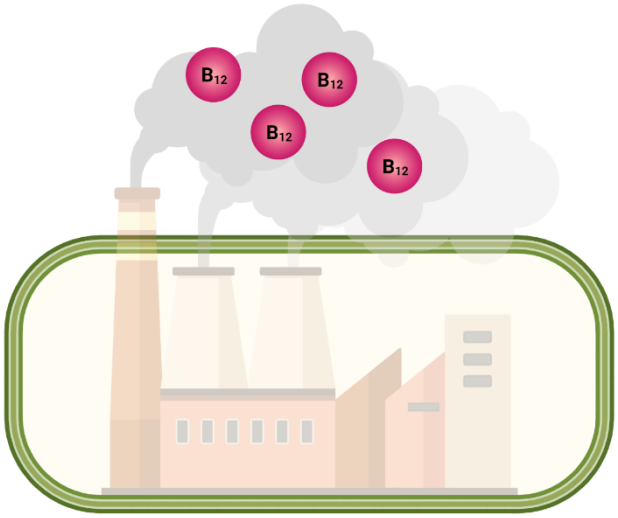
Metabolic engineering of *E. coli* as a cell factory for enhanced vitamin B_{12} production



Rozanne Stroek

February 2022

Declaration

No part of this thesis has been submitted in support of any application for a qualification at the University of Kent Canterbury or any other institute of learning.

Rozanne Stroek
February 2022

Acknowledgements

First and foremost, I would like to thank Professor Martin Warren for the opportunity to undertake a PhD research project in the group. It has been a challenging, incredible experience. I would like to thank the entire Warren Lab, past and present, for help and support throughout the course of my PhD. Special shout out to Dr. Evelyne Deery for all the help, guidance and sharing her vast knowledge with me.

Special thanks to Garry Rogerson for the interest taken in this project and the financial support that made my PhD possible.

Thank you to my friends who have provided much needed support this past 4 years. Hester, thank you for always being there, and taking the first plane when I needed you. Stephanie, for your support and being an inspiration. And of course Helen (bad devil), and Eith (Magenta) for taking part in all our adventures, and Camille – chocolatine forever!

Nick, 'one day' has become 'now'. Thank you for your love and support throughout.

And to my parents; without your endless love and support, I never would have been able to get through this. Thank you for always believing in me, even when I do not. You are my rocks.

Abstract

Vitamin B_{12} (cobalamin) is one of the most complex small molecules found in nature. Uniquely for a vitamin, its biosynthesis is restricted solely to certain prokaryotes, and humans are dependent on a dietary uptake of this essential nutrient mainly from meat, diary, eggs and fish sources. Significantly, vitamin B_{12} is absent from plants and hence those on vegetarian and especially vegan diets are prone to deficiency, which manifests in a range of neurological and haematological conditions. Vitamin B_{12} is required not only for improved human health, but it is also widely used in livestock farming for improved animal welfare and in biotechnology. There is thus a strong global demand for vitamin B_{12}, which is produced commercially through bacterial fermentation from a strain of *Pseudomonas denitrificans*. This strain employs the aerobic pathway for cobalamin synthesis and gives yields of about 200 mg/L of culture. The majority (circa 85%) of cobalamin production is located in China, where production has been variable which has caused volatile price fluctuations in recent years. To apply synthetic biology as a tool to enhance vitamin B_{12} production, researchers have turned to *E. coli* as a chassis to house the vitamin B_{12} biosynthetic pathway for the improved production of the nutrient. Previous work in the Warren lab has resulted in *E. coli* strain ED656-3B, that contains the genes of the aerobic B_{12} pathway integrated into the genome within two distinct operons, and is able to produce 40 mg/L cobalamin. To enter a competitive market, this yield needs to be increased to approach 1 g/L. The work described in this thesis involves rational approaches to enhance cobalamin yields.

The initial *E. coli* B_{12} - producing strain had the operons under the control of T7 promoters. The resulting high metabolic cost associated with T7 expression has shown that B_{12} production is shut off by the organism within 9 hours of inducing transcription of the pathway. Using CRISPR-Cas9 cloning, the two operons were placed under the control of constitutive promoters of varying strengths in an attempt to alleviate this metabolic stress, and extend the production period. However, the introduction of a weaker constitutive promoter in one of the operons resulted in decreased yields of cobalamin. The operon that was most sensitive to this change encoded genes for the entire pathway, showing a 75% drop of cobalamin yield after 24 hours when the T7 promoter was changed to a constitutive promoter. After 48 hours this yield decreases further to 10% when compared to the control strain. The other operon, which sees no dramatic effect from the change, contains an extra set of cobaltochelatase. Despite the little effect of the promoter change in this operon, a build-up of substrate intermediate hydrogeneobyrinic acid *a,c*- diamide (HBAD) indicates this point of the pathway is still vital for production improvement.

A further approach towards increased yield was to target genes *cobA* and *cobI*, that encoded rate-limiting enzymes as well as genes *SAMS* and *SAHH*, whose products contributed to the production of substrates for the B_{12} biosynthetic pathway. Investigating the effect of targeting these points in the pathway was achieved by adding the relevant genes to T7-promoter based plasmids and transforming them into the host *E. coli* strain, ED656-3B. However, all the T7 plasmid-based genes proved detrimental to B_{12} production. The genes were therefore subsequently placed under the control of a more regulated tet-operator system. Introducing CobA from *Methanosarcina barkeri* to the B_{12}-producing *E. coli* strain in such a format showed a two-fold increase in cobalamin yield, indicating this part of the pathway to be an interesting target for increased production values. CRISPR-Cas9 cloning of *M. barkeri cobA* into the genome of ED656-3B was not successful, but should be pursued further as a means of improving cobalamin yields of strain ED656-3B.

Finally, the enzyme responsible for inserting cobalt into the centre of B_{12} was investigated for cobalamin production improvement, as intermediate build-up of HBAD has shown this step to be a crucial point in the pathway. The wild-type *B. melitensis* cobaltochelatase that is incorporated into the biosynthetic B_{12} pathway of ED656-3B was characterised as well as several mutants that were created. Mutations made in the active site of cobaltochelatase subunit CobA were chosen based on previous previous research findings in the Warren lab, indicating residues involved in substrate binding of HBAD. These included R912, H1094, D1098 and Q1104. Futhermore, residues described in literature to be involved part of cobalt binding were targeted as one such mutation resulted in a more efficient turnover activity of magnesium chelatase enzyme, the function and structure of which is closely related to cobaltochelatase. The corresponding residues in *B. melitensis* CobN were E593 and D1101. A change to alanine at positions E593 and R912 resulted in the enzyme as a whole to become inactive, whereas D1101A, Q1104A, D1098A, and H1094A resulted in a drecrease in activity by 10, 20, 25, and 50%, respectively.

Of the three main approaches described, the addition of an additional *cobA* gene from *M. barkeri* was found to be the most promising avenue for improving B_{12} production by *E. coli*.

Table of Contents

Chapter 1

Introduction

1.1 Introduction

It has been nearly a 100 years since the two American physicians Minot and Murphy took the first steps in discovering vitamin B_{12} when they demonstrated the ability to cure the up till then lethal disease pernicious anaemia with crude liver extract (Minot & Murphy, 1926). Since then, research into vitamin B_{12} has become a scientific field in its own right, yielding extraordinary feats in structural biology, organic chemistry as well as biochemistry. For the purpose of clarity, a note is made here in regards to nomenclature, which can be somewhat confusing. For reasons further explained in the historical background (Section 1.2), 'true' vitamin B_{12} (B_{12}) refers to cyano-cobalamin, an unnatural cobalamin form which is the result of its chemical extraction method. The two biologically active forms of cobalamin in humans contain adenosyl- and methyl- upper ligands. Throughout this thesis, B_{12} and cobalamin are used interchangeably and refer to all these cobalamin forms, unless specifically stated otherwise.

1.2 Historical background

Like so many other ground-breaking scientific discoveries, that of vitamin B_{12} has also arisen serendipitously. George Whipple was the director of the Hooper Foundation for Medical Research at the University of California. In his study of the role of liver in haematopoiesis, he was feeding exsanguinated dogs a variety of diets, including what was meant to have been cooked liver. He noticed that the lab technician was feeding the dogs raw liver instead, which resulted in a much more dramatic response. This realisation proved crucial in determining the active liver principle (Whipple & Robscheit-Robbins, 1925). Based on the results of this study, George Minot and William Murphy, two Boston physicians, tested the effect of a diet including liver and meats as a treatment for pernicious anaemia. The results detailing an increase from 1% to 8-15% red blood cells within two weeks - including a literary review referencing Whipple's work - was published in their 1926 article in the *Journal of the American Medical Association* (Minot & Murphy, 1926). Though cautious of the findings themselves, the results were soon confirmed by many physicians throughout the world. In 1934, Minot, Murphy and Whipple were awarded the Nobel prize for physiology and medicine. As a diet of raw liver is hard to take, extracts were prepared for intramuscular injection which would be the means of treating pernicious anaemia until the 1950s (Banerjee, 1999). The extraction and purification of the actual molecule had proved difficult with conventional methods. At the Merck laboratories, Folkes and his associates introduced alumina column chromatography which they used to further purify the anti-pernicious anaemia factor. Simultaneously, Lester Smith, who

worked at the Glaxo Laboratories in the UK, worked on the purification and crystallisation of the nutrient. Both groups succeeded in isolating the active liver ingredient in 1947, describing a red crystalline compound which was subsequently named vitamin B_{12}. The findings were reported in the April issue of *Science* (Rickes et al., 1948) and *Nature* (E. L. Smith, 1948) in 1948. This launched research into the structural elucidation of the compound which proved to be a much greater task than imagined, due to the inherent complexity of the molecule.

It took the outstanding work of Dorothy Hodgkin and her group to elucidate the three-dimensional structure of vitamin B_{12} for which she was awarded the Nobel prize in 1964. She had met Lester Smith at a biochemical meeting at Oxford, two weeks after his publication on vitamin B_{12} in *Nature*. He brought his first preparation of the crystalline B_{12}, which he asked her to take some X-ray photographs of. This was not his main aim for the visit, as he wanted to get the refractive indices measured by someone else. It did however start a collaboration of a small Glaxo group sharing data, including Smith, Hodgkin, and Alexander Todd in an advisory role (Hodgkin, n.d.-a). She deduced the crystals were pleochroic which indicated a porphyrin ring structure (Hodgkin, n.d.-b). The knowledge of a central cobalt ion came two weeks later, and after calculating the electron density maps, the realisation came about that the corrin ring was not a true porphyrin. The final structure of B_{12} was published in 1956 (Hodgkin et al., 1955, 1956) (see Figure 1.1).

The biologically active coenzyme adenosylcobalamin (AdoCbl) was first crystallized by Barker *et al.* in 1958 (Barker et al., 1958) which led to the elucidation of the structure by Lenhert and Hodgkin in 1961 (Lenhert & Hodgkin, 1961). In the same year, it was found that this coenzyme B_{12} was the most predominant B_{12} found in the human liver (Toohey & Barker, 1961). In 1962, the other analogue of B_{12}, methylcobalamin (MeCbl), was synthesised by Smith and his group to investigate the role of B_{12} in methionine synthesis (Guest et al., 1962). It was soon discovered that this cofactor was the dominant form of B_{12} in human serum (Lindstrand & Ståhlberg, 1963).

Of course all these discoveries were major contributors to the overall scientific knowledge as well as important factors in treating the (no longer) pernicious anaemia. However, when telling the story of vitamin B_{12}, another magnificent feat which must be mentioned is the complete chemical synthesis of the molecule in 1973 (Eschenmoser, 1974). The groups of Robert Burns Woodward at Harvard and that of Albert Eschenmoser at ETH Zürich embarked on this task independently from each other. The ETH group started their work in December 1959 (Eschenmoser, 2015) with the Harvard group taking up the task two years later in August 1961. Since the work of the two groups was so

complementary to each other, the decision was made in 1965 to collaborate on the project of B_{12} synthesis. In the end, it took nearly 12 years for more than 100 researchers to achieve this incredibly complex synthesis which involves about 70 steps. As great as this achievement was, it was clearly not the way forward for producing the vitamin. The production history of vitamin B_{12} has been solely based on biosynthetic fermentation processes, using selected micro-organisms. The earliest of these B_{12}-producing strains was reported in the early 1980s. *Pseudomonas denitrificans* SC510, which after a decade of random mutagenesis and selection was reported to produce 50-100 mg per litre of culture (Martens et al., 2002).

1.3 Structure

The deep red vitamin B_{12} is often referred to as one of the "pigments of life"; a family of modified tetrapyrroles. Two of the most well-known members of this group are heme, giving blood its red colour and chlorophyll, responsible for the green colour in plants, algae and some bacteria. Further members include siroheme, coenzyme $_{F430}$, heme *d1* and the bilins (Layer et al., 2010). Of all the tetrapyrroles, however, cobalamin is by far the most complex. This is also reflected by the biosynthesis of the molecule, which is one of the most complex pathways found in nature. Figure 1.1 shows the relationships between the various biosynthetic pathways of this family of tetrapyrroles as well as their structures.

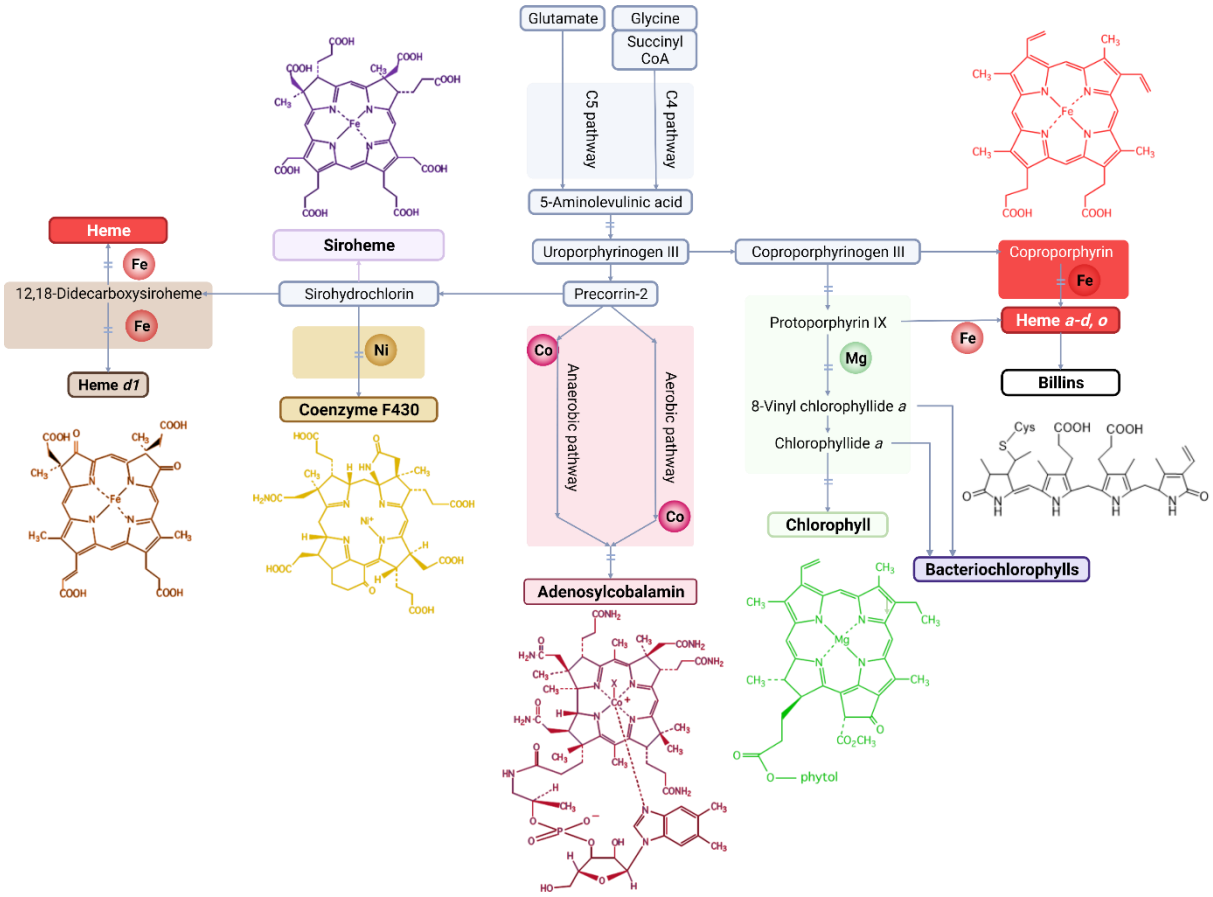


Figure 1.1 Schematic representation of the biosynthetic pathways of the pigments of life. *Showing the relationships between the tetrapyrrole biosynthetic pathways leading to heme (red), heme d1 (brown), coenzyme F430 (beige), siroheme (purple), B12 (pink), chlorophyll (green), bacteriophylls (navy) and billins (white). All pathways start with the synthesis of aminolevulinic acid (ALA) via the C4 or C5 pathway and proceed to synthesis the last common progenitor; uroporphyrinogen III. From here, branchpoints between the simplified pathways are shown, colour-coded from the first committed step into its pathway, as well as their final molecular products. The characterising metal for each tetrapyrrole is indicated at the individual pathways'chelation steps.*

The structure of B_{12} (Figure 1.2) can be considered as consisting of three parts; a central corrin ring, a lower (alpha) ligand and an upper (beta) ligand. The central cobalt ion is coordinated by the four nitrogen atoms of the four pyrrole rings as well as the upper and lower ligands. As mentioned above, vitamin B_{12} refers to cyanocobalamin, which has an upper CN-group as a result of the extraction procedure. Another commonly used dietary supplement form of vitamin B_{12} is hydroxocobalamin. In the two biologically active forms of B_{12} for humans, however, the upper ligands are actually a methyl group or an adenosyl group. The central corrin ring differs from its family of porphyrins in that it has lost the C-20 somewhere in the biosynthesis, causing the structure to have a contracted ring. This contracted ring structure is thought to be crucial to its function as energy seems stored in the helical strain of this central corrin ring (Kieninger et al., 2019). The central cobalt ion is coordinated by the four nitrogen atoms of the four pyrrole rings as well as the lower ligand. This lower ligand is a nitrogen from 5,6-dimethylbenzimidazole (DMB) in the cobalamin used by humans. There are many analogues based on variation of this lower ligand, used throughout the prokaryotic and eukaryotic branches of life (Kräutler, 2019; Kräutler et al., 1987; Stupperich et al., 1988, 1990; Stupperich & Kräutler, 1988). Many bacteria synthesize the cobalamin analogue pseudocobalamin, in which DMB is replaced for adenine (Barker et al., 1958, 1960). Pseudocobalamin is found in spirulina, bringing into question the nutritional value of these vitamin supplementation products (Watanabe et al., 1999).

