parameters are the BtuB degradation rate, the B_{12} degradation rate and the value that is used to show which strain has been used (ϵ). The effect of varying these parameters on the concentration of vitamin B_{12} and the bacterial growth rate will be considered.

III. MODEL VALIDATION

To validate the model, it is necessary to model the growth curve from (4) - (5) and compare it with the growth curve obtained from the experiments. The comparison will include different levels of vitamin B_{12} concentration (50pM - 1nM). The effect of the parameters will be described in detail. For example, the relation between the concentration of vitamin B_{12} and the growth will be considered. Bacteria have three growth phases as shown in Fig 3.

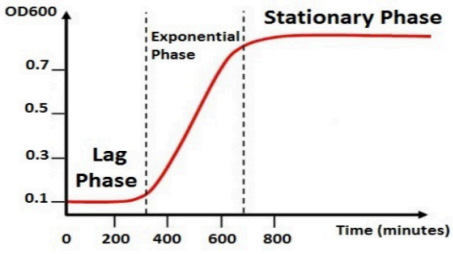


Fig. 3: Bacterial growth curve

1) **Lag phase**: when cells are transferred to fresh media, they require time to detect the environment, express specific genes, and synthesize components necessary for rapid growth. The cells are not dividing at this time.
2) **Exponential (log) phase**: binary fission occurs at a maximum rate, the cells are dividing as rapidly as possible.
3) **Stationary phase**: at this point, growth has stopped and there is no net increase or decrease in the number of cells. Bacteria use new forms of metabolism to survive in some cases producing secondary metabolites.

A model incorporating the effect of varying concentration of vitamin B_{12} with the OD600 bacterial growth output (which denotes the absorbance, or optical density, of a sample measured at a wavelength of 600 nm) is shown below:

$$OD_{600} = \begin{cases} OD_{Lag\,phase}, & \text{if } t \le t_l \\ [(\dfrac{0.057 \times P_{max}}{t_s - t_l}) \times (t - t_l)] + OD_{Lag\,phase} & \\ \text{if } t_l \le t \le t_s] & \\ OD_{t_s}, & \text{if } t \ge t_s \end{cases} \tag{6}$$

The parameters used in this equation with their biological meanings and experimental values are listed in Table II.

TABLE II: Pre-determined parameters in the growth equation (OD_{600})

Symbol	Biological meaning	Experimental Value
$OD_{Lag\,phase}$	The absorbance at OD_{600} when the growth is at the lag phase	0.1
t_l	Time to reach the log phase	400
t_s	Time to reach the saturation phase	Varies depending on vitamin B_{12} concentration
OD_{t_s}	The absorbance at OD_{600} when the growth is at the saturation phase	Varies depending on vitamin B_{12} concentration

The initial number of bacteria, bacterial type and the experimental environment depends on the particular experimental conditions. Table II presents a set of parameters consistent with the experiments performed. This choice may be justified by considering Fig. 4. P_{max} is defined to be the maximum value of the concentration of vitamin B_{12} and t_s is defined to be the time at which the maximum is reached. These values have been computed directly from the model (4) - (5) and can also be identified from the experimental results. It is clear that P_{max} and t_s vary with the change of Btub degradation rate ξ. Figure 5 shows P_{max} corresponding to different values of ξ. Changing the bacterial strain will also change P_{max} and this is demonstrated both from simulation and experimental results as shown in Fig. 6. Simulations were carried out by solving (4) - (5) and applying the simulation results in the growth equation. It is seen that by increasing the concentration of vitamin B_{12}, the absorbance at OD600

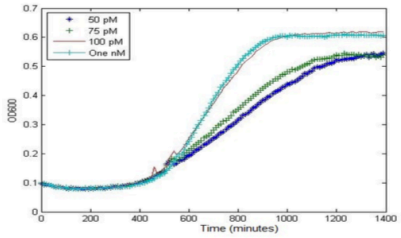
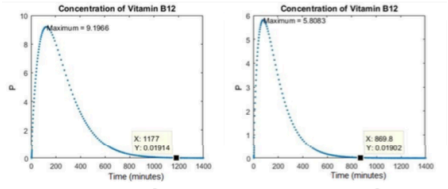
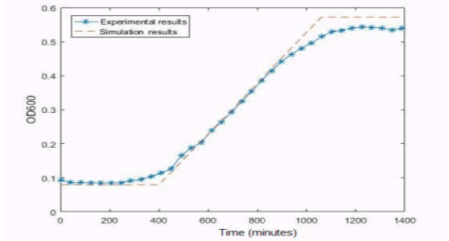


Fig. 4: Experimentally measured $E.coli$ growth curves with varying vitamin B_{12} concentration (50pM to 1nM)

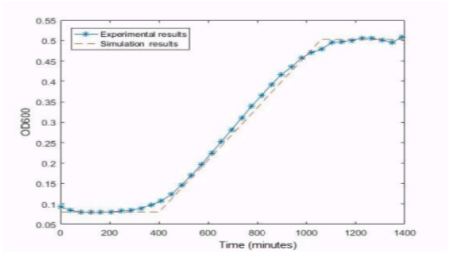


(a) 0.8×10^{-2} (b) 1.8×10^{-2}

Fig. 5: The concentration of Vitamin B_{12} when changing ξ from 0.8×10^{-2} to 1.8×10^{-2}



(a) Growth curve for 75pM when ϵ = 25



(b) Growth curve for 75pM when ϵ = 21.5

Fig. 6: Comparison between the experimental and the simulation results when changing the value of ϵ

will increase and the time required to reach the steady-state value will decrease. The time required to reach the stationary phase and the absorbance at OD600 for both simulation and experimental results for all concentrations is shown in Table III and Table IV. Fig. 4 shows the experimentally measured growth curve of *E. coli* with varying B_{12} concentration between 50pM and 1nM across three growth phases. In the lag phase, there is no change between the curves even when the concentration is changed. That is because the system requires time to detect the environment and synthesize components necessary for rapid growth while in the exponential phase. By changing the concentration of vitamin B_{12}, the growth curve changes. From Fig. 4, it is seen that the OD600 during the lag phase is 0.1 and t_s is 400 minutes. These two values vary from one plate reader to another as they are dependent on the number of bacteria present when the experiment is initialised and the atmospheric conditions in the laboratory, which may change. The exponential phase starts from 400 minutes to reach t_s. The value of t_s has a negative relation with the concentration of vitamin B_{12}, so when the concentration of vitamin B_{12} increases, the value of t_s decreases. In the stationary phase, the bacterial growth stops and there is no net increase or decrease in the number of cells. Varying the concentration of vitamin B_{12} will change the steady state value of absorbance. For example, when the concentration of vitamin B_{12} is 50pM, the value of absorbance at OD600 is 0.53 and when the concentration of vitamin B_{12} is 100pM, the value of absorbance at OD600 is 0.61 as shown in Fig. 7.