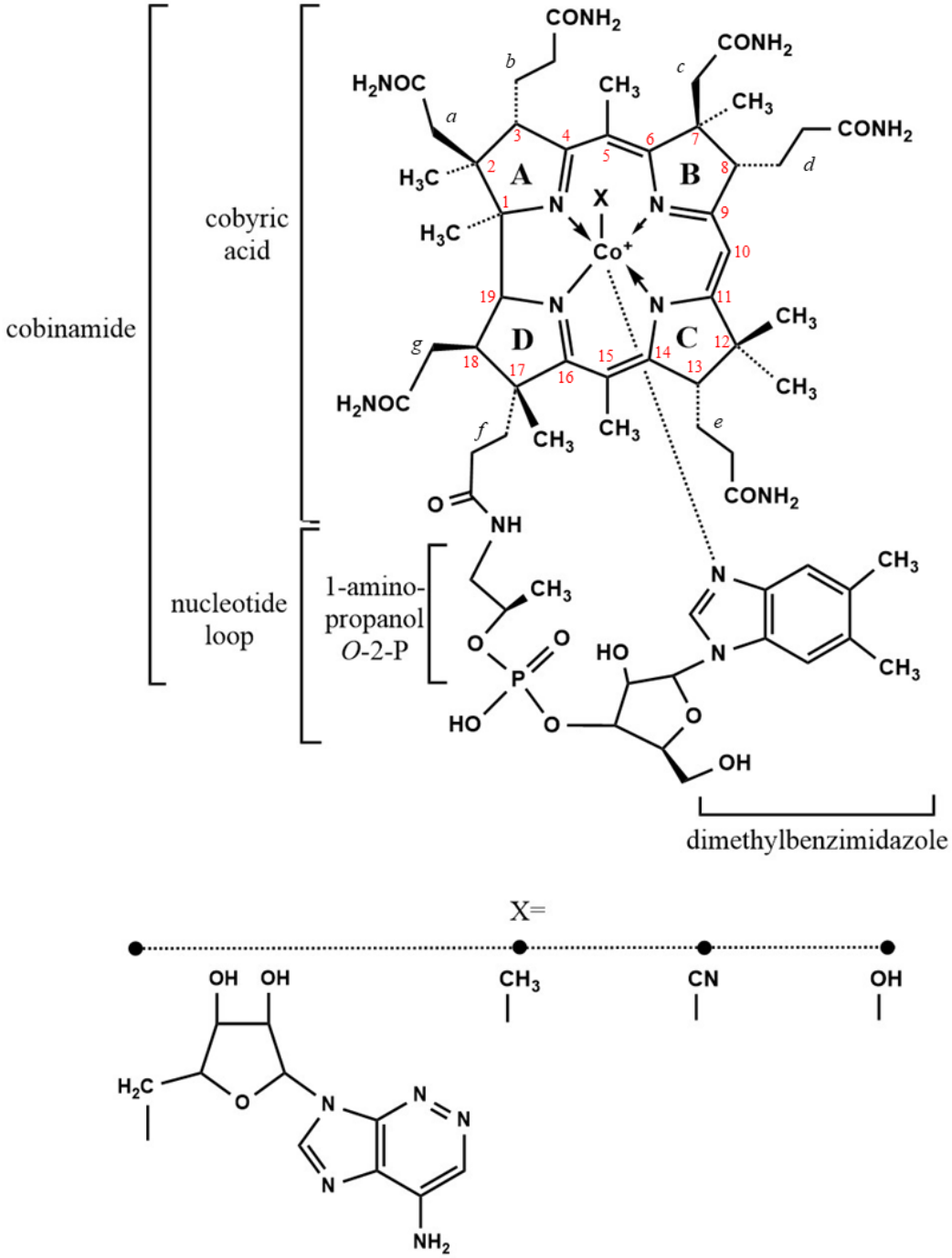


Figure 1.2 Structure of vitamin B_{12} and some of its derivatives. *True vitamin B_{12} refers to cyanocobalamin, where X is CN. In the main biological forms of cobalamin, this upper ligand is replaced with an adenosyl, methyl or hydroxyl group to form adenosylcobalamin, methylcobalamin or hydroxocobalamin, respectively. Numbering of the molecule is shown in red and the side chains are labelled in cursive.*

1.4 B_{12} - mediated reactions

The function of vitamin B_{12} has been classified in 3 enzymatic reactions. The isomerases, which were identified soon after the discovery of cobalamin, the methyltransferases, not far behind, and lastly the reductive dehalogenases (Banerjee & Ragsdale, 2003). The mechanism of this last class of B_{12}-dependent enzymes remained a field of study for many years, as they proved difficult to work with. An exciting fourth class of enzymatic reactions was identified in the form of radical SAM enzymes. Recently, B_{12} was also found to be involved in reactions where it acts neither as a cofactor nor as a coenzyme, but plays a direct role in the chemical reaction (Bridwell-Rabb & Drennan, 2017).

1.4.1 B_{12}- dependent isomerases

Soon after the identification of vitamin B_{12}, its involvement in many metabolic processes had been reported. The first direct enzymatic involvement in a rearrangement reaction was described simultaneously with the discovery of the coenzyme form (Barker et al., 1958). This became the first class of B_{12}– dependent reactions; isomerases. These enzymes use adenosylcobalamin as a coenzyme to facilitate complex rearrangement reactions. Isomerases are the largest group of B_{12}- dependent enzymes and are largely involved in fermentation reactions in bacteria. The unique isomerase which is used by bacteria as well as humans is methylmalonyl-CoA mutase, which facilitates the interconversion between methylmalonyl-CoA and succinylCoA. In some organisms, a B_{12}-dependent ribonucleotide reductase plays an important role in DNA replication and repair by catalysing the conversion of ribonucleotides to deoxyribonucleotides (Banerjee & Ragsdale, 2003). The functionality of these enzymes rest on the homolytic cleavage of the cobalt-carbon bond of Ado-Cbl, which results in the formation of a Co(II) species as well as an adenosyl radical. This adenosyl radical facilitates a hydrogen abstraction which initiates rearrangement chemistry (Banerjee, 1997).

1.4.2 B_{12}-dependent methyltransferases

The next class of B_{12}-dependent enzymes is that of the methyltransferases. Questions were raised in studying methionine synthesis when studies found there were two mechanisms wherein one was B_{12} dependent. Up till then, the only known biologically active form of B_{12} was adenosylcobalamin. The specific methylation reaction was however clearly very different from any isomerase reaction. This in fact led to the discovery of methylcobalamin

as a cofactor form of B_{12} and its subsequent function as cofactor to methionine synthase (Guest et al., 1962).

As the name implies, the reactions catalysed by B_{12}-dependent methyltransferases involve the transfer of a methyl group from methylcobalamin to an acceptor molecule. The most extensively studied B_{12}-dependent methyltransferase is methionine synthase from *Escherichia coli* (MetH) (Banerjee & Ragsdale, 2003; Matthews et al., 2008). Here, methylcobalamin transfers its methyl group to homocysteine, yielding methionine. The cofactor itself regains its methyl group from methyltetrahydrofolate, forming tetrahydrofolate (Drennan et al., 1994; Fujii & Huennekens, 1974). This interaction with folate explains why a folate deficiency so closely resembles the symptoms of a Vitamin B_{12} deficiency. The mechanism of methionine synthase is dependent on the ability of the cobalt ion in B_{12} to alternate between Co(III) and Co(I). As the cobalt ion is reduced to Co(I), it forms a strong nucleophile which forms methylcobalamin. Occasionally, Co(I) undergoes oxidative inactivation to Co(II). Methionine synthase has a separate active site where reductive reactivation of Co(II) is facilitated (Drummond et al., 1993). In this step, AdoMet is the methyl donor. Interestingly, nature has evolved a cobalamin-independent methyltransferase which is found in many different organisms, including yeast and plants. However, the selection pressure retains the B_{12}-dependent enzymes as they have a significantly higher catalytic activity.

Additional methyltransferases are described for bacteria and archea which play a role in methanogenesis, CO_2 fixation and acetogenesis (Ragsdale, 2008).

1.4.3 Reductive dehalogenases

Detailed studies on reductive dehalogenases have proven difficult as they are often membrane-bound and highly oxygen-sensitive. Recent studies have shone light on the reaction mechanism of these enzymes as well as the role which B_{12} plays in them (Bommer et al., 2014; Payne et al., 2015). Unique to this class of enzymes, cobalamin assists reductive dehalogenation by forming a cobalt-halogen bond. Three different types of reaction and electron transfer mechanisms have been described elsewhere (Bridwell-Rabb & Drennan, 2017; Fincker & Spormann, 2017). The importance of these enzymes is highlighted by their ability to neutralise a myriad of common pollutants in soil and water which are a result from improper handling and disposal of industrially produced organohalides.

1.4.4 Radical SAM enzymes

This last, relatively new class of B_{12}-dependent enzymes is that of the cobalamin-dependent radical S-adenosylmethionine (SAM) enzymes (Bridwell-Rabb & Drennan, 2017). Radical SAM enzymes are characterised by their use of an iron-sulfur cluster to reductively cleave SAM to generate a radical intermediate such as the 5'deoxyadenosyl radical, which abstracts a hydrogen atom from the substrate to initiate a radical mechanism (Frey et al., 2008; Holliday et al., 2018; Sofia et al., 2001). These enzymes are involved in different reactions such as cofactor biosynthesis, enzyme activation, peptide modification, post-transcriptional and post-translational modifications, metalloprotein cluster formation, tRNA modification, lipid metabolism, biosynthesis of antibiotics and natural products. With over 7,000 radical SAM enzymes now annotated as Cbl-dependent, these enzymes have emerged as a superfamily (Bridwell-Rabb et al., 2017). In addition to the AdoMet radical [4Fe-4S] cluster-binding motif, the members of this superfamily of enzymes have an additional N-terminal binding domain for cobalamin. As with Ado-Cbl, this [4Fe-4S] cluster-ligated molecule of AdoMet is used to catalyse the formation of 5'-dAdo.

1.4.5 Non-enzymatic functions

Some non-enzymatic functions which have been described for B_{12} are involved with transcription and translation regulation of B_{12} biosynthesis, transport, or B_{12}-related enzymatic reactions. One way B_{12} is involved in transcription and translation is via riboswitches; non-coding regions of mRNA which are able to bind metabolites. This binding causes conformational changes in the mRNA which affects gene-expression processes such as transcription termination and translation initiation (Mandal & Breaker, 2004). The *btuB* gene in *Escherichia coli* for instance, is responsible for uptake of Ado-Cbl and it has been known for some time that the coenzyme itself is involved in feedback inhibition to control the expression of these genes. In their detailed study of riboswitches, Mandal and Breaker have shown this to be via binding to a 5'UTR riboswitch of *btuB* (Mandal & Breaker, 2004). A similar riboswitch regulation mechanism is described for the *cob* and *pdu* genes of *Salmonella typhimurium* which are involved in the synthesis of vitamin B_{12} (Nahvi et al., 2004; Raux et al., 1996).

A more recent exciting mechanism has been described by *Jost et al.* where B_{12} acts as a photoreceptor (Jost et al., 2015). The CarH transcription factor makes use of the light-sensitive cobalt-carbon bond in Ado-Cbl. When bound to Ado-Cbl, CarH forms a tetramer which represses the transcription of carotenoid biosynthesis. Once the cobalt-carbon bond is severed after being exposed to light, the tetramer disassembles and leaves the DNA (Bridwell-Rabb et al., 2017; Jost et al., 2015).

1.5 Biosynthesis of cobalamin

The requirement for vitamin B_{12} is spread amongst a vast number of organisms, including humans, as cobalamin-dependent enzymes are essential in metabolic processes. In fact, plants and fungi are the only eukaryotes who seem to have become independent although it has recently been reported that certain fungi appear to be able to utilise the nutrient (Orłowska et al., 2021). It is interesting then that the biosynthesis of B_{12} is exclusive to certain prokaryotes. This might be explained by the energy cost of the production of the highly complex molecule, which takes about 30 enzymatic steps (Warren et al., 2002). There are 2 different pathways to synthetize cobalamin *de novo*; the "aerobic" and the "anaerobic" routes (Figure 1.3). There is an additional salvage pathway which enables certain bacteria to take up an intermediate and finish the biosynthesis of whichever form of cobalamin they require (Escalante-Semerena, 2007).

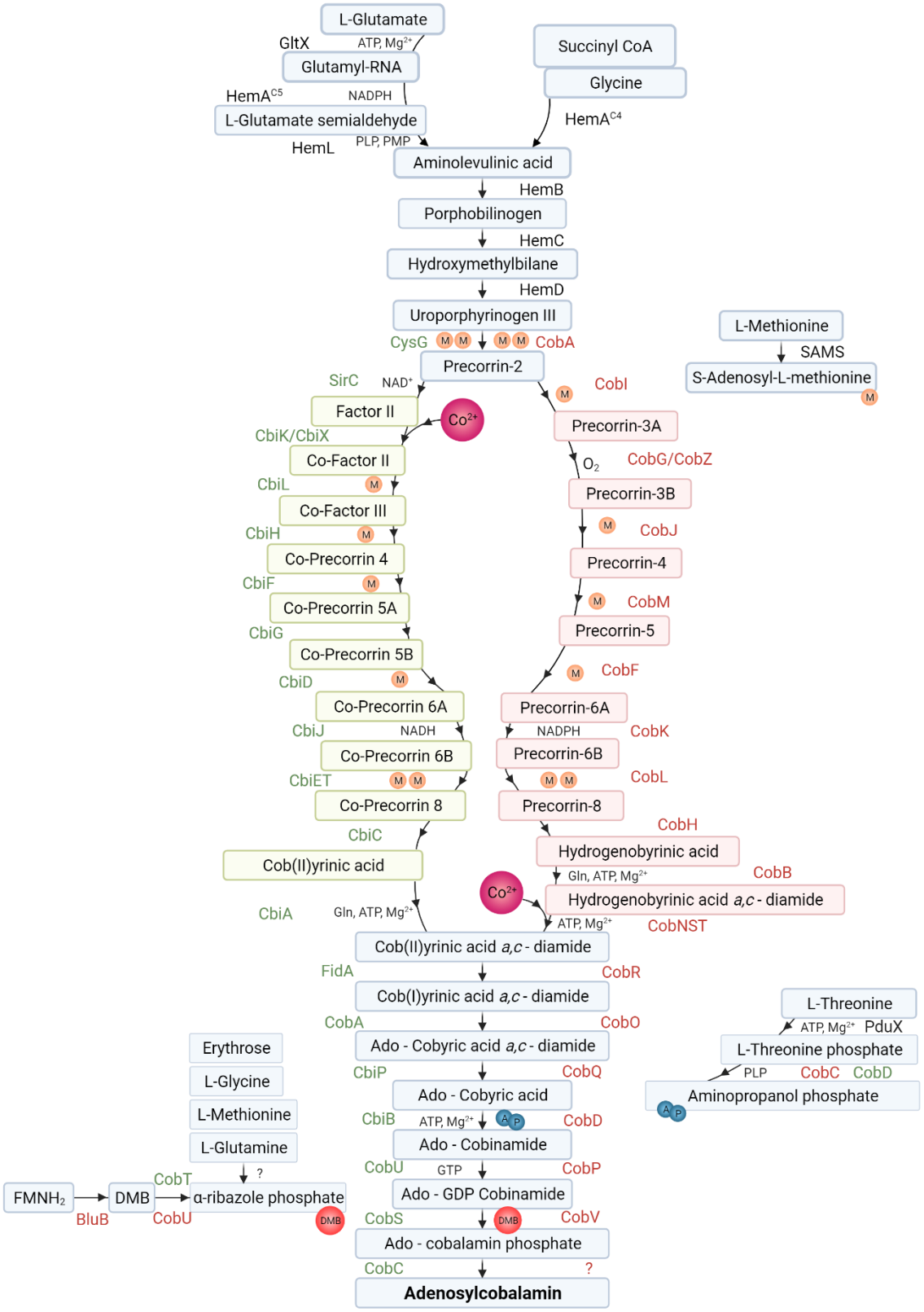


Figure 1.3 Anaerobic (green) and aerobic (red) cobalamin biosynthetic pathways. *Cobalamin intermediates are represented in boxes, colour coded where these pathways are distinct from one another. Enzymes are also colour coded according to which pathway they are active in.*

1.5.1 Elucidation of B_{12} biosynthesis

The research into the biosynthesis of vitamin B_{12} started with the investigation of the pathway in the anaerobic *Propionibacterium shermanii*, as it was a relatively good producer (Martens et al., 2002). By 1982, all steps from cobyrinic acid to cobalamin had been elucidated (Friedmann & Cagen, 1970). The steps from 5-aminolevulinic acid (ALA) to the cobalamin primogenitor uroporphyrinogen III had also been well described. Additionally, the first two intermediates of the anaerobic pathway to be isolated were Factor II and III, which were thought to represent oxidised versions of Precorrin-2 and Precorrin-3 (Uzar et al., 1987). This left a "black box" between Precorrin-3 and cobyrinic acid in which a number of methylations, amidations, decarboxylations, ring contraction as well as cobalt insertion needed to take place. It was not until Rhone-Poulenc Rhorer (ROR) started investigating into the biosynthetic pathway of an aerobic B_{12} producer, that light was shed on these individual steps. They named *Pseudomonas denitrificans* as their model organism, however, the strain is more likely to belong to the α-proteobacteria. With contributions of Battersby and Scott, the complete biosynthetic pathway was elucidated in the 1990s (Battersby, 1994, 2007; Blanche et al., 1998). This breakthrough was made possible in part by the application of molecular genetics as well as the isotopic labelling of precursors. It shifted the attention from the intermediates to the enzymes facilitating the intermediate reactions. The joint work of the French groups resulted in the isolation of 22 genes involved in the production of cobalamin (Blanche et al., 1989, 1990; Blanche, Couder, et al., 1991; Debussche et al., 1991, 1992). They were therefore given the prefix *cob* and the letter which follows indicates the order in which they are found on the operon.

Subsequently, the functions of the enzymes encoded by the *cob* genes were characterised (Blanche et al., 1990; Blanche, Couder, et al., 1991; Blanche, Thibaut, et al., 1992; Debussche et al., 1992; Thibaut et al., 1992). The oxygen requirement for the synthesis came as some surprise as many of the intermediates are oxygen sensitive, as well as the knowledge that many anaerobic organisms shared the ability to produce the vitamin. It was therefore thought that an alternative, aerobic-independent pathway must also exist. Not much time passed before these suspicions were confirmed when the biosynthetic genes of the anaerobic *Salmonella enterica* and *Bacillus megaterium* were also isolated and identified (Raux et al., 1996; Raux, Lanois, Rambach, et al., 1998; Raux, Lanois, Warren, et al., 1998; Roth et al., 1993). These genes were given the prefix *cbi* so as to distinguish them from the genes involved in the anaerobic pathway. However, the genes involved at the end of the aerobic pathway – from cobinamide to cobalamin – are also named *cob*, making the notation somewhat confusing. The main differences between the two pathways

are the point of cobalt insertion, which happens relatively early in the anaerobic pathway, whereas this is the last step in the aerobic pathway, before the two converge again at cobyrinic acid *a,c*-diamide. The second, as the names imply, is a difference in oxygen-dependence. The aerobic pathway requires oxygen to facilitate ring contraction, where the anaerobic pathway is not dependent on molecular oxygen for this step. A more correct denomination would therefore be oxygen-dependent and oxygen-independent pathway, or better yet; the late – and early cobalt insertion pathways, respectively.

1.5.2 Biosynthesis of 5-aminolevulinic acid

The biosynthesis of all tetrapyrroles all start with 5-aminolevulinic acid (5-ALA), which is produced via either the C4 pathway, also known as the Shemin pathway, or the C-5 pathway (Figure 1.4). The lesser used C-4 pathway is found in animals, fungi and certain classes of bacteria and sees the formation of ALA from succinyl-CoA and glycine. This reaction is catalysed by ALA synthase (ALAS) [E.C 2.3.1.37)] which was initially isolated by the groups of Shemin and Neuberger from photosynthetic bacteria and chicken erythrocytes, respectively (Laver et al., 1958; Shemin & Kikuchi, 1958; Shemin & Rittenberg, 1945). ALAS requires the cofactor pyroxidal 5'-phosphate (PLP) which is responsible for activation of glycine (Astner et al., 2005). The enzyme is encoded for by the *hemA* gene and has shown to be highly specific for glycine, not accepting any other amino acids (Ferreira & Gong, 1995). Transcription of ALAS is highly regulated by negative feedback of the end product of the heme pathway (Fanica-Gaignier & Clement-Metral, 1973; Hungerer et al., 1995). The C-5 pathway, which is found in the majority of phototropic organisms including plants, algae and cyanobacteria as well as many other bacteria – including *E. coli*, tRNA-bound glutamate is converted to 5-ALA in a series of enzymatic reactions (Beale, 1990). The first reaction sees the charging of glutamate to $tRNA^{Glu}$ in an ATP and Mg^{2+}- dependent reaction which is catalysed by glutamyl-tRNA synthetase (GluRS) [E.C 6.1.1.17]. GluRS is encoded for by the *gltX* gene, which is how the enzyme is referred to in Figure 1.3. The resulting glutamyl-$tRNA^{Glu}$ is used both for ALA synthesis as well as protein synthesis (Gamini Kannangara et al., 1988; Randau et al., 2004). The glutamyl-tRNA complex is subsequently reduced by glutamyl-tRNA reductase (GluTR), forming glutamate-1-semialdehyde (GSA). Confusingly, the gene that encodes GluTR is also called *hemA*, the same name as the gene which encodes ALAS in the C4 pathway. For this reason, the HemA's (as the enzymes are also called) are referred to as $HemA^{C4}$ and $HemA^{C5}$ in Figure 1.3, based on the pathway which they belong to. Glutamyl-tRNA reductase appears to be a specific enzyme for its substrate as it has been shown to be

able to discriminate between glutamyl-$tRNA^{Glu}$ complexes from different organisms (Jahn et al., 1991). Based on modelling studies, it has been suggested that GluTR physically interacts with the final enzyme of this pathway; glutamate-1semialdehyde-2, 1-aminomutase (GSAM). It has been proposed that the two enzymes form a complex in which $tRNA^{Glu}$ and GSAM simultaneously interact with GluTR (Moser et al., 2001; Rieble & Beale, 1991). Experimental support of this interaction between the two enzymes has been provided by studies in *E. coli* and *D. reinhardtii* (Lüer et al., 2005; Nogaj & Beale, 2005). GSAM [EC 5.4.38] is an aminotransferase which then catalyses the final step of ALA synthesis. The transamination of GSA to form 5-ALA happens in the presence of pyridoxal 5'-phosphate (PLP) or pyridoxamine 5'-phosphate (PAP) (Jahn et al., 1992). A nuclear gene (*gsa*) encodes this enzyme in higher plants (Grimm, 1990). The corresponding gene in *E. coli*, *B. subtilis* and *S. typhimurium* is *hemL* (Elliott et al., 1990; Hansson et al., 1991; Ilag et al., 1991).

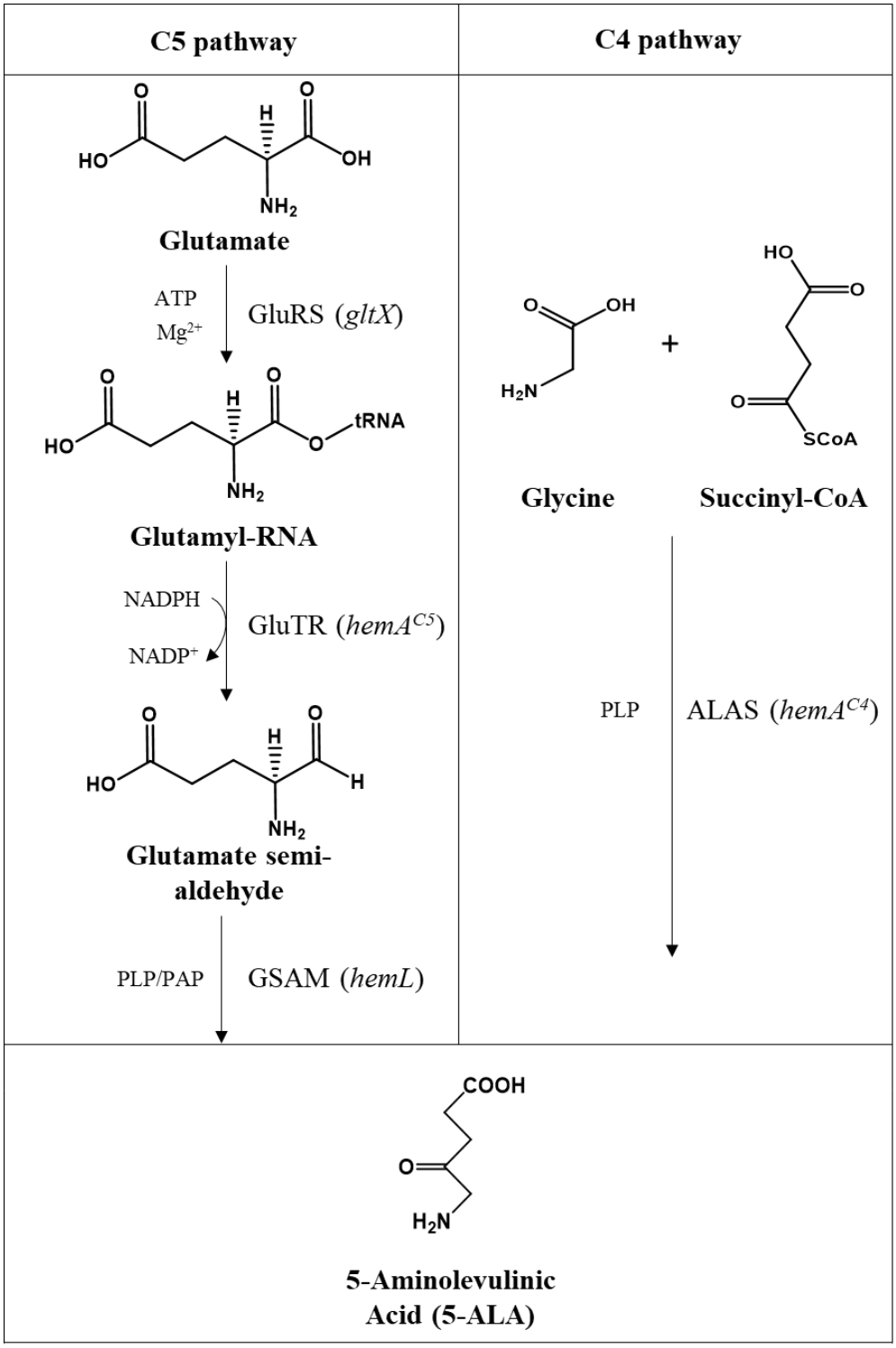


Figure 1.4 5-Aminolevulinic acid (5-ALA) synthesis via the C4 (Shemin) and C5 pathways. *The C5 pathway involves 3 enzymatic steps from glutamate to 5-ALA, whereas 5-ALA is the product of enzymatic condensation of glycine with succinyl-CoA via the C4 pathway.*

1.5.3 Conversion of 5-ALA into uroporphyrinogen III and the precorrin-2 branchpoint

The formation uropophyrinogen III (uro'gen III) from 5-ALA follows three enzymatic reactions shown in Figure 1.5. Condensation of two ALA molecules form porphobilinogen (PBG), the first pyrrolic biosynthetic intermediate. This step is catalysed by an enzyme commonly called porphobilinogen synthase (PBGS), but is also known as ALA dehydratase (Gibson et al., 1955). Four molecules of PBG are then incorporated into the linear hydroxomethylbilane (HMB) which is catalysed by HMB synthase (HMBS) and involves deamination of each substrate before forming the polymerized HMB product (Battersby et al., 1982). Uropophyrinogen III synthase (UROS) catalyses the final step in the synthesis of uro'gen III (Warren & Scott, 1990), which is the last common progenitor of the tetrapyrroles (E. Raux, Schubert, and Warren 2000). This step not only involves the cyclization of HMB, but also inverts ring D of the bilane. This is the branch point of the biosynthetic pathway towards cobalamin, siroheme, coenzyme F_{430} and heme d_1 and those resulting in the (bacterio)chlorophylls, hemes and bilins (Figure 1.1).

Methylation of uro'gen III at the C-2 and C-7 positions is catalysed by *S*-adenosyl-L-uro'gen III methyl transferase (SUMT) which results in precorrin-2 (Blanche et al., 1989; Crouzet et al., 1990), the last common intermediate for the synthesis of cobalamin, siroheme, coenzyme F_{430} and heme d_1 (A. D. Lawrence et al., 2008). In the organisms that use the aerobic pathway, the enzyme is also referred to as CobA, based on the gene which encodes the protein (Figure 1.6).

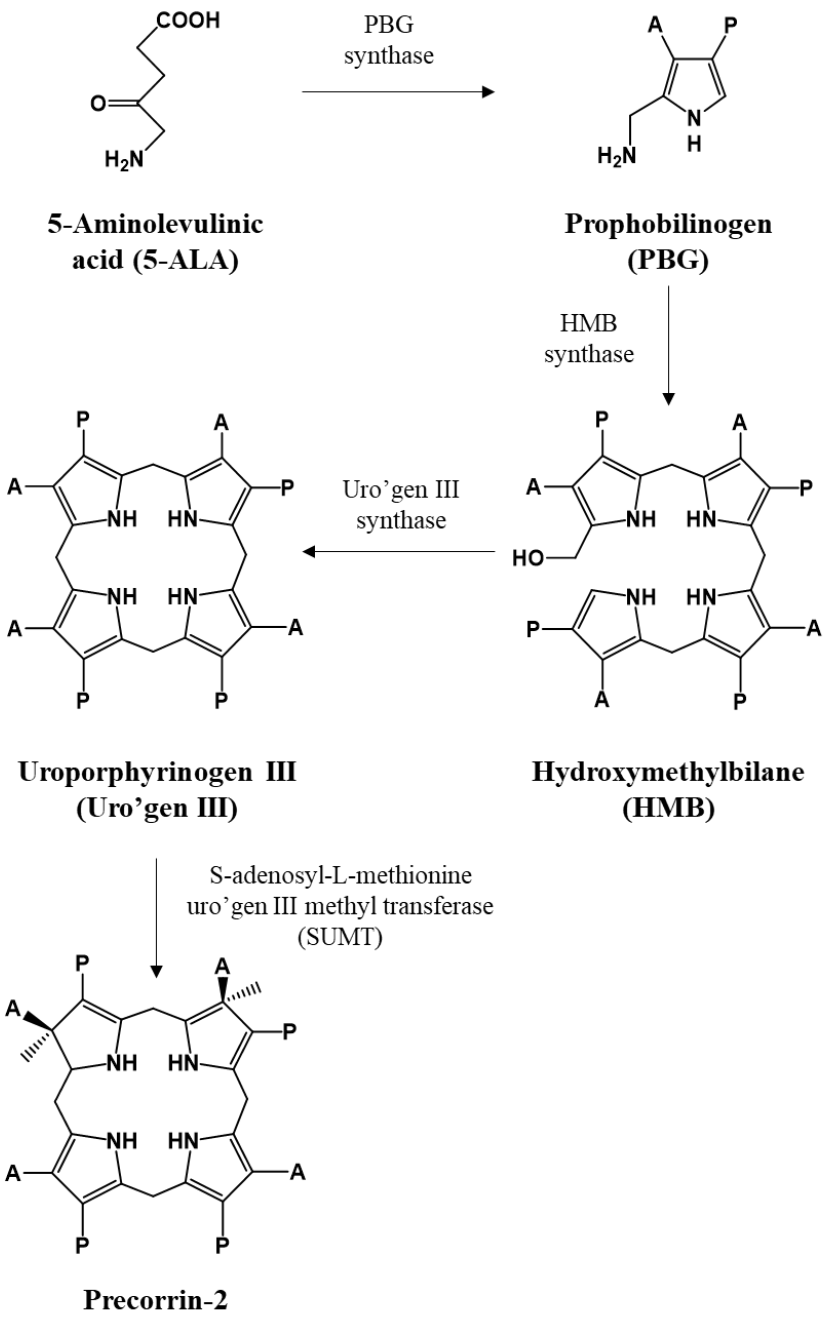


Figure 1.5 Biosynthesis of uroporphyrinogen III from 5-aminolevulinic acid. *Two molecules of 5-ALA are condensed into PBG in a reaction that is catalysed by HMB synthase. This is followed by deamination of four PBG molecules which are subsequently linked together to form HMB by HMB synthase. In the final step of uro'gen III synthesis, uro'gen III catalyses the cyclization of HMB as well as the inversion of ring D.*

1.5.4 The aerobic B_{12} pathway

Precorrin-2 represents not only the last branchpoint towards the different tetrapyrrole-derived macrocycles, but it also represents the point at which the 'aerobic' and 'anaerobic' pathways for cobalamin diverge. The separate pathway of uro'gen III towards cobalamin via the aerobic route is represented in Figure 1.6. The first committed step towards cobalamin biosynthesis is catalysed by CobI, the methyltransferase that is responsible for methylating the tetrapyrrole ring at the C-20 position to generate precorrin-3A (Thibaut et al., 1990). Next is a step that spring-loads the macrocycle in preparation for the ring-contraction, and is catalysed by CobG or CobZ, two orthologous proteins, depending on the organism. This reaction involves the insertion of a hydroxyl group at position C-20 of precorrin-3A, and the formation of a γ-lactone ring between C-1 and the acetate group of ring A in a reaction that is aided by iron-sulfur clusters on the enzymes (Leeper et al., 2012; Schroeder et al., 2009). The resulting precorrin-3B is now ready for ring-contraction, a process that is facilitated by CobJ, a bifunctional enzyme that additionally catalyzes the C-17 methylation, generating precorrin-4 (Debussche et al., 1993). Methyltransferases CobM and CobF are responsible for the methylations at positions C-11 and C-1, respectively, forming precorrin-6A. This second methylation step is preceded by the removal of the acetyl group at C-1 by CobF (Debussche et al., 1993). Precorrin-6B results from a reduction of the double bond between C-18 and C-19 in a reaction catalysed by CobK and is facilitated by the cofactor NADPH (Blanche, Thibaut, et al., 1992; Kiuchi et al., 1992). Precorrin-8 synthesis involves two methylations at positions C-5 and C-15 as well as the decarboxylation of the acetate side chain of ring C. All these reactions are catalysed by one multifunctional enzyme; CobL (Blanche, Famechon, et al., 1992).

CobH then catalyses the final enzymatic step which finalises the corrin macrocycle. This involves the migration of the methyl group at C-11 to the C-12 position (Thibaut et al., 1992). At this point the molecule is no longer a precorrin and is called hydrogenobyrinic acid (HBA) and the change in conjugation of the macrocycle is visible as an adjustment in colour from yellow to orange. Following the completion of the corrin ring, the HBA side chains *a* and *c* (C-2 and C-7, respectively), are amidated by CobB (Debussche et al., 1990). Cobalt insertion results in the formation of cob(II)yrinic acid *a,c*-diamide, and represents the point at which the aerobic and anaerobic pathways rejoin (Leeper et al., 2012). This chelation is facilitated by CobNST, a large enzyme complex of 590 kDa that is comprised of three subunits. The largest subunit at 140 kDa is CobN, which contains the active site where cobalt chelation occurs. The remaining two proteins, CobS and CobT, form a complex of 450 kDa with each forming a hexameric ring that are stacked upon each other.

This complex is responsible for ATP hydrolysis which powers the chelation reaction that takes place in CobN (Debussche et al., 1992).