IV. Discussion

To validate the model, a comparison has been made between the experimental results and the simulation results as shown in Fig. 6. Table III shows a comparison between the time required to reach the stationary phase in simulation and experimentally while varying the concentration of vitamin B_{12}. It is clear that by increasing the concentration, growth is faster and it reaches the stationary phase more rapidly. In addition, by increasing the concentration the maximum value of the simulated concentration will increase and thus

TABLE III: Comparison between the time required to reach the stationary phase experimentally and the time required for the concentration to reach zero mathematically.

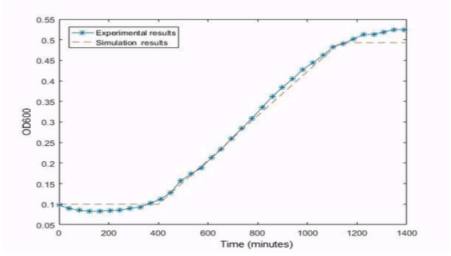
Concentration of Vitamin B_{12}	Time required to reach the stationary phase in Fig. 6, 7 (mins) Simulation Results	Time required to reach the stationary phase in Fig. 6, 7 (mins) Experimental Results
50 pM	1330	1497
75 pM	1170	1170
85 pM	1140	1135
100 pM	1090	1061

the absorbance at 600 nm will increase as shown in Table IV. Growth reaches the stationary phase when the concentration

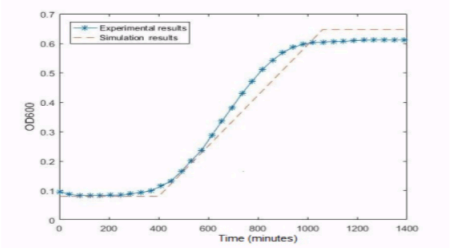
TABLE IV: Comparison between the absorbance at OD600 in the experiment and in the simulation.

Concentration of vitamin B_{12}	Simulation	Experiment
50 pM	0.503	0.532
75 pM	0.57	0.53
85 pM	0.61	0.57
100 pM	0.66	0.61

of vitamin B_{12} goes to zero. By increasing the concentration, the time to reach the stationary phase decreases. Fig. 7 shows a comparison between the experimental results and the simulation results for the growth curve when the concentration of vitamin B_{12} is 50pM and when it is 100pM. The simulation results have a similar trend to the experimental observations in terms of bacterial growth, peak values and steady-state values. The value of ϵ is related to the strain for each bacteria



(a) 50pM



(b) 100pM

Fig. 7: Comparison between the experimental and the mathematical growth curve with different concentration of Vitamin B_{12}

that has been used. Here, *E. coli* and *S. enterica* are used and they have different growth curves as shown in Fig. 6. The experiments correspond to the wild-type *E. coli* and 375 mutant strain *S. enterica* [10]. The values of ϵ employed in these simulations are as follows: ϵ = 25 (wild-type strain), ϵ = 21.5 (375 mutant strain). The value of absorbance at OD600

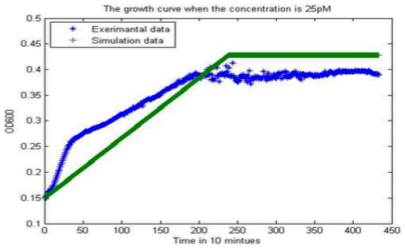
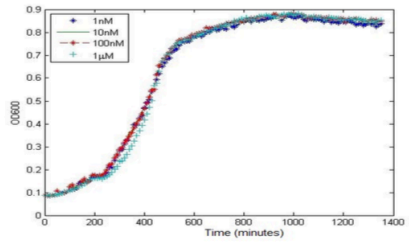


Fig. 8: Comparison between the simulation and experimental growth curve for E.coli with concentration of 25pM

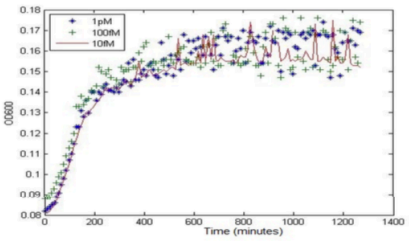
increases when the value of ϵ increases. From Fig. 6, the absorbance at OD600 is 0.63 when ϵ = 25 and it decreases to 0.52 when ϵ = 21.5.

After the model has been validated, an expected growth curve when the concentration of vitamin B_{12} is 25pM has been generated based only on the mathematical model. Experimental results were then obtained to verify the predictions. Fig. 8 shows the comparison between the simulated and experimental growth curve for *E. coli* with a concentration of 25pM. The model parameter t_l in (6) is zero and OD600 at the lag phase is 0.15. This setting is based on the plate that has been used for the experiments where it should be noted that results vary between plates.

It can be concluded that there is a positive relation between the concentration of vitamin B_{12} and bacterial growth, as when the concentration of vitamin B_{12} increases, the growth curve increases. Moreover, when the concentration of vitamin B_{12} increases to a certain level, it should be noted that bacterial growth will saturate. Comparing simulation and experimental results, it can be seen that the value of the concentration when the growth saturates is 1 nM; increasing the concentration of vitamin B_{12} by more than 1 nM, produces no change in the growth curve. Using the experimental results shown in Fig. 9a which are the experimental growth curve results for *E. coli* when the concentration of vitamin B_{12} is between 1nM and 100 μM, the experimental results agreed with the observed simulation results. In addition, this shows that for all the concentrations (1nM - 100 μM), the bacterial growth curve is approximately the same. Fig. 9b shows the growth curve for *E. coli* when varying the concentration of vitamin B_{12} between 1 pM and 10 fM. This also shows that when the growth is in the exponential phase, the growth is increasing slowly and the absorbance at OD600 is low. This also agrees with the simulation results, as when the concentration is low, the amplitude of the output is low. The growth curve remains in the exponential phase. When the concentration is decreased further, the absorbance at OD600 will also decrease to reach the same value as in the lag phase as shown in table V.



(a) high concentration



(b) low concentration

Fig. 9: Bacterial growth curve when the concentration of vitamin B_{12} are high and low

TABLE V: Simulation results for absorbance at OD600 with low vitamin B_{12} concentration.

Concentration of vitamin B_{12}	Simulation results
10 pM	0.2
5 pM	0.16
1 pM	0.11

V. CONCLUSIONS

In this paper, the effect of the vitamin B_{12} riboswitch has been tested at both the cellular and population level. For the cellular level, the concentration of $btub$ mRNA, the concentration of the BtuB and the concentration of vitamin B_{12} have been studied to determine how they affect the cells during transcription and translation. The dynamics of the vitamin B_{12} riboswitch have been incorporated in a model with *E. coli* bacterial growth. At the population level, the effect of varying the concentration of vitamin B_{12} has been tested on the bacterial growth curves of *E. coli*. The same analysis has been performed with different bacterial strains. The effect of changing the concentration of vitamin B_{12} in bacterium *S. enterica* has been tested. The results at both the cellular and population level have been linked and the simulation results obtained replicated the experimental results. The effect of enzyme inhibitors which are vitamin B_{12} analogues will be considered in future work. The long term objective of the study is to provide modelling and simulation tools to assist in the study of the impact of the vitamin B_{12} on health.