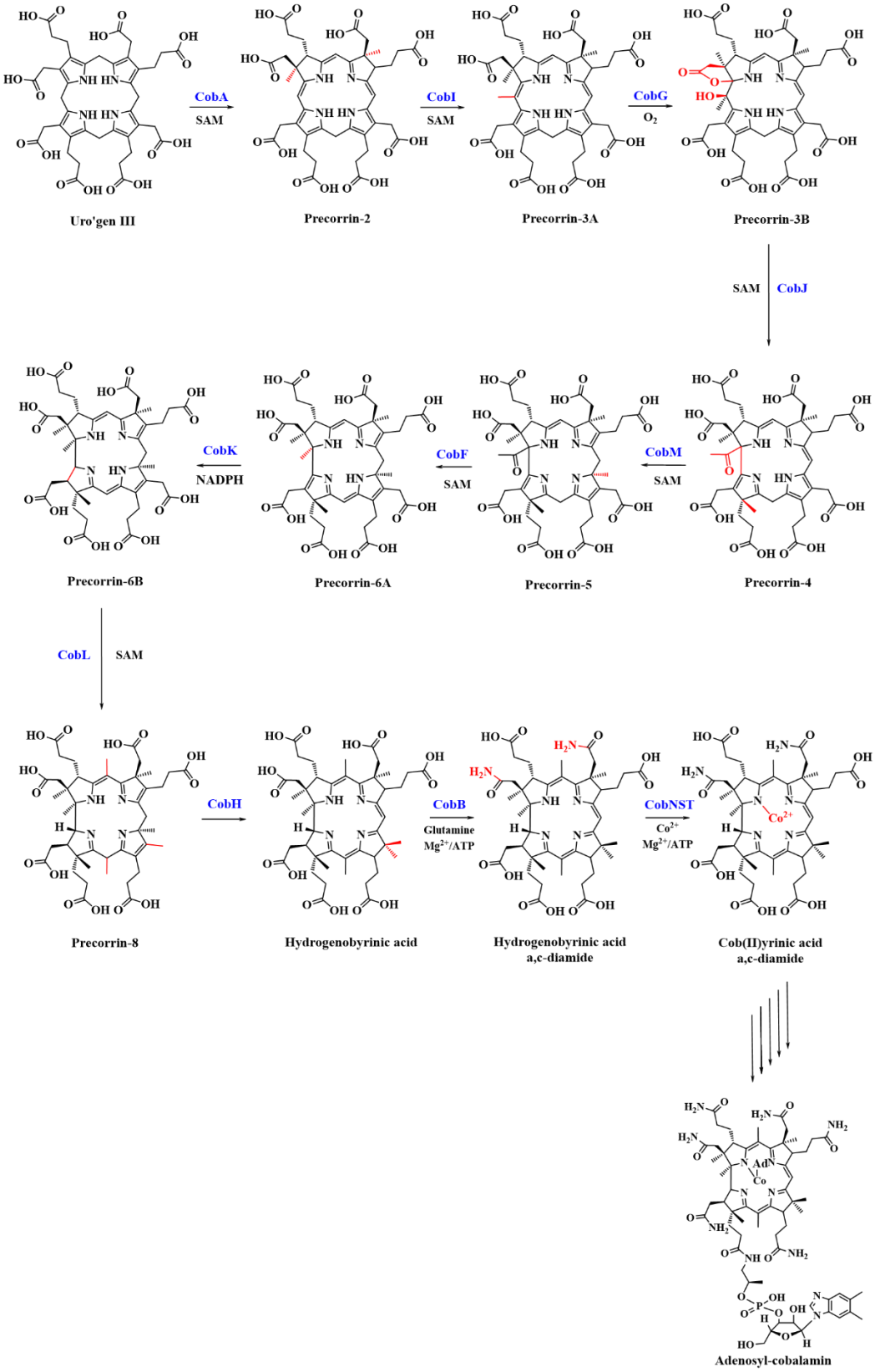


Figure 1.6 The aerobic biosynthetic cobalamin pathway. *The reaction steps are shown from uropophyrinogen III up to where the aerobic pathway converges with the anaerobic pathway, at cob(II)yrinic acid a, c-diamide, indicating the following steps from there result in the final product; adenosylcobyric acid. Added or affected groups in each step are shown in red.*

1.5.5 The anaerobic pathway

In contrast to the aerobic pathway, cobalt insertion occurs as the second step in the anaerobic pathway. This results in the generation of many unstable cobalto-complexed intermediates throughout this route. An oxidation of precorrin-2 precedes this step in an NAD-dependent reaction mediated by SirC, resulting in the generation of Factor II. The cobaltochelatase (CbiX or CbiK) subsequently catalyses the cobalt insertion resulting in cobalt(II)-Factor II (Moore & Warren, 2012). The following methylation step at C-20 is catalysed by CbiL, resulting in cobalt(II)-Factor III. CbiL has sequence similarity to CobI in the aerobic pathway, however, its function differs significantly in that it does not methylate precorrin-2, recognizing only the cobalt-complexed Factor II (Frank et al., 2007). The next enzyme in this pathway is the multifunctional CbiH which catalyses the methylation at C-17 as well as the formation of a δ-lactone ring as part of the ring contraction step. The sequence mechanism is not fully understood yet as CbiH seems to be able to catalyse the ring contraction with cobalt(II)-precorrin III as well as cobalt(II)-Factor III. This step also includes a reduction, forming cobalt(II)-precorrin-4. When the reduction takes place is uncertain. This is in part due to the product being hard to isolate and very unstable, readily converting back to cobalt(II)-Factor-IV (Leeper et al., 2012; Moore & Warren, 2012). Cobalt(II)-precorrin-4 is methylated at the C-11 position by CbiF, creating cobalt(II)-precorrin-5A, before the lactone ring is removed by CbiG, resulting in cobalt(II)-precorrin-5B. From here up to cobyrinic acid, a "black box" existed for many years, in which the order of intermediates was based on educated guesses, until 2013, when the final mysteries of B_{12} biosynthesis were elucidated. Moore *et al*. confirmed the methylation of cobalt(II)-precorrin-5B at the C-1 position by CbiD, resulting in cobalt(II)-precorrin-6A and the subsequent cobalt(II)-precorrin-6B formation by the NADH-assisted reduction of ring D. It goes on to describe the decarboxylation of the C-12 side chain and methylation at the C-15 position, catalysed by the dual-functional CbiT part of the fusion protein CbiET, resulting in cobalt(II)-precorrin-7. Cobalt(II)-precorrin-8 is formed after a methylation at C-5 by the CbiE moiety. CbiC catalyses the transfer of the C-11 methyl group to the C-12 position which finalises the synthesis of the macrocyle, generating cob(II)yrinic acid (Moore et al., 2013). The two pathways merge again after cob(II)yrinic acid is amidated on side chains *a* and *c* by CbiA (Fresquet et al., 2004).

1.5.6 Final part of the B_{12} pathway: Cobyrinic acid to Adenosylcobalamin

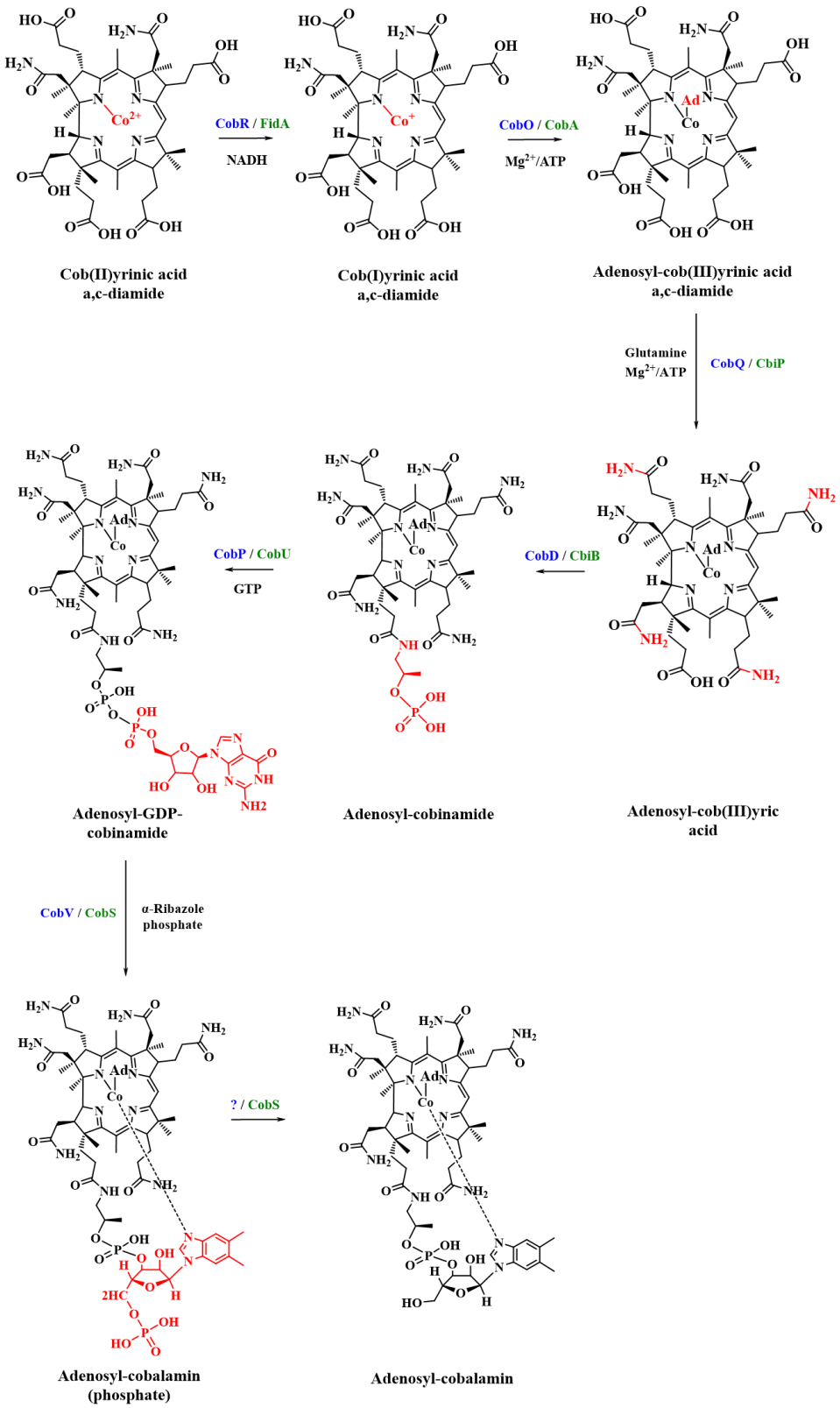


Figure 1.7 Steps from cobyrinic acid a,c diamide to adenosylcobalamin in the vitamin B_{12} pathway. *Enzymes catalysing the steps of the aerobic and anaerobic pathways are shown in blue and green, respectively.*

From the point where the aerobic and anaerobic pathways merge the molecule undergoes further amidations as well as the lower and upper ligand attachments to form the final product; cobalamin (Figure 1.7). Firstly, however, the central cobalt ion that up to this point was Co(II) is reduced to the Co(I) state. The enzyme responsible for this NADPH-dependent reaction has been identified as CobR in the aerobic pathway (Blanche, Maton, et al., 1992; A. D. Lawrence et al., 2008). A similar reaction is mediated by flavodoxin (FldA) in the anaerobic pathway. This reduction is an energetically unfavourable reaction creating cob(I)yrinic acid *a,c*-diamide, a very strong nucleophile, priming the molecule for the attachment of the upper ligand to the cobalt ion. This reaction, in which an adenosyl group is transferred from ATP, is facilitated by the adenosyltransferase CobO in the aerobic B_{12} producer *Pseudomonas denitrificans* (Debussche et al., 1991). The adenosyltransferase found in the anaerobic *Salmonella typhimurium* was called CobA (Suh & Escalante-Semerena, 1993) - not to be confused with the methyltransferase CobA found at the beginning of the aerobic pathway. In *S. enterica* the reduction of the cobalt is mediated either by flavodoxin or by reduced flavin when the substrate is bound to CobA, where the square planar binding of the substrate helps reduce the redox potential of the cobalt-corrin complex. In this way, CobA facilitates the reduction as well as the subsequent adenosylation reaction (Stich et al., 2005). Adenosylcob(III)yrinic acid *a,c*-diamide is then amidated on the *b*, *d*, and *e* propionic side chains as well as the acetic acid side chain *g*, resulting in adenosylcobyric acid (Blanche, Couder, et al., 1991). All four amidations are catalysed by a single enzyme; cobyrinic acid synthase which is encoded by *cobQ* in the aerobic pathway (Crouzet et al., 1991) and *cbiP* in the anaerobic pathway (Raux, Lanois, Rambach, et al., 1998). The next step, in which an aminopropanol phosphate group is attached to side chain *f* of adenosylcobyric acid, is the first in the construction towards the lower ligand that is characteristic of vitamin B_{12} as it provides the linker to the lower base. This aminopropanol phosphate group is derived from threonine in two enzymatic steps (Friedmann & Cagen, 1970). CobD in *P. denitrificans* and CbiB in *S. enterica* were identified as the enzymes responsible for the incorporation of the aminopropanol moiety, forming adenosylcobinamide (Blanche et al., 1995; Zayas & Escalante-Semerena, 2007). A guanosyl monophosphate group (GMP) is added to this linker to form adenosyl-GDP-cobinamide. This GTP-dependent step is catalysed by CobP in the aerobic *P. denitrificans* (Blanche, Debussche, et al., 1991) and CobU in the anaerobic *S. enterica* (O'Toole & Escalante-Semerena, 1995).

In the next step, this GDP group is exchanged for α-ribazole which contains the characteristic base 5,6-dimethylbenzimidazole (DMB) that provides the lower axial ligand to the central cobalt ion. There is a great variety in cobamide lower ligand bases, but the

only one which is relevant to humans is the DMB group that is found in vitamin B_{12} (Johnson & Escalante-Semerena, 1992). DMB can be synthetized via various biosynthetic routes. In aerobic conditions, the single-step oxidation of flavin mononucleotide (FMN) catalysed by BluB results in DMB (Gray & Escalante-Semerena, 2007). In the anaerobic *Eubacterium limosum*, DMB is constructed from 5'-aminoimidazole ribotide in five enzymatic steps (Hazra et al., 2015). DMB is then incorporated into α-ribazole by the nicotinate mononucleotide (NaMN):DMB phosphoribosyltransferase (PRTase), which, as the name implies, transfers the ribosyl-5'-phosphate moiety from NaMN to DMB. These enzymes were identified as CobU in the aerobic *P. denitrificans* (Cameron et al., 1991) and CobT in *S. enterica* (Trzebiatowski et al., 1994). The exchange of the GDP moiety for α-ribazole, forming adenosylcobalamin 5'-phosphate is catalysed by CobV (Cameron et al., 1991) and CobS (Maggio-Hall et al., n.d.) in the aerobic and anaerobic pathways, respectively. Finally, adenosylcobalamin is completed by the removal of the 5'-phosphate group. The timing of this step has been debated as both CobV and CobT recognise α-ribazole as well as α-ribazole-5'-phosphate as substrates. Higher cellular levels of α-ribazole were found within *P. denitrificans* than α-ribazole-5'-phosphate, which indicated the prior to be the true physiological substrate for CobV (Crouzet et al., 1991). In *S. enterica* however, α-ribazole 5'-phosphate appears to be a true intermediate and CobC has been identified as a phosphatase which recognises adenosylcobalamin 5'-phosphate as a substrate (Zayas & Escalante-Semerena, 2007).

Historically, organisms using the oxygen-dependent pathway are the main producers used for industrial cobalamin production as they are generally easier and cheaper to work with. Of these, *Pseudomonas denitrificans* is reported to be the highest producer. Microbes that use the anaerobic pathway are still interesting from a production point of view depending on their growth preferences and resource availability.

1.6 Biotechnological applications

Whilst there are many applications of vitamin B_{12}, the most important sectors of the vitamin B_{12} market are those of the feed industry as well as supplementation and pharmaceutical industry. In this section, these individual market branches will be explored and described more fully.

1.6.1 Pharmaceutical

As previously mentioned, vitamin B_{12} was initially linked to the devastating disease pernicious anaemia, which had an ultimately fatal result. It was believed to be the result of nutritional deficiency, but it has since been recognised that an autoimmune disease which attacks the parietal cells in the stomach can also give rise to the condition (Doniach et al., 1963; Taylor et al., 1962). Since the most common form of pernicious anaemia in developed countries was caused by this autoimmune disorder, pernicious anaemia is now referred to as an autoimmune disorder, which is somewhat misleading.

Absorption of the nutrient starts with the release of food-bound B_{12} from proteins by gastric acid, before being bound to R protein haptocorrin (sometimes referred to as transcobalamin I). This protein is secreted by the salivary glands as well as the gastric mucosa, protecting the vitamin from degradation by the acidic environment in the stomach (R. H. Allen et al., 1978). Further down the digestive tract in the small intestines, haptocorrin is degraded by pancreatic proteases, releasing the B_{12} to bind with intrinsic factor (IF). As a complex, IF/B_{12} binds to specific receptors found in the final section of the small intestine, causing internalisation of the complex into enterocytes. From here, B_{12} enters the bloodstream where it is bound to the plasma binding proteins haptocorrin and transcobalamin II (TCII). While haptocorrin binds approximately 80% of all B_{12} carried in the blood, it is TCII that delivers B_{12} to the tissues through specific receptors for TCII (Medicine, 1998).

Some patients suffer from an autoimmune disease that attacks intrisic factor (IF), thereby causing decreased vitamin B_{12} uptake (Toh et al., 1997). This disorder is called pernicious anaemia but should more accurately be called autoimmune metaplastic atrophic gastritis; the reason for this is that pernicious anaemia is a symptom resulting from autoimmune metaplastic atrophic gastritis. Nowadays, a multitude of causes are recognised for vitamin B_{12} deficiency. In higher-income countries, these include autoimmune metaplastic atrophic gastritis, as well as a vegetarian (or vegan) diet, gastrointestinal surgeries and certain medication. In lower-income countries B_{12} deficiency is largely a result of lower intake of the nutrient, or gastrointestinal infections (Green et al., 2017).

On a cellular level, the depletion of B_{12} results in a lack of cofactor MeCbl and coenzyme AdoCbl. As MeCbl is involved in methylating homocysteine, the upshot is reduced methionine synthesis and a build-up of its precursor, homocysteine (Hcy). A consequence of a lack of AdoCbl leads to an accumulation of methylmalonic acid (MMA) as it is involved in the conversion of methylmalonyl-CoA to succinyl-CoA. The neurological manifestations of vitamin B_{12} deficiency include neuropathy and demyelination of neurons. It has also been shown that vitamin B_{12} treatment improves cognitive function and plays a part in preventing dementia. Haematologically, B_{12} functions in the synthesis of red blood cells and prevents megaloblastic anaemia as well as hyperhomocysteinemia (L. H. Allen et al., 2018). It is important to note that the neurological manifestations can precede haematological changes, or even occur without any haematological symptoms (Lindenbaum et al., 1988). Even in patients with clinical pernicious anaemia, 28% do not have anaemia (Carmel, 2000). This would indicate that the haematological testing which is used to diagnose vitamin B_{12} deficiency is wildly inadequate. An alternative test was proposed in 2015 which would incorporate data on serum B_{12}, B_{12}-binding transporter protein holo-transcobalamin, and metabolites Hcy and MMA, to form a "combined B_{12} indicator", which would give a more comprehensive outcome (Fedosov et al., 2015). The practicality of such extensive testing as well as the interpretation of it needs more developing however.

The testing of MMA and Hcy have indicated there is a progression from normalcy to clinical deficiency which passes through a stage of inadequacy. The cut-off value of B_{12} deficiency is set at 148 pmol/L by the World Health Organisation but numerous studies suggest "pre-clinical" insufficiency levels below 200 pmol/L are linked to many undesirable outcomes (A. D. Smith et al., 2018). The identification of subtler degrees of deficiency adds to the idea that we are only seeing the tip of the iceberg of what might be attributed to B_{12} (pre-clinical) deficiency. In the pharmaceutical industry hydroxocobalamin (OHCbl) is used rather than cyanocobalamin (CNCbl) due to its higher uptake and a higher sustained serum level (Vandamme & Revuelta, 2016). For the same reason, MeCbl and AdoCbl are used in this industry. Another, albeit smaller, application of B_{12} in its hydroxocobalamin form is to treat acute cyanide poisoning. All forms of B_{12} in this industry are of pharmaceutical grade which is the highest purity, and are administered either intravenously or in pill form.

1.6.2 Supplementation

Even though vitamin B_{12} is made by enteric bacteria living in the gut, humans are dependent food of animal origin for the intake of the nutrient. The lack of vitamin B_{12} uptake from our own enteric bacteria is explained by their location as they mainly live in the large intestine, whereas the B_{12}-uptake mechanism is located in the terminal ileum of the small intestine (Stabler & Allen, 2004). The rise in vegetarian and vegan diets has contributed to the growth of the vitamin supplementation and food fortification subclasses of the vitamin B_{12} market.

1.6.3 Animal feedstock

The second major B_{12} market is that of animal feedstock. Nearly a billion heads of cattle are produced each year, (*Cattle/Cow Population Worldwide 2012-2020*, n.d.) making up a global market of 945.7 billion USD (Markets, 2019). Regardless of the rise of vegetarianism/veganism, the global meat market still grows annually as the global population increases. Although animals have the ability to absorb B_{12} which is synthetized by bacteria in the gut, animal feed fortification is described globally. The effect of vitamin B_{12} deficiency in cattle is described to result in poor appetite and growth, lacrimation, muscular weakness, demyelination of peripheral nerves and emaciation of the animals. How this deficiency is caused is unclear, but it is linked to occur during 'special conditions'. This will include age, as the prevalence is greater in young animals (Hogan et al., 1973; Lassiter et al., 1953), as well as stress such as is experienced during lambing/calving season (Duplessis et al., 2017). It is likely that a lack of cobalt in animal diets contributes to a secondary vitamin B_{12} deficiency. Albeit the more expensive alternative, it makes sense to add B_{12} to the feed rather than cobalt, which is toxic in its pure form.

As is the case with humans, there are subtler degrees of B_{12} insufficiency before showing the aforementioned symptoms. Vitamin nutrition for animals can be aimed solely to prevent deficiency signs, however, it is suggested to look at feed fortification to optimise the animals' health and product quality (McDowell, 2000). As no toxic level of vitamin B_{12} has been detected, fortification of animal feed has been suggested to only have beneficial effects (Winter & Mushett, 1950).

1.6.4 Biotech fermentation

The final contributor to the vitamin B_{12} market is that of biotech fermentation processes. Here, B_{12} is added to fermentation cultures, as it is an essential media additive for the metabolism of the production organism. One such process is the biological production of 1,3-propanediol, a component of emerging polymer business and terephthalic acid. The B_{12}-dependent glycerol dehydratase is part of the metabolic pathway from glycerol to 1,3-propanediol, rendering the addition of B_{12} to the fermentation batch a necessity (Dunn-Coleman et al., 2006; Nakamura & Whited, 2003).

Another process in which B_{12} is widely used, is the production of many pharmaceutics via cell culture. In the production of monoclonal antibody therapeutics, cyanocobalamin specifically is a vital media component (Schnellbaecher et al., 2019).

1.7 Market evaluation

The total available market (TAM) for vitamin B_{12} production stands at 280 USD (*Vitamin B12 Market Size Worldwide 2021 Forecast*, n.d.) and is reported to have an annual growth rate of 3.6% (Data,https://www.reportsanddata.com, n.d.). This total market can be split into the different applications of human consumption, further divided into nutritional food fortification, vitamin supplementation and direct medical application; animal feedstock; nd bioprocessing, in which fermentation batches are supplemented by vitamin B_{12} to assist the production of propane diol and pharmaceuticals. Whilst it is difficult the predict the exact submarkets' worth without freely available market reports, the majority of the market is reported to be occupied by the food grade segment at 63.8% (Data,https://www.reportsanddata.com, n.d.).

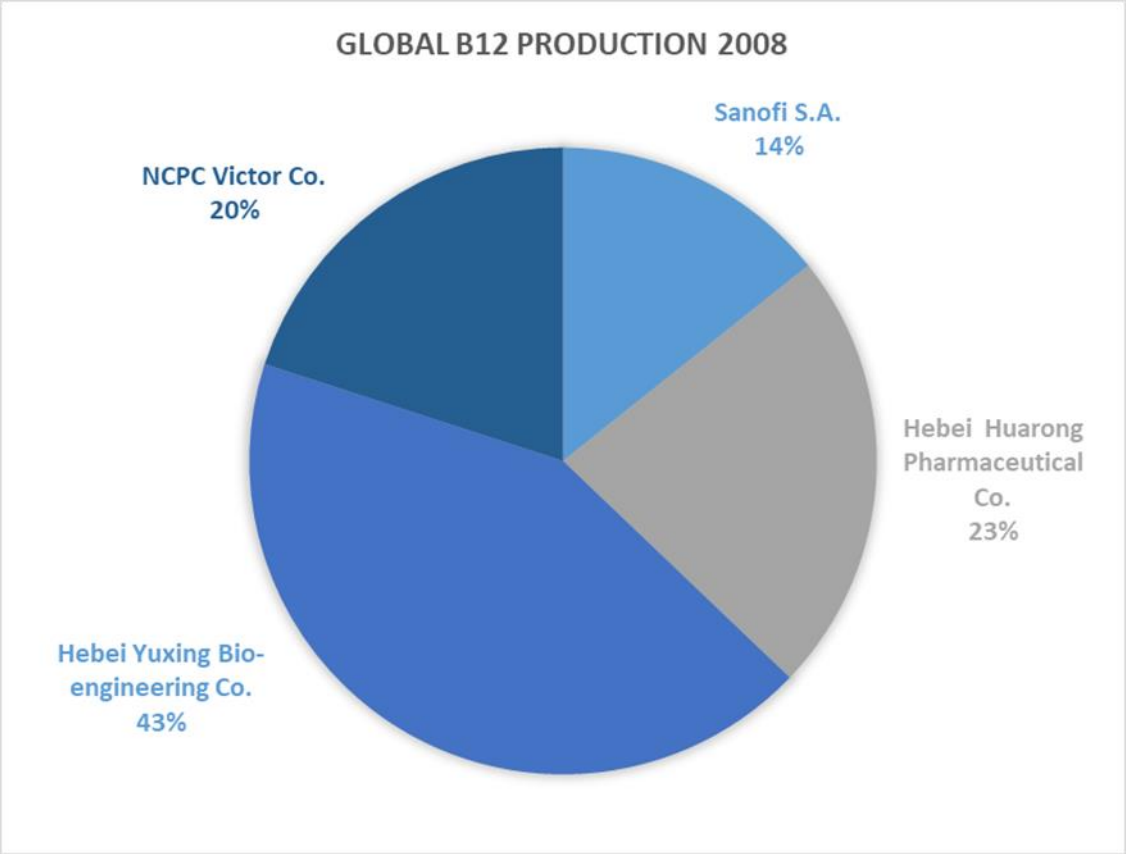


Figure 1.8 Market breakdown of global B_{12} production in 2008 based on Zhang Yemei China Chemical market report 2009.

After being major players in the elucidation of the oxygen-dependent pathway of B_{12}, Rhone-Poulenc continued to produce cyanocobalamin. The merge with German chemical company Hoechst in 1999 formed Aventis and has since become Sanofi-Aventis. To this day Sanofi-Aventis is the only European producer of cyanocobalamin. At present, the majority of global B_{12} production is achieved in China, by Hebei Huarong Pharmaceutical Co., Ltd, NCPC Victor Co., Ltd, and Hebei Yuxing Bio-Engineering Co.

A market monopoly has become evident with 86% of B_{12} production output located in China. As Aventis tends to focus more on the European market, its product is in a different price category, giving the Chinese producers some pricing power. Unfortunately, this has also given rise to major price fluctuations as companies have attempted to outcompete each other (Figure 1.9).

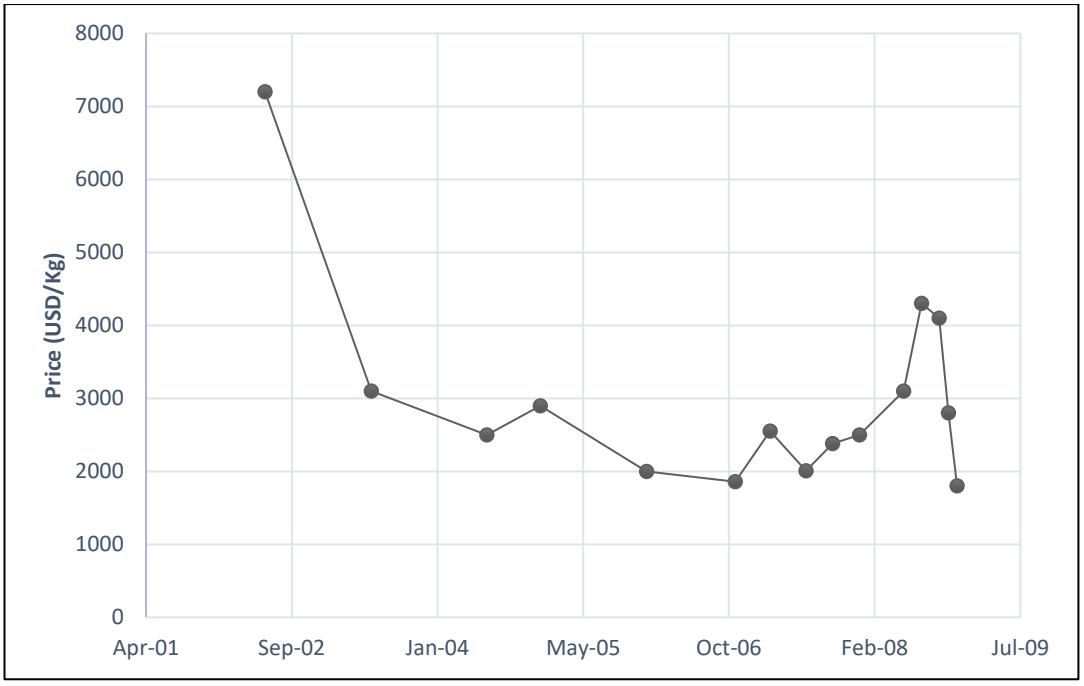


Figure 1.9 B_{12} price timeline from April 2001 up to June 2009, based on figures from Zhang Yemei China Chemical market report 2009.

The dip seen in 2005 correlates with the entrance of Hebei Yuxing Bio-Engineering Co. to the B_{12} production market, after which fierce price competition has kept the price of B_{12} around 2000 USD/Kg, a price which approaches the production costs (Y. Zhang, 2009). In more recent years, anti-pollution factory closures in China have caused major rises in B_{12} prices, from 2700 USD/kg in July 2017 to 13,000 USD/kg by January 2018. Whilst the price has since decreased to 4000 USD/kg, this has highlighted to need for stability in the market. With the increased costs of environment protection, labour, raw material and energy and power supply, production costs have risen and will likely continue to do so. The metabolic cost of B_{12} production is incredibly high with around 60 ATP per molecule of B_{12}. The media requirements are therefore higher than that of minimal media used in most fermentation processes. The high environmental costs are caused by the use of cobalt, a key component of cobalamin, and cyanide which is used to extract the B_{12}. This extensive chemical extraction procedure itself will take a week, adding to the overall cost.

Strains used in the fermentation of B_{12} conceivably produce around 250 mg per litre of culture. One batch of fermentation using these current strains will take a week, which, when using a 100,000 L fermenter will produce 25 kg B_{12} per week. At 50 weeks of production, 1.25 tons of B_{12} can be achieved each year. At 43% of the global output of 35 tons, Hebei Yuxing Bio-Engineering Co produced 15 tons of B_{12} in 2008 (Y. Zhang, 2009). Any attempt at entering the B_{12} market will require a drastically lower production cost either in

the form of cheaper media requirements or the use of a novel producing strain with a considerably higher yield.

1.8 Production improvements

Current production systems utilise a couple of strains which have been chosen for their naturally high vitamin B_{12} productivity and relatively rapid growth. Amongst these, the most commonly used strains are described to be *Pseudomonas denitrificans* and *Propionibacterium shermanii*. Using *P. denitrificans* has its clear benefits in that it is an aerobic organism which grows and produces in a cycle of 2-3 days. However, it does not have GRAS (generally recognised as safe) status.(Martens et al., 2002) *P. Shermanii* has this status, but requires a much more complicated growth procedure. As a microaerophilic organism, it will only produce B_{12} in high yields under anaerobic conditions. However, the synthesis of the DMB lower loop requires oxygen. The production is therefore split in two ways. In the first stage, the organism is grown anaerobically for 3 days to produce the cobamide intermediate. Afterwards, DMB synthesis is encouraged via gentle aeration to complete the molecule (Hargrove & Abraham, 1955). Improvements of these natural producers have been made in numerous ways, with the most common strategy being random mutagenesis. RPR combined this method along with genetic engineering to increase the yield of their *P. denitrificans* by 30%. This was achieved by amplification of 8 genes in the *cobF-cobM* operon, as well as the *cobA* and *cobE* genes via multicopy plasmids (Blanche et al., 1998). A similar approach has been made by DSM using *P. freundenii* (Pel & Hopper, 2010). Multiple studies have been made into additional production improvements of B_{12} by *Pseudomonas denitrificans*. Approaches which proved successful were dissolved oxygen control, pH control and continuous betaine feeding strategies (Li et al., 2012; Li, Liu, Chu, et al., 2008; Li, Liu, Li, et al., 2008).

Whilst being successful in increasing B_{12} production, the cost of production is unchanged or even increased by continuous feeding. To that effect, a study in 2015 describes the growth and B_{12} production of *P. denitrificans* on alternative, cheaper carbon sources such as beet molasses, maltose syrup and corn steep liquor, all industry by-products (Xia et al., 2015). Whilst these studies have been relatively successful, the maximum production levels for these organisms have been approached. An exciting alternative way forward in vitamin B_{12} production is the use of alternative organisms. A study from 2014 describes the cloning of genes required for B_{12} production into *Bacillus megaterium* and reaching a yield of 200 µg/L growing on cheap media (Moore et al., 2014). More recently, the attention has turned to *Escherichia coli* as a host for the production of valuable chemicals as it has a short

fermentation cycle and cheap media requirements. Additionally, *E. coli* is a model organism which has been studied extensively making it a convenient choice for metabolic engineering (Rosano & Ceccarelli, 2014). Indeed, *E. coli* has been used as a host for the heterologous production of many products, including terpenoids (Martin et al., 2003; C. Zhang & Hong, 2020), non-natural alcohols (K. Zhang et al., 2008), and biofuels (Fatma et al., 2018). However, in the era of synthetic biology the benefits of using *E. coli* cannot be overlooked. As a well-studied model organism, *E. coli* is a very convenient choice for metabolic engineering (Rosano & Ceccarelli, 2014). As such, it has been used as a host for the heterologous production of multiple products. These range from terpenoids (Martin et al., 2003; C. Zhang & Hong, 2020) to non-natural alcohols (K. Zhang et al., 2008) and biofuels (Fatma et al., 2018).

E. coli relies on either the uptake of intact corrinoids or the salvage of incomplete corrinoids for its B_{12} production; in the case of the latter *E. coli* is able to only to reconstruct B_{12} when provided with complex intermediates at or beyond cobinamide (J. G. Lawrence & Roth, 1995). The first use of *E. coli* to produce cobalamin elucidated the enzymes which made up the anaerobic pathway of *Salmonella typhimurium*. The genes of this operon were introduced on plasmids and *de novo* synthesis of B_{12} was established in *E. coli* under anaerobic conditions (Raux et al., 1996). The next major step in metabolic engineering of *E. coli* towards cobalamin production comes via research into the elucidation of the aerobic cobalamin pathway (Deery et al., 2012). Here, the cobalamin pathway up to the hydrogenobyrinic acid (HBA) intermediate is introduced to *E. coli* on plasmids. These plasmids were shared with the group of Zhang at Tianjin Institute of Industrial Biotechnology. Based on these, strain development resulted in the publication of a metabolically engineered *E. coli* strain which had the ability to produce cobalamin de novo to a yield of 307.00 μg g^{-1} cdw (Fang et al., 2018a). Only one recombinant *E. coli* strain had been reported previously which had the ability of expressing B_{12} biosynthetic genes from *P. denitrificans* ATCC 13867 under aerobic and anaerobic conditions (Ko et al., 2014). The genes were again introduced via plasmids and the yield was estimated to be 0.65 μg g^{-1} cdw. In parallel, continued developments in the Warren group has led to an *E. coli* strain that has the capacity to produce vitamin B_{12} to around 40 mg/L of culture. This strain was named ED656-3B and contains the genes for the aerobic vitamin B_{12} pathway housed within two artificial operons that had been integrated into the host's genome.