REFERENCES

[1] S. P. Stabler and R. H. Allen, "Vitamin B12 deficiency as a worldwide problem." *National Center for Biotechnology Information Search database*, 24:299-326, 2004.

[2] W. R. Best, J. M. Becktel, J. W. Singleton and F. Kern,"Development of a Crohn's Disease Activity Index," *Gastroenterology* , Volume 70 , Issue 3 , 439 - 444, 1976.

[3] M. Santillan and M. C. Mackey, "Dynamic behavior of the B_{12} riboswitch", *IOPscience*, 2005.

[4] A. L. Edwards and R. T. Batey, "Riboswitches: A Common RNA Regulatory Element," *Nature Education* 3(9):9, 2010.

[5] C. L. Beisel and C. D. Smolke, "Design Principles for Riboswitch Function," *PLOS Computational Biology*, 2009, doi: 10.1371/journal.pcbi.1000363

[6] B. Bandyopadhyay, S. Janardhanan and S. K. Spurgeon, *Advances in Sliding Mode Control: concept, theory and implementation.* Springer International Publishing, 2013.

[7] B. C. Goodwin, "Oscillatory behavior in enzymatic control processes," *Advances in Enzyme Regulation*, pp. 425 - 437, 1965, doi: 10.1016/0065-2571(65)90067-1

[8] J. M. Berg, J. L. Tymoczko and L. Stryer, "Eukaryotic Transcription and Translation Are Separated in Space and Time", *Biochemistry. 5th edition*, 2002.

[9] J. A. Stapleton, K. Endo, Y. Fujita. K. Hayashi, M. Takinoue, H. Saito and T. Inoue, "Feedback Control of Protein Expression in Mammalian Cells by Tunable Synthetic Translational Inhibition", *ACS Synth Biol*, 2012, doi: 10.1021/sb200005w

[10] A. Gudmundsdottir, C. Bradbeer and R. J. Kadner, "Altered binding and transport of vitamin B12 resulting from insertion mutations in the Escherichia coli btuB gene," *Journal of Biological Chemistry*, 1988.

[11] S. K. Spurgeon and A. J. N. Anelone, "Modelling and Simulation of the Dynamics of the Antigen-Specific T Cell Response Using Variable Structure Control Theory,"*PLOS ONE*, 2016.

Soon to be published:

Using sliding mode observers to estimate btuB concentration from measured vitamin B_{12} concentration

Mohammad Abbadi[1], Sarah Spurgeon[2], Martin Warren[3] and Naziyat Khan[4]

Abstract— A simple model for the B_{12}-riboswitch regulatory network in *Escherichia coli* is first described and the same analysis is applied when changing the strain to *Salmonella enterica*. Model validation is undertaken by linking the dynamics of the riboswitch model to bacterial growth and comparing the results obtained with in vivo experimental measurements of bacterial growth for both cases. Measurements of bacterial growth are relatively straightforward to obtain experimentally, but experimental measurements relating to the operation of the riboswitch are more difficult. Using the validated model, observer design methods from the domain of control engineering are used to estimate BtuB given measurements of the concentration of vitamin B_{12}. Validation of the estimates of BtuB has been undertaken by comparing the relationship between the BtuB and vitamin B_{12} concentrations estimated from the observer with the relationship between green fluorescent protein production and the concentration of vitamin B_{12} obtained experimentally. The results establish that using an observer as a soft sensor is a useful approach to explore the operation of a vitamin B_{12} riboswitch if measurements of the concentration of vitamin B_{12} are available.

I. INTRODUCTION

Vitamin B_{12}, the cyano-derivative of cobalamin, is a water-soluble vitamin whose biological forms play key roles in metabolism that affect the normal functioning of the brain and nervous system in humans. Herein, we use the term B_{12} generically to refer to cobalamin. B_{12} is unique among the vitamins in that it is made exclusively by only certain prokaryotes and there is significant interest in how B_{12} production is controlled within microbial communities and how the nutrient makes its way through the food chain into animals. This motivates the use of modelling to better understand

[1,2] Mohammad Abbadi and Sarah Spurgeon are from the Depart-ment of Electronic & Electrical Engineering, University College Lon-don, Gower St, Bloomsbury, London WC1E 6BT, United Kingdom uceemdo@ucl.ac.uk, s.spurgeon@ucl.ac.uk

[3,4] Martin Warren and Naziyat Khan are from the School of Biosciences, University of Kent, Canterbury CT2 7NZ, United Kingdom M.J.Warren@kent.ac.uk, nk344@kent.ac.uk

the dynamics of vitamin B_{12}. To understand the control of B_{12} availability study of the the gene regulatory network (GRN), which is a collection of molecular regulators that interact with each other and with other substances in the cell to govern the gene expression [3], is required. In many bacteria aspects of B_{12} production and uptake are controlled by riboswitches. Riboswitches are naturally occurring RNA-based regulatory components that essentially function by first sensing specific metabolites such as cofactors, amino acids and nucleotides and then regulating the expression level of proteins in the corresponding metabolic pathways [4]. The switches undergo a conformational change in response to a small-molecule ligand whereby increasing ligand concentrations either increase or decrease protein production. The steadily increasing number of examples of such natural RNA regulators that control protein production through diverse mechanisms in different organisms has fostered growing interest in using RNA to build synthetic controllers. Both naturally occurring and synthetic riboswitches are seen to be highly tunable components capable of regulating metabolic circuits. Some work to model riboswitch function has taken place and useful guidelines have been developed for tuning, but this work has been performed under steady-state operating conditions.

Beisel et al. [5] modelled the kinetics of riboswitch function when it is operating under known regulatory mecha-nisms. They used the models to develop a generalized frame-work for examining quantitative aspects of riboswitch tuning which is complicated by the integration of various regulatory mechanisms and processes such as RNA folding. The tuning of riboswitches also focusses primarily on the riboswitch as an individual element rather than as a component within a wider biological pathway. For the purposes of designing synthetic riboswitches, Sha Gong et al. [6] studied computational methods for modelling the structure and kinetics of riboswitch aptamers. Manja et al [7] studied the regulation levels which occur at either the level of transcription or translation in riboswitches with the objective of regulating gene expression in synthetic biology. The prior literature has also considered riboswitch modelling by linking the aptamer-ligand binding and expression platform [8], studying the structure and the dynamics in RNA [9] and by desiging artificial riboswitches [10], [11].

The mathematical model which is used in this paper is based on the riboswitch regulatory pathway [3]. This riboswitch is capable, in principle, of regulating BtuB, which is an outer membrane transporter that mediates high affinity binding and TonB- dependent active transport of vitamin B_{12} across the outer membrane. Regulation at the transcriptional

level occurs when a B_{12} molecule binds the riboswitch aptamer domain and causes the formation of a terminator structure in the riboswitch expression platform. After estimating the parameters of the model, the system has been validated by creating an equation for the bacterial growth and it is proved experimentally that the model reproduces the system behaviour and captures the system dynamics within suitable bounds. For validation purposes, the growth levels in vitamin B_{12}-dependent bacteria were determined when grown in environments containing different concentrations of vitamin B_{12}. *Escherichia coli* and *Salmonella enterica* strains were used in the experiments and the growth levels were measured using optical density data recorded on triplicates at 600 nm. The verification process was performed by comparing simulation and experimental results. In addition, the assessment of the saturation effect when the concentration is high or low has been studied. The primary function of BtuB is that of a transporter; it binds with a number of different molecules in order to carry them across the cell membrane and into the cell. In control theory, a state observer is a dynamic system that provides an estimate of the internal state of a given physical system, from measurements of the input and output (the available measurements) of the real system. Knowing the system state is necessary to solve many control theory problems; for example, stabilizing a system using state feedback or monitoring the condition of the dynamics. An observer is a dynamical system model which is driven by the input and output of the physical system, an output error is formed by comparing the output of the observer and the output of the physical system; this error is used to construct a feedback loop around the observer with the aim of driving the error between the observer output and the physical system output to zero. An observer can be considered a soft sensor as it provides estimates of unmeasured states. An observer provides robustness in the face of uncertainty between the observer model and the physical system. In this paper, the importance of vitamin B_{12} concentration in different bacterial strains have been studied for the fact the the experimenters can measure it and a tool from engineering is to use an observer either is used to measure states or to reconstruct unknown input and then to use a robust observer to estimate BtuB concentration. The organisation of the paper is as follows: first, a preliminary analysis for the model of vitamin B_{12} is conducted. Second, the model validation is undertaken. Third, two state observer design methods are used to estimate BtuB concentration from measurements of the vitamin B_{12} concentration and finally, an experimental results are used to validate the simulated observer results.

II. ANALYSIS OF A SIMPLE MODEL OF THE VITAMIN B_{12} RIBOSWITCH

The mathematical model which is used in this paper is based on the riboswitch regulatory pathway. Gene expression and protein production can be regulated at several levels in a cell. Differential gene transcription regulates which genes are allowed to be transcribed into RNA while selective RNA processing regulates messenger RNA levels. Genes can be regulated any time before, after or during the processes of translation and transcription. *E. coli* is incapable of synthesizing vitamin B_{12}. Instead, these bacteria actively transport the vitamin from the environment. BtuB is an important component of the vitamin B_{12} transporter system. The *btuB* gene is negatively regulated by vitamin B_{12} via a riboswitch. To develop a mathematical model for the B_{12}-riboswitch regulatory network in *E. coli*, Santillan et al. (2005) [3] identified three normalised model variables m, e and p which respectively represent the concentration of *btuB* mRNA, the concentration of BtuB and the concentration of vitamin B_{12}. The model itself is described by:

$$\dot{m} = \gamma\,[\phi(p) - m] \tag{1}$$

$$\dot{e} = \xi\,[m\,\theta(p) - e] \tag{2}$$

$$\dot{p} = \delta\,[\epsilon(P_{ext})\,e - p] \tag{3}$$

The definition of the parameters in (1) - (3) are shown in Table I. The functions (p) and (p) denote the B_{12}-governed regulation at the transcriptional and translational levels re-spectively. These functions are given by the Michaelis-Menten equations:

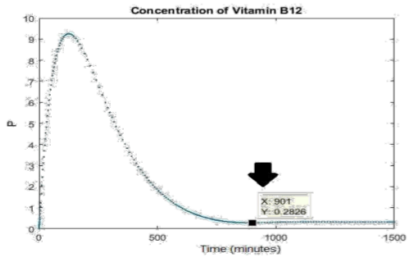
$$\phi(p) = \frac{K_\phi}{K_\phi + p}$$

$$\theta(p) = \frac{K_\theta}{K_\theta + p}$$

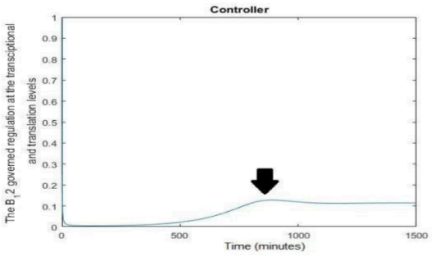
Santillan et al. [3] initially developed this model and presented simulation results for only 600 minutes. When the simulation is run for a longer time (1500) minutes with the parameters given in [3], the concentration of vitamin B_{12} increases again after 900 minutes as shown in Fig. 1. This is an unexpected response and does not match the expected experimental results because the system should switch off when the concentration goes to zero. In the original paper [3], the dissociation constants at the transcription (K) and translation (K) levels are considered to be 1/4. Based on the literature [12], [13] and [15] the dissociation constant for the reaction can be expected to have a smaller magnitude. By changing the dissociation constant to 2.25 10^6 in [14], the concentration output is correct and matches the results expected from experiments. The new

TABLE I: Pre-determined parameters in the mathematical model (1) - (3)

Symbol	Biological meaning	Original / New Experimental Value
γ	The mRNA degradation rate	7.1×10^{-2}
ξ	The btuB degradation rate	0.8×10^{-2}
δ	The B_{12} degradation rate	0.8×10^{-2}
$\epsilon(P_{ext})$	Represent the type of the strain	25
K_ϕ	The dissociation constant at the transcriptional level	1/4 / 2.25×10^{-6}
K_θ	The dissociation constant at the translation level	1/4 / 2.25×10^{-6}



(a) Concentration of vitamin B_{12}



(b) Controller

Fig. 1: The simulation results for the concentration of vitamin B_{12} and the controller in the original paper [3]

parameter values are shown in Table I. The simulations using the modified parameters were carried out by solving (1) - (3) with MATLAB as shown in Fig. 2. Initial conditions of the normalised concentration of *btuB* mRNA, the concentration of BtuB and the concentration of vitamin B_{12} were chosen as m(0) = 1, e(0) = 1 and p(0) = 0. All the simulations shown in this study used the same initial conditions. It can

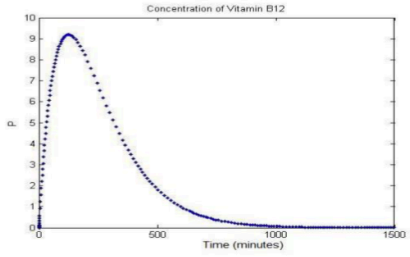


Fig. 2: Concentration of Vitamin B_{12} following simulation for 1500 minutes

be seen from the simulation results of model (1) - (3) that the concentration of *btuB* mRNA has a rapid transient and thus it is assumed that this has reached steady-state in the subsequent analysis. The system's order is thus reduced to yield:

$$\dot{e} = \xi\left[\phi(p)\,\theta(p) - e\right] \tag{4}$$

$$\dot{p} = \delta\left[\epsilon(P_{ext})\,e - p\right] \tag{5}$$

using the steady-state value of m given by (p).