1.9 Aim

As mentioned above, previous work towards improved vitamin B_{12} production in the Warren group has led to an *E. coli* strain which houses the genes of the aerobic B_{12} biosynthetic pathway in the genome which has enabled the organism to produce the nutrient. In this strain, called ED656-3B, these genes are incorporated into two operons, called the "A-site", containing most of the B_{12} biosynthetic pathway from *Rhodobacter capsulatus*, except *cobG*, *cobC* and *cobE*, which originate from *Brucella melitensis*, and the "B-site" (Figure 1.10), which contains a second cobaltochelatase as well as *btuR*, *cobR* and a second *cobB* from *Brucella melitensis*. These two synthetic operons are controlled via T7 promoters. ED656-3B yields 40 mg vitamin B_{12} per litre of culture, however, to enter a competitive market, 1g/L needs to be approached. This describes the ultimate aim of the work presented in this thesis.

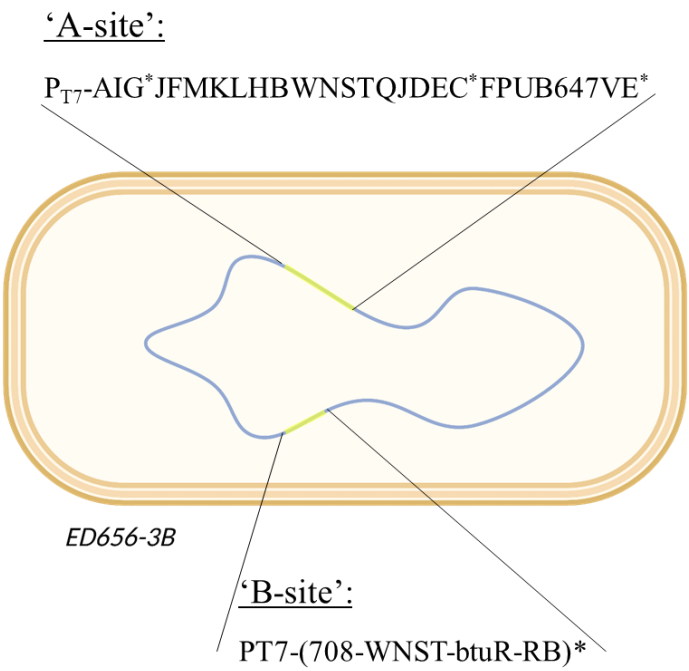


Figure 1.10 Schematic representation of cobalamin-producing *E. coli* strain ED656-3B. *The genes of the B_{12} biosynthetic pathway are housed in two operons, called the "A-site" and the "B-site".*

Improving the cobalamin yields involves three rational approaches, targeting the production in different ways. Firstly, the high metabolic cost associated with T7 expression has shown ED656-3B to shut off cobalamin production approximately 9 hours after induction. Placing the operons under the control of constitutive promoters of weaker, varying strengths, could alleviate this effect. Furthermore, enzymes within the pathway present targets for production improvement as they are involved in rate-limiting steps, as is the case for ALA synthesis by HemA. This also includes CobA and CobI, the latter of which also represents the first committed step in the (aerobic) B_{12} biosynthetic pathway. Similarly, genes that encode enzymes involved in the production of substrates for the B_{12} biosynthetic pathway present targets for improved yields. These include SAM synthetase and SAH hydrolase. Lastly, the cobaltochelatase enzyme that is responsible for inserting cobalt into the centre of cobalamin shows potential for improved catalytic turnover as the build-up of intermediate HBAD is seen in production cultures, despite the strain containing two genes encoding for cobaltochelatases. To this end, the enzyme, which in the pathway originates from *Brucella melitensis*, will be characterised. This includes the mutations of several residues that were identified as critical for catalytic turnover in closely related chelatases (Adams et al., 2020; J.-H. Zhang et al., 2021).

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Chemicals

Most chemicals and antibiotics were purchased from Sigma – Aldrich Ltd. Other materials used were purchased from manufacturers as follows: tryptone, yeast extract and bacteriological agar from Oxoid; ampicillin, IPTG and agarose from Melford Laboratories Ltd; chelating fast flow sepharose, disposable and empty PD10 columns from GE Healthcare; restriction enzymes from Promega and New England Biolabs, Inc; DNA extraction QIAprep® Spin Miniprep Kit, QIAquick® Gel Extraction Kit and Ni-NTA Spin Columns from QIAGEN, PCR Thermofischer, Roche FastStart High Fidelity PCR System from Roche Diagnostics GmbH.

2.1.2 Bacterial Strains

Bacterial strains were purchased from Novagen, Invitrogen or Promega. All bacterial strains used in this study can be found in Table 2.1.

Table 2.1: Bacterial strains used in this study. *Including genotype and/or genotype and origin*

Strain	Genotype and/or phenotype	Description	Reference
Escherichia Coli			
DH10β	F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lacZ*ΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ (*ara, leu*)7697 *gal*U *gal*K λ- *rps*L *nup*G /pMON14272 / pMON7124		Invitrogen
JM109	*end*A1, *rec*A1, *gyr*A96, *thi*, *hsd*R17 (r_{k-}, m_k^+), *rel*A1, *sup*E44, Δ(*lac-pro*AB), [F´ *tra*D36, *pro*AB, *laq*I^qZΔM15].		Promega
BL21 (DE3)	F- *ompT hsdSB (rB-mB-) gal dcm* (DE3)		Invitrogen
BL21 star (DE3)	F- *ompT hsdS*B (rB-mB-) *gal dcm rne131* (DE3)		Invitrogen
BL21 star (DE3) pLysS	*ompT hsdS*B (rB-mB-) *gal dcm rne131* (DE3) pLysS (CamR)		Invitrogen

BL21 star (DE3) pLysS-DNAJ	*ompT hsdS*B (rB-mB-) *gal dcm* *rne131* (DE3) pLysS (CamR) DnaJ	pLysS contains chaperone protein DnaJ	Dr Evelyne Deery
ED614	BGEC043 *fim (T7P)-AIG*JFMKLHBWN STQJDEC*FPUB64 7VE*ΔlacZYAΩ(T7RNAP)*	contains all the genes necessary to make cobalamin	Dr. Evelyne Deery
ED656-3B	ED614 with (T7Pro)- Bmei 708-cobWNST-btuR-cobR-cobB integrated with CRISPR in 5' of the *E. coli cobUST* to form a longer operon	contains all the genes necessary to make cobalamin	Dr. Evelyne Deery
ED661	*E. coli* with T7P-AIG*JFMKLH-AlvBQ-E*-lacZYA integrated in the fim operon + with the ΔlacZYAΩT7RNAP	Contains all the genes necessary to make HBAD and HBAH	Dr. Evelyne Deery
ED722A	*E. coli* with T7P-AIG*JFMKLHBWN STQJDEC*FPUB647VE* integrated in the fim operon + with the ΔlacZYAΩT7RNAP. Deletion of cobBWNST, insertion of Alv cobB and *Bmei* cobR-btuR instead.	Used for cobaltochelatase complementation	Dr. Evelyne Deery
RS01	ED656-3B where T7 is swapped to A2 in PT7-(708-WNST-btuR-RB)* ('B-site' promoter)	Genes from *R. capsulatus*, except *, which are from *B. Melitensis*	This study
RS02	ED656-3B where T7 is swapped to M2 in PT7-(708-WNST-btuR-RB)* ('B-site' promoter)	Genes from *R. capsulatus*, except *, which are from *B. Melitensis*	This study
RS03	ED656-3B where T7 is swapped to A2 in P_{A2}-AIG*JFMKLHBWNSTQJDEC*FPUB64 7VE*	Genes from *R. capsulatus*, except *, which are from *B. Melitensis*	This study

	(‘A-site’ promoter)		
RS05	ED656-3B where T7 is swapped to M2 in P_{A2}-AIG*JFMKLHBWNSTQJDEC*FPUB647VE* (‘A-site’ promoter)	Genes from *R. capsulatus*, except *, which are from *B. Melitensis*	This study
RS06_2	ED656-3B where T7 is swapped to A3 in PT7-(708-WNST-btuR-RB)* (‘B-site’ promoter)	Genes from *R. capsulatus*, except *, which are from *B. Melitensis*	This study
RS07_1	RS004_4 where T7 is swapped to M2 in PT7-(708-WNST-btuR-RB)* (‘B-site’ promoter)	Genes from *R. capsulatus*, except *, which are from *B. Melitensis*	This study
RS08	RS005 where T7 is swapped to A2 in PT7-(708-WNST-btuR-RB)* (‘B-site’ promoter)	Genes from *R. capsulatus*, except *, which are from *B. Melitensis*	This study
RS09_1	RS005 where T7 is swapped to A3 in PT7-(708-WNST-btuR-RB)* (‘B-site’ promoter)	Genes from *R. capsulatus*, except *, which are from *B. Melitensis*	This study
RS10_1	RS001 where T7 is swapped for A3 in in P_{A2}-AIG*JFMKLHBWNSTQJDEC*FPUB647VE* (‘A-site’ promoter)	Genes from *R. capsulatus*, except *, which are from *B. Melitensis*	This study
Salmonella typhimurium			
AR2680	MS1868 *Leu*$^{+}$ *Sm*^{r} *cbiB metE*	Used as indicator strain for	Raux *et al*. 1996

		quantitative cobinamide assays	
AR3612	MS1868 *Leu*$^{+}$ *Sm*^{r} *cysG metE*	Used as indicator strain for quantitative cobyric acid assays	Raux *et al*. 1996

2.1.3 Plasmids

All plasmids used in this study can be found in Table 2.2.

Table 2.2: Plasmids. *Sp: Streptococcus pyogenes, Bmei: Bruscella Meilitensis (also denoted with *), Rc: Rhodobacter capsulatus, Mbar: Methanosarcina barkeri, Pd: Pseudomonas denitrificans, Bt: Bacteroides thetaiotaomicron.*

Plasmid	Genotype and/or phenotype	Description	Reference
pET3a	Over-expression vector with T7 promoter, Amp^R		Novagen
pET14b	Cloning and expression vector; N terminal His_6 tag; T7 promotor; Amp^R		Novagen
pET14b BtuF	BtuF, Amp^R		Dr Evelyne Deery
pA2-HmCherry	pET3a where T7 is swapped for A2, *mCherry*	Adapted Tet promoter, strong expression levels	Dr. Rokas Juodeikis
pA3-HmCherry	pET3a where T7 is swapped for A3, *mCherry*	Adapted Tet promoter, medium expression levels	Dr. Rokas Juodeikis
pM2-HmCherry	pET3a where T7 is swapped for M2, *mCherry*	Adapted Tet promoter, low expression levels	Dr. Rokas Juodeikis
pCas9	Origin of replication: repA101ts, Kan^R, λ Red *exo*, *bet*, *gam* under araBAD promoter, gRNA scaffold, tracrRNA, *SpCas9* under lacIq promoter.	Expresses the cas9 protein which cleaves the target DNA in CRISPR-Cas9 cloning, Lambda red recombinase proteins needed for homologous recombination and RNA structures needed to target	Sheng Yang Addgene plasmid # 62225

		N20 and guide Cas9 protein.	
pTargetF	Origin of replication; pMB1, aadA^{R}, sgRNA containing an N20 sequence to target xxx replicon under *trc* promoter.	Introduces target DNA during CRISPR-Cas9 transformation.	Sheng Yang Addgene plasmid # 62226
pTarget-Afe	pTargetF with modified SgRNA region to include Afe restriction site for easier N20 cloning.	Introduces target DNA during CRISPR-Cas9 transformation.	Dr. Evelyne Deery
pTarget-N20A	pTarget-Afe, cloned with N20A target sequence. Afe restriction site is lost.	Cloned by PCR *SpeI* and *Afe* into pTarget-Afe Targets the A-site promoter.	This study
pTarget-N20B	pTarget-Afe, cloned with N20B target sequence. Afe restriction site is lost	Cloned by PCR *SpeI* and *Afe* into pTarget-Afe Targets the B-site promoter	This study
pTarget-N20 cobA	pTarget-Afe, cloned with N20 cobA target sequence. Afe restriction site is lost.	Cloned by PCR *SpeI* and *Afe* into pTarget-Afe Targets the A-site promoter.	This study
pET3a-*Bmei 708*	*Bmei 708*		Dr. Evelyne Deery
pET3a-*Rc cobA*	*Rc cobA*		Dr. Evelyne Deery
pA2-*Bmei 708*	A2 promoter, *Bmei 708*	*XbaI-BamHI* fragment of pET3a-*Bmei 708* into	This study

		Xba-BamHI of pA2-*HmCherry*	
pA3-*Bmei 708*	A3 promoter, *Bmei 708*	*XbaI-BamHI* fragment of pET3a-*Bmei 708* into *Xba-BamHI* of pA3-*HmCherry*	This study
pM2-*Bmei 708*	M2 promoter, *Bmei 708*	*XbaI-BamHI* fragment of pET3a-*Bmei 708* into *Xba-BamHI* of pM2-*HmCherry*	This study
pA2-*Rc cobA*	A2 promoter, *Rc cobA*	*XbaI-BamHI* fragment of pET3a- *Rc cobA* into *Xba-BamHI* of pA2-*HmCherry*	This study
pA3-*Rc cobA*	A3 promoter, *Rc cobA*	*XbaI-BamHI* fragment of pET3a- *Rc cobA* into *Xba-BamHI* of pA3-*HmCherry*	This study
pM2-*Rc cobA*	M2 promoter, *Rc cobA*	*XbaI-BamHI* fragment of pET3a- *Rc cobA* into *Xba-BamHI* of pM2-*HmCherry*	This study
pYoe-A2-*Bmei 708*	5'flanking region, A2 promoter, *Bmei 708*	Cloned by PCR *BglII* and *SalI* into pA2-*Bmei 708*. CRISPR-cas9 donor DNA	This study
pYoe-A3-*Bmei 708*	5'flanking region, A3 promoter, *Bmei 708*	Cloned by PCR *BglII* and *SalI* into pA3-*Bmei 708*. CRISPR-cas9 donor DNA	This study

pYoe-M2-*Bmei 708*	5'flanking region, M2 promoter, *Bmei 708*	Cloned by PCR *BglII* and *SalI* into pM2-*Bmei 708*. CRISPR-cas9 donor DNA	This study
pNanC-A2-*Rc cobA*	5'flanking region, A2 promoter, *Rc cobA*	Cloned by PCR *BglII* and *SalI* into pA2-*Rc cobA*. CRISPR-cas9 donor DNA	This study
pNanC-A3-*Rc cobA*	5'flanking region, A3 promoter, *Rc cobA*	Cloned by PCR *BglII* and *SalI* into pA3-*Rc cobA*. CRISPR-cas9 donor DNA	This study
pNanC-M2-*Rc cobA*	5'flanking region, M2 promoter, *Rc cobA*	Cloned by PCR *BglII* and *SalI* into pM2-*Rc cobA*. CRISPR-cas9 donor DNA	This study
pET3a-*Mbar cobA*	*Mbar cobA*		Dr Evelyne Deery
pET3a-*Bmei cobI*	*Bmei cobI*		Dr Evelyne Deery
pET14b-*Pd* $cobA^{K187A}$	*Pd* $cobA^{K187A}$	Point mutations at bp 618/619 from AAA to GCA , Lysine 187 to Alanine	Dr Evelyne Deery
pET3a-*Pd* $cobA^{K187A}$	*Pd* $cobA^{K187A}$	Cloned by PCR *NdeI* and *SpeI* into pET3a- *Mbar cobA* (*Mbar cobA* is removed)	This study

pET14b-*ySAMS*	*Yeast SAMS*	yeast SAM synthetase	Dr Evelyne Deery
pET28-*hSAMS*	*Human SAMS*	Formerly called pET28-hMAT2A, human SAM synthetase	Dr Evelyne Deery
pET28-*rSAH*	*Rat SAHH*	Rat SAH hydrolase	Dr Evelyne Deery
pET28-*hSAMS-rSAHH*	*hSAMS, rSAHH*		Dr Evelyne Deery
pET14b-*Rc hemA*	*Rc hemA*		Dr Evelyne Deery
pET-*Yoe-HisT-tetR*-PtetA	Ptet promoter, *tetR*, Amp^R		Dr Evelyne Deery
pET-tetR-*Mbar cobA*	Ptet promoter, *tetR, Mbar cobA*	*BglII – SpeI* fragment (tetR promoter) cloned into *BglII – XbaI* sites of pET3a- *Mbar cobA*	This study
pET-tetR-*Bmei cobI*	Ptet promoter, *tetR, Bmei cobI*	*BglII – SpeI* fragment (tetR promoter) cloned into *BglII – XbaI* sites of pET3a- *Bmei cobI*	This study
pET-tetR-*Mbar cobA -I**	Ptet promoter, *tetR, Mbar cobA, Bmei cobI*	*BglII – SpeI* fragment (tetR promoter) cloned into *BglII – XbaI* sites of pET3a- *Mbar cobA-I**	This study
pET-tetR-*Pd* $cobA^{K187A}$	Ptet promoter, *tetR, Pd* $cobA^{K187A}$	*BglII – SpeI* fragment (tetR promoter) cloned into *BglII – XbaI* sites of pET3a- *Pd* $cobA^{K187A}$	This study