The validity of the model can be demonstrated by modelling the growth curve from (4) - (5) and comparing it with the growth curve obtained experimentally. A model incorporating the effect of varying concentration of vitamin B_{12} with the OD_{600} bacterial growth output (which denotes the absorbance, or optical density, of a sample measured at a wavelength of 600 nm) has been presented in [14] and is shown below:

$$OD_{600} = \begin{cases} OD_{Lag\,phase}, & \text{if } t \le t_l \\ [(\frac{0.057 \times P_{max}}{t_s - t_l}) \times (t - t_l)] + OD_{Lag\,phase} & \\ \text{if } t_l \le t \le t_s] & \\ OD_{t_s}, & \text{if } t \ge t_s \end{cases} \tag{6}$$

The parameters used in this equation with their biological meanings and experimental values are listed in Table II.

TABLE II: Pre-determined parameters in the growth equation (OD_{600})

Symbol	Biological meaning	Experimental Value
$OD_{Lag\,phase}$	The absorbance at OD_{600} when the growth is at the lag phase	0.1
t_l	Time to reach the log phase	400
t_s	Time to reach the saturation phase	Varies depending on vitamin B_{12} concentration
OD_{t_s}	The absorbance at OD_{600} when the growth is at the saturation phase	Varies depending on vitamin B_{12} concentration

The initial number of bacteria, bacterial type and the experimental environment depends on the particular experimental conditions. Table II presents a set of parameters consistent with the experiments performed. P_{max} is defined to be the maximum value of the concentration of vitamin B_{12} and t_s is defined to be the time at which the maximum is reached. These values have been computed directly from the model (4) - (5) and can also be identified from the experimental results shown in Fig. 3.

Following validation of the model [14], an expected growth curve when the concentration of vitamin B_{12} is 25 pM has been generated based only on the mathematical model. Experimental results were then obtained to verify the predictions. Fig. 4 shows the comparison between the simulated and experimental growth curve for *E. coli* with a concentration of 25 pM. The model parameter t_l in (6) is zero and OD_{600} at the lag phase is 0.15. This setting is based on the plate that has been used for the experiments where it should be noted that results vary between plates.

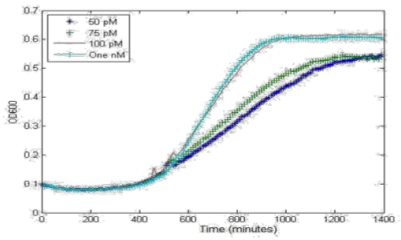


Fig. 3: Experimentally measured E:coli growth curves with varying vitamin B_{12} concentration (50pM to 1nM)

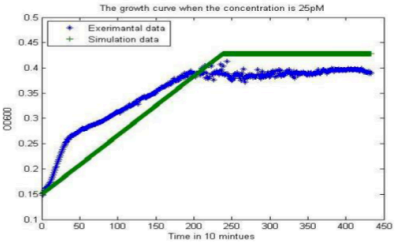


Fig. 4: Comparison between the simulation and experimental growth curve for E:coli with vitamin B_{12} concentration of 25pM

III. ANALYSIS OF A SIMPLE MODEL OF A btuB OBSERVER

In this section, two observer candidates will be presented. Before developing the observers, it is necessary to verify that the model is observable. A system with an initial state, x(t0) is observable if and only if the value of the initial state can be obtained from measurements of the system output p(t) that have been obtained over a finite time $t_0 < t < t_f$. If the initial state cannot be obtained, then the system is unobservable. A straightforward observability check is to produce the input-output form of the system. Input-output form is a mathematical model of a physical system as a function of the input and output and the derivatives of the input and output. In equations (4), (5), the product (p) (p) is set to be the input signal and p is set to be the output. The corresponding input-output form is given by:

$$p = [(P_{ext})[U \quad e] \quad p] \tag{7}$$

where U = (p) (p) = $\frac{K}{K+p}\ \frac{K}{K+p}$ It is clear that the system state can be observed from knowledge of the output measurement p(t) and the corresponding control inputs. Two observers will now be presented based on the model developed in the previous sesction. The objective is to formulate a soft sensor to reconstruct BtuB concentration based on measurements of the concentration of vitamin B_{12}. The first observer strategy is based on an unknown input observer formulation and uses a first order model representation. The second observer uses a second order model representation to directly estimate e(t).

A. AN UNKNOWN INPUT APPROACH TO ESTIMATE btuB

This observer will use the measured concentration of vitamin B_{12}, p(t) as the known output and use the corresponding dynamical equation (5) to reconstruct BtuB concentration. An unknown input observer technique will be used [17]. Referring to Eq. (5), [(P_{ext}) e] will be considered as the unknown input. The assumed observer is as follows where is an observer injection to be designed:

$$\bar{p}\hat{} = \quad p + \tag{8}$$

Define the error between the plant and observer by

$$e_1 = p\hat{} \quad p \tag{9}$$

The error dynamics is then given by

$$e\_1 = p\hat{} \quad p = \qquad (P_{ext})\ e\hat{} \qquad - \tag{10}$$

The injection must be designed so the error is driven to zero. When this holds, the designed injection compensates for the unknown input and the BtuB concentration can then be estimated as follows

$$\hat{e} = \frac{\mu}{\delta\,\epsilon(P_{ext})} \tag{11}$$

can be designed using numerous techniques. Appealing to sliding mode concepts, define:

$$\mu = Q\,sgn(e_1) \tag{12}$$

Q should be greater than (P_{ext}) in magnitude to enforce a sliding mode, whereby $e_1 = 0$. This selection ensures the error becomes zero in finite time. An alternative approach is to use a Luenberger observer paradigm:

$$\mu_1 = Q_1\,e_1 \tag{13}$$

The performance of the Luenberger observer can be expected to approach that of the sliding mode observer when Q_1 is selected to be very large, so that a high gain observer results. Figure 5 compares the two observers with no uncertainty or disturbance considered. In practice the observer dynamics is an approximation of the system dynamics. The observer performance will now be considered in the presence of a realistic range of the parameter variations in the system dynamics (4), (5). The sliding mode observer is expected to exhibit greater robustness than the Luenberger observer; infinite gain Q_1 would be required for the Luenberger observer to achieve the same performance as the sliding mode observer in the presence of parameter variations. The first parameter variation considered is the bacterial growth rate which will be bacteria dependent. The growth rate of E:coli cultured in a minimal medium can be calculated from:

$$= \frac{\ln(2)}{t_d} \tag{14}$$

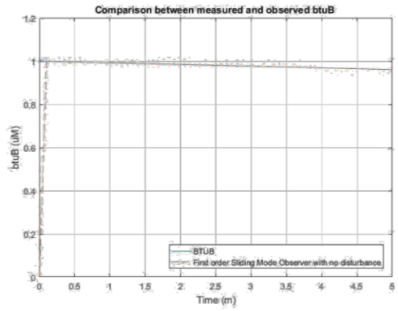


Fig. 5: Time evaluation for estimated and measured system btuB with no uncertainty

where t_d is the time required to double the number of cells. The bacterial growth rate can be calculated in practice using OD_{600} data as shown in Fig. 3 when the vitamin B_{12} concentration is 100 pM and using equations:

$$\ln OD \quad \ln OD_0 = \quad t \tag{15}$$

$$\ln OD_2 \quad \ln OD_1 = 2{:}303 \quad (\log OD_2 \quad \log OD_1) = \quad (t_2 \quad t_1) \tag{16}$$

$$= \frac{2{:}303 \quad (\log OD_2 \quad \log OD_1)}{t_2 \quad t_1} \quad = 0{:}008 \tag{17}$$