pET-tetR- *Pd cobA*^{K187A}*-I**	Ptet promoter, *tetR*, *Pd cobA*^{K187A} *Bmei cobI*	*BglII – SpeI* fragment (tetR promoter) cloned into *BglII – XbaI* sites of pET3a- *Pd cobA*^{K187A}*-I**	This study
pET-tetR- *Rc hemA*	Ptet promoter, *tetR*, *Rc hemA*	*BglII – SpeI* fragment (tetR promoter) cloned into *BglII – XbaI* sites of pET14b- *Rc hemA*	This study
pET-tetR- *ySAMS*	Ptet promoter, *tetR, yeast SAM* synthetase	*XhoI – XbaI* fragment (*ySAMS*) cloned into *XhoI – SpeI* sites of pET-*Yoe-HisT-tetR*-PtetA	This study
pET-tetR- *hMAT2A*	Ptet promoter, *tetR*, *hMAT2A*	*XhoI – XbaI* fragment (*hMAT2A*) cloned into *XhoI – SpeI* sites of pET-*Yoe-HisT-tetR*-PtetA	This study
pET-tetR- *rat SAHH*	Ptet promoter, *tetR*, *rat SAH* hydrolase	*XhoI – XbaI* fragment (*rat SAHH*) cloned into *XhoI – SpeI* sites of pET-*Yoe-HisT-tetR*-PtetA	This study
pET-tetR- *rat SAHH-hMAT2A*	Ptet promoter, *tetR*, *rat SAHH-hMAT2A*	*XhoI – XbaI* fragment (*rat SAHH – hMAT2A*) cloned into *XhoI – SpeI* sites of pET-*Yoe-HisT-tetR*-PtetA	This study
pET3a-*Rc cobI*	*Rc cobI*		Dr Evelyne Deery
pNanC-T7-*Mbar cobA*	5'flanking region, *Mbar cobA*	*SalI-BglII* fragment of pNanC-A3-*Rc cobA* into *SalI-BglII* of pET3a-*Mbar cobA*	This study

pNanC-T7-*Mbar cobA-Rc cobI*	5'flanking region, *Mbar cobA, Rc cobI* (3'flanking region)	*XbaI-HindIII* fragment of pET3a-*Rc cobI* into *SpeI-HindIII* of pNanC-T7-*Mbar cobA*	This study
pNanC-T7-*Pd* $cobA^{K187A}$	5'flanking region, *Pd* $cobA^{K187A}$	*SalI-BglII* fragment of pNanC-T7-*Mbar cobA* into *SalI-BglII* of pET3a-*Pd* $cobA^{K187A}$	This study
pNanC-T7-*Pd* $cobA^{K187A}$ – *Rc cobI*	5'flanking region, *Pd* $cobA^{K187A}$, *Rc cobI* (3'flanking region)	*XbaI-EcoRI* fragment of pET3a-*Rc cobI* into *SpeI-EcoRI* of pNanC-T7-*Pd* $cobA^{K187A}$	This study
pET14b- *Bmei cobN*	N-terminally his-tagged *Bmei cobN*		Dr. Evelyne Deery
pET3a- *Bmei* $cobS^{H}T$	*Bmei cobS*, N-terminally his-tagged *Bmei cobT*		Dr. Evelyne Deery
pET14b- *Bt BtuG2_Nde3*	*BtBtuG2,* N-terminally his-tagged *Bt BtuG2*		Dr. M. Liang
pET14b- *Bmei* $cobN^{5933A}$	*Bmei* $cobN^{E593A}$	Point mutation at bp 1841 from GAG to GCG, Glutamic acid 593 to Alanine, template pET14b- *Bmei cobN*	This study
pET14b- *Bmei* $cobN^{H1094A}$	*Bmei* $cobN^{H1094A}$	Point mutations at bp 3343/3344 from CAC to GCC, Histidine 1094 to Alanine, template pET14b- *Bmei cobN*	This study
pET14b- *Bmei* $cobN^{D1089A}$	*Bmei* $cobN^{D1089A}$	Point mutation at bp 3356 from GAC to GCC, Aspartate 1089 to	This study

		Alanine, template pET14b- *Bmei cobN*	
pET14b- *Bmei cobN*^{Q1104A}	*Bmei cobN*^{Q1104A}	Point mutation at bp 3362 from GAT to GCT, Aspartate 1100 to Alanine, template pET14b- *Bmei cobN*	This study
pET-*tetR*- PtetA- *Bmei cobWNST*	Ptet promoter, *tetR* *Bmei cobW, cobN, cobS* and *cobT*		Dr. Evelyne Deery
pET-*tetR*- PtetA- *Bmei cobWN*^{E593A}*ST*	Ptet promoter, *tetR* *Bmei cobW, cobN*^{E593A}, *cobS* and *cobT*	*cobN*^{E593A} cloned into BamHI and StuI sites of partially digested *pET-tetR- PtetA-Bmei cobWNST*	This study

2.1.4 Primers

The primers used in this study were purchased from Eurofins scientific. A list of all primers can be found in Table 2.3. Restriction sites are referred to in bold. The corresponding restriction enzymes are listed in the last column of the table. The underlined nucleotides of the N20 primers denote the linker, while the highlighted section is the N20 region on the corresponding genome. In red, point mutations are shown.

Table 2.3 List of primers used in this study

Primer name	Sequence	Restriction site
Sal-nanC	CAT**GTCGAC**CGTCCATTGATCATCAAG	**SalI**
BglII-nanC	GAC**AGATCT**TACGCTTATTGATGACGC	**BglII**
Site A – F check	CAGGACAGTTTTACGTAG	
Sal-yoe2	CAT**GTCGAC**ACTCTAATGGCCAGAGTGATG	**SalI**
BglII-yoe	CAT**AGATCT**GGCATACAAACTGGACGTATTG	**BglII**
Yoe-check	CCGGTTAATTGCTGATAC	
Primer Yoe-Avr	CAT**CCTAGG**TTTATCACACTCAGAGC	**AvrII**
Pd cobA fw primer	CAGCCATATGATCGACGACC	
PdCobA-SpeI	CAT**ACTAGT**TTATGCCGGGTTCCTG	**SpeI**
cobA rev	GAAAGATCGTCGAACAGC	
Primer N20–A 1F	CTAGTTCACTATAGGGAGACCACAAGTTTTAGA	
Primer N20–A 1R	TCTAAAACTTGTGGTCTCCCTATAGTGAA	
Primer N20-B 2F	CTAGTGAGCAACGTTCTTCAGCAGGGTTTTAGA	
Primer N20-B 2R	TCTAAAACCCTGCTGAAGAACGTTGCTCA	
Primer N20–cobA 1F:	CTAGTGTTCGACGATCTTTCCGCCGGTTTTAGA	
Primer N20–cobA 1R:	TCTAAAACCGGCGGAAAGATCGTCGAACA	
BmeiN_AmpTAA_ Rev	GATTAAGCATTGGTAACGTAACGCGTGATCAGAAG	
Bmei N3	GAACGGTGGAAGTGTTGC	

BmeiN-mut-Nde	CAT**CATATG**AAGCCTCGGGCGTC	**NdeI**
T7T	GCTAGTTATTGCTCAGCGG	
BmeiCobN_H554A_F	GACCATATGGGCAAG**GC**TGGCAATCTGGAATGG CTG	
BmeiCobN_E593A_F	GAACGATCCGGGCG**C**GGGCACGCAG	
BmeiCobN_E593A_R	CTGCGTGCCCGC**G**CCCGGATCGTTC	
BmeiCobN_H1094A_ F	GGATAATCGCGAA**GC**CGACCTTCTCGACAG	
BmeiCobN_H1094A_ R	CTGTCGAGAAGGTCG**GC**TTCGCGATTATCC	
BmeiCobN_D1098A_ F	GACCTTCTCG**C**CAGTGATGACTATTAC	
BmeiCobN_D1098A_ R	GTAATAGTCATCACTG**G**CGAGAAGGTC	
BmeiCobN_D1100A_ F	CTCGACAGTG**C**TGACTATTACC	
BmeiCobN_D1100A_ R	GGTAATAGTCA**G**CACTGTCGAG	
BmeiCobN_Q1104A_ F	GTGATGACTATT**ATGCAT**TCGAGGGCGGC	**NsiI**
BmeiCobN_Q1104A_ R	GCCGCCCTCGAA**TGCAT**AATAGTCATCAC	**NsiI**
MSB – cobA STOP	AGGATCCAAAACTAGTTAAAAGTCAACTCCTGT CCG	

2.1.5 Media and solutions for bacterial work

Luria-Bertani (LB) broth:	Tryptone	10 g
	Yeast extract	5 g
	NaCl	5 g

The media was made up to 1 L with distilled H_2O and autoclaved.

Luria-Bertani (LB) broth + agar:

15 g of bacteriological agar was added to 1 L of LB broth before autoclaving.

Super Optimal broth and Catabolite repression (SOC):

	Tryptone	2 g
	Yeast Extract	0.5 g
	1M NaCl	1 mL
	2M Mg^{2+} stock (see below)	1 mL
	20 % (w/v) glucose	1 mL
	1M KCl	0.25 mL

The media was made up to 1 L with distilled H_2O and autoclaved.

2YTNN media:	Tryptone	16 g
	Yeast extract	10 g
	NaCl	5 g
	Na_2HPO_4	10.99 g
	NaH_2PO_4	2.71 g
	NH_4Cl	1 g

This was made up to 1 L with distilled H_2O and autoclaved.

10X M9 salts:	Na_2HPO_4	60 g
	KH_2PO_4	30 g
	NH_4Cl	10 g
	NaCl	5 g

This was made up to 1 L with distilled H_2O and autoclaved.

YE-M9-G media:	10X M9 salts	100 mL
	50% glycerol	10 mL
	0.1 M CaCl	1 mL

	1 M $MgSO_4$	2 mL
	25.03 mg/mL CoCl	2 mL
	1 mM $MnCl_2$ $6H_2O$	50 µL

The above additives were added to YE solution which was prepared separately by adding 10 g yeast extract to 885 dH_2O and autoclaved separately.

YE-M9-GG fermentation media:

This media was prepared as described above to which 0.2% glucose was added.

Minimal media:	10x M9 salts	100 mL
	20% glucose	10 mL
	1 M $MgSO_4$	2 mL
	0.1 M $CaCl_2$	1 mL
	dH_2O	877 mL

All solutions were made up separately and autoclaved before being mixed. For minimal media agar plates to re-streak AR3612 on, 10 mL 5 mg/mL L-cysteine was added as well as 10 mL 5 mg/mL Methionine and 15g bacterial agar before autoclaving. In preparing plates for AR2680, L-cysteine is not added. For the bioassay plates, the same preparation is upheld, with the exception of methionine, which is left out.

20% glucose solution:

20 g of D-glucose was made up to 200 mL with dH_2O and then autoclaved.

1M $MgSO_4$ solution:

24.07 g of $MgSO_4$ was made up to 200 mL with dH_2O and then autoclaved.

0.1M $CaCl_2$ solution:

2.22 g of $CaCl_2$ was made up to 200 mL with dH_2O and then autoclaved.

0.9% NaCl solution:

0.18 g NaCl was made up to 20 mL and filter sterilised using a 0.2 µm syringe filter.

1M IPTG solution:

4.76 g of IPTG were dissolved in 10 mL dH2O, filter sterilised using a 0.2 µm syringe filter, aliquoted into 400 µL and stored at -20°C.

$CaCl_2$/glycerol solution for preparing competent cells:

	$CaCl_2$	0.1 M
	Glycerol	10% (w/v)

This was made up to 100 mL with dH_2O, autoclaved and stored in the fridge.

TB solution (Inoue et al., 1990):

Pipes	10 mM
$CaCl_2 \cdot H_2O$	12 mM
KCl	20 mM
$MnCl_2 \cdot H_2O$	55 mM

All components were made up to solution with dH_2O except $MnCl_2 \cdot H_2O$. The pH was adjusted to 6.8 with 2M KOH, after which the $MnCl_2 \cdot H_2O$ was added. The solution was filter sterilised using a 0.2 µm syringe filter and stored at 4°C.

L-Arabinose:

3 g of L-arabinose was made up to 20 mL with dH_2O and filter sterilised using a 0.2 µm syringe filter. The solution was stored at -20°C.

Antibiotics:

Table 2.4 Antibiotics used in this study.

Antibiotic	Stock concentration	Final concentration
Ampicillin	100 mg/mL in 50% dH_2O, 50% Ethanol	100 µg/mL
Chloramphenicol	34 mg/mL in Ethanol	34 µg/mL
Kanamycin	25 mg/mL in dH_2O	100 µg/mL
Spectinomycin	50 mg/mL in dH_2O	50 µg/mL

All antibiotics were filter sterilised using a 0.2 µm syringe filter, aliquoted into 400 µL and stored at -20°C.

2.1.6 Media and solutions for DNA work

TE buffer:	Tris-HCl, pH 8.0	10 mM
	EDTA, pH 8.0	1 mM
5x DNA loading buffer:	Bromophenol blue (w/v)	0.25 %
	Glycerol (v/v)	50 %
	TE buffer (v/v)	50 %
2x Ligase buffer:	Tris-HCl pH7.8	60 mM
	MgCl2	20 mM
	DTT	20 mM
	ATP	2 mM
	PEG 8000 (v/v)	10 %

Bioline 1kb molecular size hyperladder:

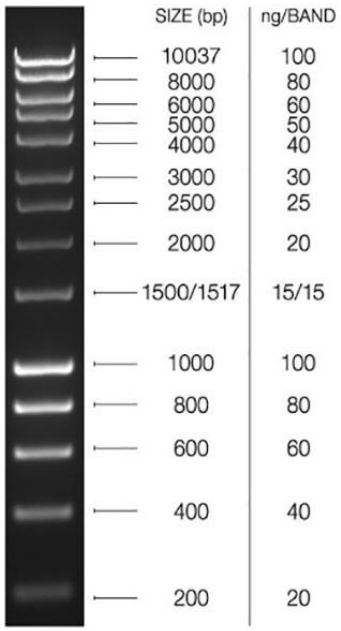


Figure 2.1 5 µL Hyperladder I applied to 1% agarose gel

2.1.7 Media and solutions for protein work

- **Solutions for His-bind column**

Charge buffer:	NiSO4	100 mM
Binding buffer A:	Tris-HCl, pH 7.5	20 mM
	NaCl	100 mM
Wash buffer A.I:	Tris-HCl, pH 7.5	20 mM
	NaCl,	500 mM
	Imidazole	50 mM
Wash buffer A.II:	Tris-HCl, pH 7.5	20 mM
	NaCl	500 mM
	Imidazole	100 mM
Elution buffer A:	Tris-HCl, pH 7.5	20 mM
	NaCl	500 mM
	Imidazole	400 mM
Binding buffer B:	HEPES, pH 7.5	20 mM
	NaCl	500 mM
	Imidazole	10 mM
Wash buffer B:	HEPES, pH 7.5	20 mM
	NaCl,	500 mM
	Imidazole	60 mM
Elution buffer B:	HEPES, pH 7.5	20 mM
	NaCl	500 mM
	Imidazole	400 mM
Strip buffer:	Tris-HCl, pH 8.0	20 mM
	EDTA	100 mM
	NaCl	0.5 M

All buffers were made up by weighing out the individual components, making them up to ~100 mL less than the final volume with MilliQ water before adjusting the pH. After this,

the buffers were made up to the final volume and filtered using a 0.20 µM nylon Millipore filter.

- **Solutions for desalting steps**

Dialysis buffer	Tris- HCL, pH 7.5	20 mM
	NaCl	100 mM
PD10 desalting buffer	HEPES, pH 7.5	20 mM
	NaCl	100 mM

- **Solutions for Sodium Dodecyl Sulfate polyacrylamide gel electrophoresis (SDS-PAGE):**

10x Running buffer:	Tri-HCl	30 g/L
	Glycine	144 g/L
1x Running buffer:	10x running buffer	100 mL
	10% (w/v) SDS	10 mL

Made up to 1 L with dH_2O.

2x Laemmli sample buffer:	0.5M Tris-HCl, pH 6.8	2.5 mL
	Glycerol	2 mL
	SDS 10% (w/v)	4 mL
	ß-mercaptoethanol	1 mL
	Bromophenol Blue 0.08% (w/v)	0.5 mL

Made up to 20 mL with dH_2O.

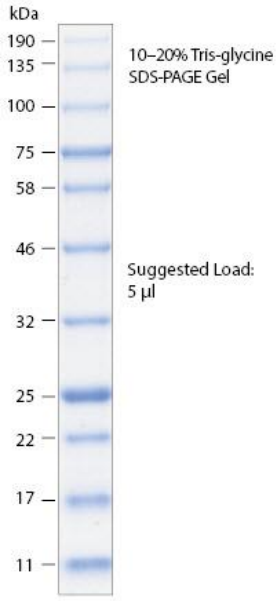
Coomassie blue stain:	Trichloroacetic acid 100%	250 mL
	Coomassie blue R250	0.6 g
	SDS 10% (w/v)	0.1 g
	Tris-HCl	0.25 g
	Glycine	0.15 g

Made up to 500 mL with dH_2O.

Table 2.5: SDS PAGE gel compositions

SDS Gel				
Running gels	**10%**	**12.5%**	**Stacking gel**	**5%**
dH_2O	4.7 mL	3.4 mL	dH_2O	3.4 mL
30% Acrylamide	5 mL	6.3 mL	30% Acrylamide	1.5 mL
1.5 M Tris-HCl, pH 8.8	3.8 mL	3.8 mL	0.5 M Tris-HCl, pH 6.8	1.9 mL
10% SDS	1.5 mL	1.5 mL	10% SDS	750 µL
10% APS	150 µL	150 µL	10% APS	75 µL
TEMED	10 µL	10 µL	TEMED	10 µL

NEB Broad Range prestained Protein Standard (P7706):



- **Solutions for gel filtration column (size exclusion)**

Sample buffer:	HEPES, pH 7.5	20 mM
	NaCl	100 mM

Table 2.6. Contents of molecular weight standards solution for gel filtration.

Protein	Approximate molecular weight (Da)	Concentration
Albumin, bovine serum	66,000	10 mg mL-1
Alcohol dehydrogenase, yeast	15,0000	5 mg mL-1
β-amylase, sweet potato	200,000	4 mg mL-1
Apoferritin, horse spleen	443,000	10 mg mL-1
Thyroglobulin, bovine	669,000	8 mg mL-1
Blue dextran	2,000,000	2 mg mL-1

2.1.8 Media and solutions for HBAD work

- **Solutions for Tris-acetate EDTA gel**

Table 2.7. TAE gel composition

TAE gel, 30% acrylamide/bisacrylamide	
50X TAE buffer	200 µL
dH2O	3.1 mL
30% acrylamide/bisacrylamide	6.6 mL
APS	100 µL
TEMED	5 µL

Enzyme assays were measured using a FLUOstar omega microplate reader by BMG. labtech.