The full set of parameter ranges considered for the model are as given in Table III. The growth rates of different types of bacteria are compared in [16]. The btuB and vitamin B_{12} degradation rates depend on the bacterial growth rate and a realistic range for both rates is shown in table

III. The assumed range for P_{ext} is based on the work of Santillan et al. [3] which demonstrates that the wild-type, 375 mutant, 434 mutant, 343 mutant strains are represented by 25; 21:5; 15:75 and 5:8, respectively. The dissociation constant at the transcriptional and translation level should be small and the selected range is based on the results reported in [12], [13] and [15]. In the tests that follow the model used to represent the physical system which generates p(t) and the model used to define the observer are assumed to take different parameter values from the defined ranges. This test is replicating the fact that the parameters used to define the observer will be an approximation of any physical system.

TABLE III: Pre-determined parameters in the mathematical model (1) - (3)

Symbol	Biological meaning	Realistic Range
	The btuB degradation rate	[0.693 - 0.4 10 ^{2}]
	The B_{12} degradation rate	[0.693 - 0.4 10]
(P_{ext})	Represent the type of the strain	[25 - 5.8]
K	The dissociation constant at the transcriptional level	[2 10 ^{9} - 2 10 ^{4}]
K	The dissociation constant at the translation level	[2 10 ^{9} - 2 10 ^{4}]

B. A FULL ORDER btuB OBSERVER

In this section, a second order observer will be implemented using equations (4) and (5) to determine the model. Luenberger [18] and Utkin [19] injection signals will be used to force the observation error to zero. The corresponding Luenberger observer is given by:

$$\dot{\hat{e}} = [\ (p)\ (p)\ \hat{e}] + L_1 \quad (\hat{p}\ \ p) \tag{18}$$

$$\dot{\hat{p}} = [\ (P_{ext})\ \hat{e} \quad \hat{p}] + L_2(\hat{p} \quad p) \tag{19}$$

If the error between the plant and the observer is defined by:

$$\begin{aligned} e_1 &= e \quad \hat{e} \\ e_2 &= p \quad \hat{p} \end{aligned} \tag{20}$$

the error dynamics between the plant and the observer may be expressed by

$$\begin{aligned} \dot{e}_1 &= e_1 + L_1 e_2 \\ \dot{e}_2 &= e_1 \quad e_2 + L_2 e_2 \end{aligned} \tag{21}$$

The poles of the closed-loop error dynamics (21) are the roots of the following characteristic equation:

$$s^2 + s(-L_2 + \delta + \xi) + \xi(\delta - L_2) - L_1\delta\epsilon \tag{22}$$

Appropriate adjustment of the gains L_1 and L_2 enables the closed-loop poles of the observer to be placed in any desired location. With the parameter values from Table I and = 25, the selection $L_1 = 1000$ and $L_2 = 30$ yields closed-loop observer poles at 10 and 20.

The second observer will be based on an Utkin observer formulation. The observer is defined based on the nominal model in equations (4) and (5) as follows:

$$\dot{\hat{e}} = \xi\,[\phi(p)\,\theta(p) - \hat{e}] + L\mu \tag{23}$$

$$\dot{\hat{p}} = \delta\,[\epsilon(P_{ext})\,\hat{e}] - \mu \tag{24}$$

Where and L define the observer injection to be designed so the error between the measured and the observed con-centration of vitamin B_{12} is driven to zero. Using the error definitions in (20), the error dynamics are given by

$$\begin{aligned} \dot{e}_1 &= e_1 \quad L \\ \dot{e}_2 &= e_1 \quad e_2 + \end{aligned} \tag{25}$$

To induce a sliding motion on the output error e_2 so that the sliding condition $e_2 = 0$ is enforced, it is necessary to ensure e_2 and $\dot{e}_2$ have opposite signs. Consider defined by

$$= \qquad M\ \mathrm{sign}(e_2) \tag{26}$$

It follows that for sufficiently large M, e_2 and $\dot{e}_2$ have opposite signs and $e_2 = 0$ will be satisfied. When this sliding mode is attained, it follows that $e_2 = 0$ and $\dot{e}_2 = 0$. Using the principle of the equivalent injection it follows from (25) that on average, the effect of the applied injection signal (26), denoted $_{eq}$, is given by

$$_{eq} = \qquad e_1 \tag{27}$$

To complete the observer design it is necessary to ensure the e_1 subsystem in (25) exhibits stable dynamics in the sliding mode. Substituting (27) in (25)

$$e\_1 = (\quad L \quad)e_1 \tag{28}$$

It follows that for stability of (28)

$$L < \frac{\xi}{\delta \epsilon} \tag{29}$$

The Utkin observer (23)-(24) with the injections defined by (26) and (29) will ensure stable evolution of the error trajectories as shown in Figure (6). The parameters are selected as M = 1 and L = 100.

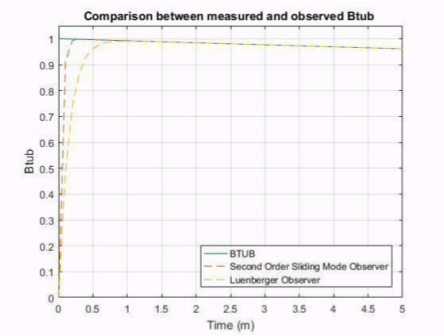


Fig. 6: Time evaluation for measured and observed btuB dynamics when applying Utkin and Luenburger observation methods with no disturbance

To test the robustness of the observers in the presence of parameter uncertainty, two tests have been performed. The first considers variation in the parameter . The second test assumes that the parameters of the observer have been configured for a different bacteria than is present in the real system. This corresponds to changing all the parameters in equations (4) and (5).

ξ is the BtuB degradation rate, and from [3], it is determined by the summation of the bacterial growth rate and the BtuB degradation rate. The value of the BtuB degradation rate is assumed negligible in [3] and because of that, the value of is set equal to the bacterial growth rate, which is 0:008 for *E. coli.* In this robustness test, the BtuB degradation rate is not neglected and an additional value will be added to the bacterial growth rate to define

. Fig 7 shows compares the dynamics of BtuB obtained from the model with = 0:016 as well as the extimates obtained from the Utkin and Luenberger observers, where in the observers = 0:008. It is seen from Fig 7 that varying the parameter does not affect the estimate of the BtuB concentration. Despite the uncertainty, both observers converge to observe the correct value of the BtuB concentration based only on

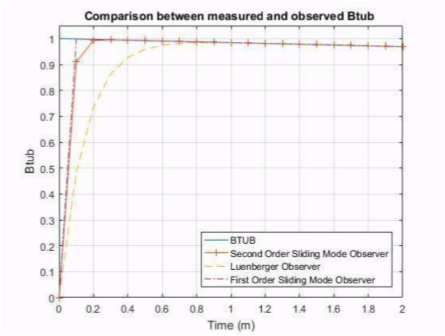


Fig. 7: Comparison between the estimated and the actual value of BtuB when changes between the system and observer from 0.016 to 0.008

knowledge of the concentration of vitamin B_{12} in less than 1 minute.

The second test changes the bacterial type. Previously, *E. coli* was used as the bacterial stain in the system model. This test uses the parameters for salmonella in the system model; the observers remain parameterised as for *E. coli.* In the system model the growth rate is 0:007 and thus

= = 0:007; bacterial strain changes to = 21:5. In the observers the model parameters are as in Table I. The

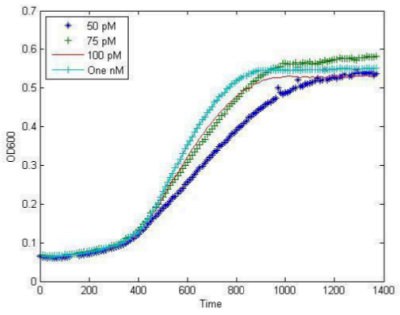


Fig. 8: Experimentally measured *S. enterica* growth curves with varying vitamin B_{12} concentration (50 pM to 1 nM)

corresponding observation results after changing the bacterial type are shown in Fig. 9-10.

It is seen from Fig. 9 that the steady-state performance is reasonable with the observers requiring around 200 minutes to track the btuB signal from the system model. From Fig. 10 it is seen that the Utkin observer has a better transient performance that the Luenberger observer as has

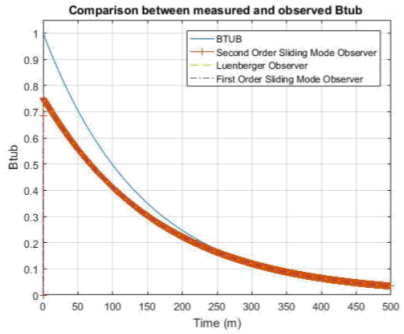


Fig. 9: Comparison between the estimated and the actual value of BtuB when the bacteria changes between the system and observers - steady-state performance

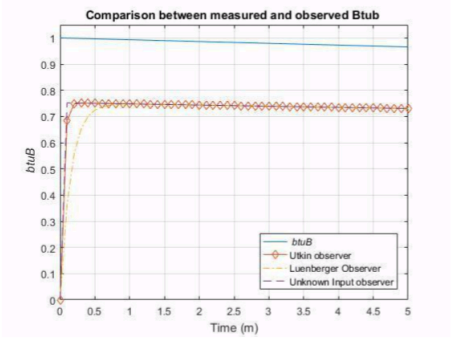


Fig. 10: Comparison between the estimated and the actual value of BtuB when the bacteria changes between the system and observers- early transient

been observed in the nominal simulations.

IV. EXPERIMENTAL VALIDATION OF OBSERVER RESULTS

Observers have been presented for estimation of the concentration of BtuB from measurements of the concentration of vitamin B_{12} . The robustness and performance validation has thus far been based on results from experimentally validated mathematical models. Validation of the estimates of btuB concentration obtained from the observers against experimental measurements is desirable. It is, however, difficult to measure the concentration of BtuB directly experimentally. Fig. 11 shows a comparison of the variation in the concentration of vitamin B_{12} against the construction of a plasmid in which green fluorescent protein (GFP) production is controlled by a vitamin B_{12} riboswitch. This is an alternative way of viewing BtuB production experimentally. GFP is used because its levels can be readily detected after the reaction is complete by use of anti-GFP antibodies and detection on a western blot. When this plasmid is subject to transcription/translation (in an in vitro protein synthesis kit), in the absence of B_{12}, GFP production will occur as normal. However, if B_{12} (the preferred riboswitch ligand being adenosylcobalamin) is present, it is thought that this will bind to the riboswitch and change its structure so that the ribosome can no longer bind and translation can no longer occur. Therefore, using this method, various B_{12} variants have been tested by adding them to the *in vitro* protein synthesis reactions, at increasing concentrations, in order to see the effect on the vitamin B_{12} riboswitch and in turn, on GF P production. Figure 11 contains two curves, one obtained with

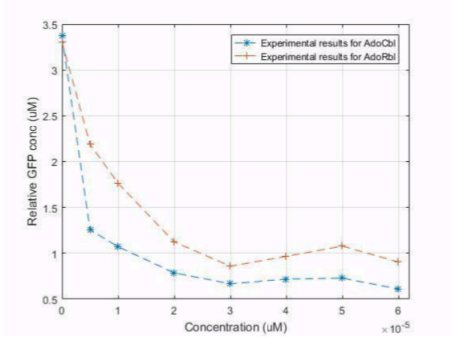


Fig. 11: Experimental results to show the relation between GF P and vitamin B_{12} for AdoCbl, AdoRbl

adenosylcobalamin AdoCbl and the other using adenosylrhodibalamin AdoRbl. The only differences between the variants is that in AdoRbl, the central cobalt ion is replaced with rhodium (Reference – Widner et al 2016). The plot of shows that increasing vitamin B_{12} concentration will decrease the GFP concentration and changing B_{12} variants will affect the magnitude of the response.

In order to validate the observer results experimentally, the observed concentration of BtuB, e^(t) is plotted against the concentration of vitamin B_{12}, p(t) for the case of the Utkin observer. The observer was parameterised using the strategy reported in [14] by considering the evolution of the concentration of vitamin B_{12} over time from the bacterial growth curve. A comparison between particular forms of vitamin B_{12} variants in [20] found that each vitamin B_{12} analogue has a specific degradation rate. In this case, each degradation rate will give a specific concentration of vitamin B_{12}. Fig. 12 shows the relation between the observed concentration of BtuB and the concentration of vitamin B_{12}. The simulation results were computed when in (5) was

= 0:008 for AdoRbl and = 0:0054 for AdoCbl.

The simulation results in Fig. 12 show that the BtuB con-

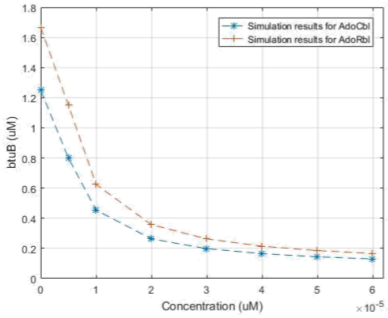


Fig. 12: Simulation results to show the relation between btuB and vitamin B_{12} for AdoCbl and AdoRbl

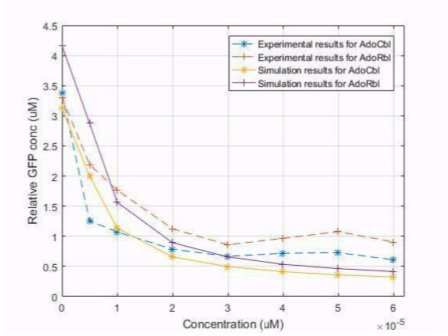


Fig. 13: Comparison between the simulated and experimental GF P for AdoCbl and AdoRbl.

centration decreases when the concentration of vitamin B_{12} increases in line with the observations from the experimental results. The rate of decay is noted to be similar. It is desirable to be able to directly compare the relationship between the evolution of the observed BtuB and the concentration of vitamin B_{12} and the relationship between GFP and the concentration of B_{12}. In particular it is required to find the relationship between GFP production and BtuB production, as in the experiments the GFP gene replaces the *btuB* gene. It follows that the concentration of GFP should correlate with the concentration of BtuB. Certain factors will affect the absolute values of GFP. The initial conditions for the experimental observations will be impacted by the experimental set-up. The size of each protein will influence the experimental results. If one was 30 kDa and another 15 kDa, then it may be expected that twice the amount of the smaller one may be produced because the RNA polymerase and protein translation apparatus can only work so fast. As both figures have the same dynamics, a factor has been introduced to adjust the magnitude observed in Fig. 12. Effectively it is assumed that a relationship GFP = e^(t). Figure 13 shows a comparison between the simulation and experimental results when = 2:5. The correlation between the experimental results and the estimate of the concentration of BtuB ob-tained by the observer is reasonable. The experimental results further validate that the presented observer approach is able to provide good estimates of the concentration of BtuB.

V. CONCLUSIONS

In this paper, the effect of the vitamin B_{12} riboswitch has been tested at both the cellular and population level. For the cellular level, the concentration of btub mRNA, the concentration of the BtuB and the concentration of vitamin B_{12} have been studied to determine how they affect the cells during transcription and translation. The dynamics of the vitamin B_{12} riboswitch have been incorporated in a model with *E. coli* bacterial growth. At the population level, the effect of varying the concentration of vitamin B_{12} has been tested on the bacterial growth curves of *E. coli*. The same analysis has been performed with different bacterial strains. The effect of changing the concentration of vitamin B_{12} in bacterium *S. enterica* has been tested. The results at both the cellular and population level have been linked and the simulation results obtained replicated the experimental results. The effect of enzyme inhibitors which are vitamin B_{12} analogues will be considered in future work. The long term objective of the study is to provide modelling and simulation tools to assist in the study of the impact of the vitamin B_{12} on health.

REFERENCES

[1] Stabler SP, Allen RH, "Vitamin B_{12} deficiency as a worldwide problem.," National Center for Biotechnology Information Search database, 2004, 24:299-326.
[2] M. T. Croft, A. D. Lawrence, E. Raux-Deery, M. J. Warren, and A. G. Smith,"Development of a Crohn's Disease Activity Index," Gastroenterology , 1976, , Volume 70 , Issue 3 , 439 - 444
[3] Moises Santillan, Michael C Mackey, "Dynamic behavior of the B_{12} riboswitch", IOPscience, 2005.
[4] Edwards, A. L. , Batey, R. T. "Riboswitches: A Common RNA Regulatory Element.", Nature Education, 2010.
[5] Chase L. Beisel, Christina D. Smolke, "Design Principles for Riboswitch Function," PLOS Computational Biology, 2009, doi: 10.1371/journal.pcbi.1000363
[6] Sha Gong, Yanli Wang, Zhen Wang and Wenbing Zhang, "Com-putational Methods for Modelling Aptamers and Designing Ri-boswitches," International journal of Molecular sciences , 2017, doi: 10.3390/ijms18112442
[7] Wachsmuth, M., Domin, G., Lorenz, R., Serfling, R., FindeiB, S., Stadler, P. F. and Morl , M . " Design criteria for synthetic riboswitches acting on transcription" RNA biology vol. 12,2 (2015): 221-31. doi: 10.1080/15476286.2015.1017235
[8] Fareed Aboul-ela, Wei Huang, Maaly Abd Elrahman, Vamsi Boyapati, Pan L "Linking aptamer-ligand binding and expression platform fold-ing in riboswitches: prospects for mechanistic modelling and design," WIREs RNA 6:631650, 2015, doi: 10.1002/wrna.1300
[9] Liu, Chang C. ; Qi, Lei ; Lucks, Julius B. ; Segall-Shapiro, Thomas H. ; Wang, Denise ; Mutalik, Vivek K. ; Arkin, Adam P. / An adaptor from translational to transcriptional control enables predictable assembly of complex regulation. In: Nature Methods. 2012 ; Vol. 9, No. 11. pp. 1088-1094.

[10] Atsushi Ogawa, Artificial Riboswitches . Humana Press, doi: 10.1007/978-1-62703-755-6.
[11] Findei, Sven et al. Design of Artificial Riboswitches as Biosensors Sensors (Basel, Switzerland) vol. 17,9 1990. 30 Aug. 2017, doi:10.3390/s17091990
[12] Berg JM, Tymoczko JL, Stryer L, " Eukaryotic Transcription and Translation Are Separated in Space and Time", Biochemistry. 5th edition, 2002.
[13] James A. Stapleton, Kei Endo, Yoshihiko Fujita, Karin Hayashi, Masahiro Takinoue, Hirohide Saito and Tan Inoue, "Feedback Control of Protein Expression in Mammalian Cells by Tunable Synthetic Translational Inhibition", ACS Synth Biol, 2012, doi: 10.1021/sb200005w
[14] Abbadi, M; Spurgeon, S; Khan, N; Warren, M; (2018) "Understanding the control of a vitamin B_{12} riboswitch ", In: 2018 UKACC 12th International Conference on Control (CONTROL). (pp. pp. 474-479). doi: 10.1109/CONTROL.2018.8516881
[15] Bintu, Lacramioara et al. "Transcriptional regulation by the numbers: applications " Current opinion in genetics and development, 2005, vol. 15,2 : 125-35.
[16] Mason, M M. A Comparison of the Maximal Growth Rates of Various Bacteria under Optimal Conditions Journal of bacteriology,1935, vol. 29,2 : 103-10.
[17] Floquet, T., Edwards, C., and Spurgeon, S.K., " On sliding mode observers for systems with unknown inputs", International Journal of Adaptive Control and Signal Processing, Volume 21, Issue 8-9 , 638 656, 2007.
[18] Davis J.H. " Luenberger Observers. In: Foundations of Deterministic and Stochastic Control. Systems and Control: Foundations and Applications." pp 245-254, 2002, Birkhauser, Boston, MA, doi.org/10.1007/978-1-4612-0071-0 8
[19] Spurgeon S.K. " Sliding mode observers: a survey." pp 751-764, 2008, International Journal of Systems Science, Volume 39,8, doi:10.1080/00207720701847638
[20] Paul C, and Brady D.M. "Comparative Bioavailability and Utilization of Particular Forms of B_{12} Supplements With Potential to Miti-gate B12-related Genetic Polymorphisms" Integr Med (Encinitas), 2017;16(1):42-49.