2.2 Microbiological methods

2.2.1 Sterilisation of reagents

Unless stated otherwise, all media and buffers were sterilised at 121°C and 1 bar pressure for 15 minutes in an autoclave. Temperature sensitive reagents were filter sterilised (0.2 μm diameter pore size).

2.2.2 Liquid cultures

Liquid cultures were prepared by inoculating 5 mL of LB with a single colony from an agar plate. Cultures were then incubated overnight at ~150 rpm and 37°C, unless stated otherwise.

2.2.3 Storage of bacteria

For long-term storage of bacteria, glycerol stocks were prepared. Sterile glycerol was added to overnight liquid cultures, to a final concentration of 15% (v/v) before storage at -80°C.

2.2.4 Plate cultures

Plate cultures were prepared either by streaking out a glycerol stock or via transformation of competent cells with a plasmid of interest (Section 2.2.6). Plates were then incubated overnight at 37°C, unless stated otherwise.

2.2.5 Preparation of competent *E. coli* cells

A 100 mL of LB was inoculated 1:100 with liquid starter culture and was incubated at 37°C at 180 rpm until an OD_{600} value of 0.4 was reached. The cells were then chilled on ice for 30 minutes and centrifuged at 2,700 x g for 10 minutes at 4°C. The pellets were gently resuspended in 20 mL ice cold $CaCl_2$/glycerol solution and incubated on ice for 15 minutes. Cells were collected again by centrifugation and re-suspended in 2 mL of ice cold $CaCl_2$/glycerol solution. 50 μL aliquots were flash frozen in dry ice before being stored at -80°C.

2.2.6 Transformation of *E. coli* competent cells

For plasmid amplification, 10 μl competent cells was defrosted on ice for 10 min before adding 0.5 μl plasmid DNA to 10 μl cells. For transformation of ligation mixture (Section 2.3.6), 5 μl plasmid was mixed with 50 μl competent cells. The mixture was then incubated on ice for 15 minutes and heat-shocked at 42°C for 1 minute before rapid transfer to ice and incubation for 2 minutes. After the addition of 200 μl SOC media, the cells were incubated at 37°C for 30-60 min to allow antibiotic resistance expression. The mixture was

then spread on a LB agar plate containing the required antibiotics and incubated overnight at 37°C, unless stated otherwise.

2.2.7 Recombinant protein overproduction in *E. coli:*

BL21 Star (DE3) or BL21 Star (DE3) pLysS *E. coli* competent cells were transformed with a vector containing the gene(s) of interest. A 5 or 10 mL LB starter culture containing the appropriate antibiotics was inoculated with a single colony and incubated overnight at 37°C. 1 L of LB containing ampicillin (and chloramphenicol for pLysS strain) was inoculated from the starter culture, grown at 37°C and shaken at ~160 rpm until the culture reached an OD_{600} ~ 0.6. Protein production was induced by adding 400 µM IPTG, after which expression was left at 19°C overnight. For IMAC, the cells were harvested by centrifugation at 4000 rpm for 20 min at 4°C. The pellet was re-suspended in 25 mL of binding buffer (Section 2.1.7.1) and either directly used for protein purification (Section 2.4.1) or stored at -80°C.

2.2.8 Preparation of ultracompetent *E. coli* cells for CRISPR-Cas9

ED656-3B was transformed with plasmid pCas9 as described in Section 2.2.6, with the exception of the grown temperature. All cells containing pCas9 were grown at 28°C to accommodate the temperature-sensitive origin of replication. From this plate culture, a 5 mL LB starter culture was prepared containing 25 mg/L kanamycin. 10 mL of LB containing kanamycin was inoculated with 100 µL starter culture and grown for 1.45 hours whilst shaken at 150 rpm before inducing expression of Lambda red recombinase genes with 100 µL 1M L-Arabinose. The cells were then grown further until an OD_{600} of 0.4-0.6 was reached. At this point the culture was incubated in ice water slurry for 15 min before the cells were harvested by centrifugation at 3000 rpm for 10 min at 4°C. Cells were then washed with 5 mL ice cold TB and incubated on ice for 25 min. The centrifugation step was repeated as described above and cells were resuspended in 1 mL of ice cold TB. 100 µL of 100% glycerol was added to a final concentration 20% and the cells were aliquotted in 80 µL. These were flash frozen and stored at -80°C.

2.2.9 Lysis of cells by sonication

Harvested cells were lysed by sonication using a Sonics Vibracell Ultrasonic processor equipped with a Jencons solid titanium alloy probe in 30 sec bursts with 30 sec breaks repeated 6 times at 55% maximum amplitude. The sonicated cells were centrifuged at 18,000 rpm for 20 min at 4°C to remove cell debris. The soluble cell extract (supernatant) was then purified by IMAC (Section 2.4.1).

2.2.10 Bacterial production of cobalamin in 250 mL shake flasks

All bacterial growth for cobalamin production took place at 28°C to accommodate for the majority of the pathway's genes originating from *Rhodobacter capsulatus*, the optimal growth temperature of which is 28°C. Plate cultures were made from glycerol stocks of cobalamin-producing strains as described in Section 2.2.4. From these plates, liquid pre-cultures were made as described in Section 2.2.2. Plastic 250 mL Erlenmeyer flasks, containing 25 mL of YE-M9-G media (Section 2.1.7), were inoculated with 2.5 mL pre-culture, grown at 28°C and shaken at ~160 rpm for 6 hours before inducing T7 promoters with 400 µM IPTG. Expression was continued at 28°C under constant shaking at ~160 rpm. Samples were taken for analyses after 24 hours and 48 hours.

2.2.11 Bacterial production of HBAD

E. coli strain ED661 contains the B_{12} pathway up to intermediate hydrogenobyrinic acid a,c-diamide (HBAD), as well as enzymes assisting in further amidations of the central corrin ring. To produce HBAD, liquid cultures were made as described in Section 2.2.2 which were used to inoculate 1L of 2YTNN media 1:100. The cultures were incubated at 28°C, shaking at 160 rpm until an OD_{600} of 0.6 was reached. HBAD production was induced by adding 0.4 mM IPTG. The cultures were placed 18°C overnight at constant shaking ~160 rpm. Cells were harvested via centrifugation at 4000 rpm for 20' at 4°C leaving a distinctly pink pellet (Fig 2.2.1). Supernatant was collected for potential HBAH purification.

Figure 2.2 Pellet and supernatant of HBAD-producing strain ED661

2.2.12 Bacterial production of cobalamin in fermenter.

For fermentation tests, Infors minifors 2 fermenters were used, connected to an Eve system. A pre-culture of vitamin B_{12}-producing strains was made as described in Section 2.2.2, which was grown at 28°C. This was used to inoculate 500 mL YE-M9-GG media (Section 2.1.5) 1:100. The pH was stabilised at 7.0 using 25% sulphuric acid and 25% ammonia base solutions. The O2 level was set to 100%. 0.5 mL of polypropylene glycol was added during inoculation as anti-foam. Those cultures containing T7 promoters were induced by adding 0.4 mM IPTG 6 hours after inoculation.

2.2.13 Preparation of *Salmonella* bioassay plates

B_{12} yields of cultures were determined using *Salmonella typhimurium AR3612* or *AR2680* bioassay plates. The strains were streaked out from glycerol stocks (Section 2.2.4) and subsequently re-streaked on minimal media plates containing L-cysteine and methionine for AR612. For re-streaking AR 2680, the plates only contained methionine. This re-streaking was done so a film of bacterial growth would cover the entire plate. These plates were incubated at 38°C overnight. The cultured layer was scraped off and collected in a 1.5 mL Eppendorf tube. The methionine was removed by adding 0.9% sterile NaCl solution to the cells and spinning this down at 6000 rpm for 2 min. The supernatant was removed and this wash step was repeated 3 times. The bioassay plates were prepared as minimal media plates, to which the *Salmonella* was added when the media had cooled down to 40°C. Plates were stored at 4°C for 7 days max.

2.3 Molecular biology methods

2.3.1 DNA extraction via Miniprep

A QIAprep® Miniprep Kit (Qiagen) was used for the purification of plasmid DNA as described in the handbook using the microcentrifuge protocol.

2.3.2 DNA digest

Plasmid or PCR product DNA was digested using the relevant restriction enzymes at 10 U μL^{-1}, in the optimal buffer from Promega or New England Biolabs. Reactions were incubated at the temperature required by the restriction enzyme for 0.3 – 1 hour. Restricted fragments were separated according to length via gel electrophoresis (Section 2.3.3).

Table 2.8 Typical DNA pipetting scheme

Component	**Double digest**
Restriction enzyme 1	0.5 µL
Restriction enzyme 2	0.5 µL
10x buffer	1 µL
Plasmid/PCR DNA	5 µL
MilliQ H_2O	3 µL

Partial digests were used to extract bands which were otherwise unable to be retrieved from a full digest. A typical pipetting scheme is show below.

Table 2.9: Pipetting scheme partial digest

	Units				
Component	**1.2**	**1**	**0.5**	**0.25**	**0.1**
Plasmid DNA (µL)	10	10	10	10	10
10 X Cutsmart Buffer (µL)	2	2	2	2	2
0.2 U μL^{-1} Enzyme 1 (µL)	6	5	2.5	1.25	0.5
10 U μL^{-1} Enzyme 2 (µL)	1	1	1	1	1
MilliQ (µL)	1	2	4.5	5.75	6.5

2.3.3 PCR reaction

All PCR reactions were performed in an Eppendorf Mastercycler 5341 PCR machine using the FastStart High Fidelity PCR System (Roche). The basic PCR reaction and the cycles used are outlined below:

Table 2.10 PCR reaction mix.

Component	**µL**
MilliQ H_2O	32.5 – 34.5
10x FSHF buffer	5
DMSO	0 – 2
5 mM dNTPs	2
10 µM 5' forward primer	2
10 µM 3' reverse primer	2
DNA template	1
Taq polymerase (5 U μL^{-1})	0.5

2.3.4 Gel electrophoresis

DNA fragments were separated by DNA gel electrophoresis. 1 % agarose gels were prepared in 1x TAE buffer with the addition of ethidium bromide to a final concentration of 0.5 µg mL^{-1}. Ethidium bromide allows visualisation of the DNA as it binds between the base pairs and then upon exposure to UV radiation re-emits at a wavelength that can is visible on a UV transilluminator. DNA samples containing 20 % (v/v) DNA loading buffer (Section 2.1.6) were loaded into the wells of the agarose gel and electrophoresis was carried out at 80 V for approximately 50 min using a SubCell GT electrophoresis tank (BioRad) connected to a Power PAC 300 power supply (BioRad).

2.3.5 Gel extraction

DNA bands of interest were excised from an agarose gel using a scalpel blade. Subsequent gel extraction was carried out using a QIAquick® Gel Extraction Kit (Qiagen) using the microcentrifuge protocol in the handbook.

2.3.6 Ligation of DNA

Vectors and inserts were digested with the relevant restriction enzymes as described in Section 2.3.2 and extracted from the gel (Section 2.3.4). DNA fragments were ligated into vectors at room temperature for 2 hours or at 4 °C overnight. Reaction components are listed in Table 2.3.3.

Table 2.11 Ligation pipetting scheme

Component	Volume (µL)
Insert	2.5
vector	1.5
10x T4 DNA ligase buffer (Promega)	1
milliQ H_2O	4.5
T4 DNA Ligase (3 U μL^{-1}, Promega)	0.5

2.4 Biochemical methods

2.4.1 Protein purification

- **Immobilised metal ion affinity chromatography**

Columns containing approximately 5mL of Chelating Sepharose Fast Flow Resin (GE Healthcare) were washed with dH_2O and subsequently charged with $NiSO_4$ (charge buffer). Following equilibration with binding buffer, the soluble fraction of cell extract was applied to the column. ^{6x}His-tagged protein bind to the Ni^{2+}. Binding buffer followed by wash buffers I and II were applied to the column to remove unbound protein. The ^{6x}His-tagged protein was eluted in 2 mL fractions with buffer containing imidazole. The column was re-generated with strip buffer and washed with dH_2O. Buffer compositions can be found in Section 2.1.7.

- **Desalting column PD10**

Eluted protein was desalted by buffer exchange into imidazole-free buffer using pre-packed, disposable PD-10 columns (GE Healthcare). The column, containing Sephadex G25 resin, was equilibrated with 25 mL of imidazole-free buffer followed by application of 2.5 mL of eluted protein. The protein was then eluted with 3.5 mL of imidazole-free buffer.

2.4.2 Bradford protein assay

Rough Bradford assays were done to determine which fractions held protein (and would subsequently loaded onto an SDS PAGE gel). This was done by mixing 15 µL of MilliQ with 5 µL bradford reagent and starting the reaction with 2 µL of fraction sample. Bradford reagent contains coomassie brilliant blue G-250 which is undergoes a colour change from red/brown to blue upon binding to protein's carboxyl groups.

2.4.3 A280 protein concentration estimation

Protein concentrations were determined by measuring absorbance at 280 nm in a cuvette with a path length of 1 cm. Blanks were set using 1 mL of buffer. The molar extinction coefficient (ε) was calculated using the EXPASY protparam bioinformatics tool (Gasteiger et al., 2005). Protein concentrations were calculated using the Beer Lambert law;

$$A = \varepsilon * c * l$$

where A is absorbance, c is the concentration (M) and l is the path length in cm.

2.4.4 SDS PAGE

Proteins were separated according to their mass using SDS-PAGE. Gels were prepared with a 15% running gel and a 5% stacking gel. Samples were denatured by the 1:1 addition of Laemmli sample buffer (Section 2.1.7) and boiling for 10 minutes. 5 - 10 µL of denatured sample was loaded into each well and 7 µL of molecular mass marker (NEB P7706) run on each gel to estimate the relative molecular mass of the proteins of interest. Electrophoresis was run at a constant voltage of 200 V using an Atto Dual Mini-Slab system and an Atto Mini Power electrophoresis power supply. Gels were stained using Coomassie blue stain (Section 2.1.7) and de-stained with dH_2O.

2.4.5 HBAD purification

All HBAD purification steps were done in the dark as it is susceptible to light-induced degradation.

- **BtuG-IMAC column**

BtuG is a B_{12}-binding protein which has shown to be promiscuous to cobalamin derivatives. This property was exploited in this method by executing the binding and wash steps of BtuG-containing cell lysate on Ni^{2+}-Speharose columns and subsequently loading the soluble fractions of lysed ED661 cultures (Section 2.2.11). Binding buffer and Wash buffers I and II were applied to wash away unbound protein and impurities before eluting the BtuG-bound HBAD off the column with imidazole.

- **Buffer exchange via dialysis**

The BtuG-HBAD was dialysed against 3 L of Dialysis buffer (Section 2.1.7.2) under constant stirring for 1 hour. The dialysis bag was then transferred to another 3 L of Dialysis buffer which was left stirring overnight at 4°C.

- **Release of HBAD by heat treatment**

To release HBAD from BtuG, the solution was boiled for 15 min (Figure 2.3). The precipitated BtuG was then spun down at 4000 rpm for 20 min at °C and the soluble fractions was collected for further purification.

Figure 2.3 HBAD-BtuG precipitate. *Heat treatment of the mixture by boiling for 15 minutes caused the protein to release HBAD.*

- **Separation of multi-amidated B_{12}-intermediates**

As mentioned in Section 2.2.11, strain ED661 has the ability to produce all of HBA to HBAH, and those in between. To separate the various amidated forms and purify the desired HBAD, the sample was loaded onto a Diethylaminoethyl sepharose (DEAE) (Sigma Aldrich) column. The anion exchange column bound the negatively charged HBA-H intermediates at pH 7.5 which were eluted separately by increasing the salt concentration.

- **Concentration of fractions by reverse phase chromatography**

The eluted fractions of Section 2.4.5.4 were adjusted to pH 4.5 with glacial acetic acid before being applied to 25-40 uM LiChroprep reverse-phase C18 columns (Merck). After all sample was applied, the HBA-intermediates were eluted with 2 mL with 100% MeOH which is subsequently evaporated Spin-vac (MiVac DUO concentration, Genevac).

- **TAE gel**

Tris-Acetate gels were used for the purpose of separating B_{12} intermediates based on their number of amidated groups as the EDTA in the buffer sequested divalend cations. The gels were prepared with 20% acrylamide/bisacrylamide. The samples loaded onto the gel were made up out of 20 µL sample and 5 µL 50% glycerol. Electrophoresis was run at a constant voltage of 200 V using an Atto Dual Mini-Slab system and an Atto Mini Power electrophoresis power supply. Samples were visualised by placing the gel on a UV transilluminator.

- **HPLC-MS analysis of HBAD sample**

Samples were separated on an Agilent 1100 series HPLC coupled to a micrOTOF-Q II (Bruker) mass spectrophotometer using an Ace 5 AQ column (2.1 x 150 mm; Advanced Chromatography Technologies) maintained at 30˚C and with a flow rate of 0.2 mL/min. The mobile phase consisted of 0.1% TFA in water (solvent A) and acetonitrile (solvent B). The initial conditions were set at 100% of solvent A. The concentration of solvent B was increased with a linear gradient to 100% over 50 min and back to the initial conditions at 60 min. The mass spectrometer was operated in the positive electrospray mode.

2.4.6 Enzymatic cobalt chelation of HBAD

All Enzymatic reactions were followed in a BMG labtech Omega plate reader.

Mass spectra were collected using a Finigan MAT 95s LCQ classic electrospray mass spectrometer. Additional spectra were collected using a Bruker micro QtoF II with electrospray ionisation.

Chapter 3

Changing promoters in a vitamin B_{12}-producing strain of *E. coli* to enhance B_{12} production

3.1 Introduction

B_{12} production has historically been dependent on bacterial strains that had been selected due to naturally high levels of the nutrient, strains that have contained either the native aerobic or anaerobic pathway. Although strains such as *Pseudomonas denitrificans* and *Propionibacterium shermanii* are able to produce relatively high vitamin B_{12} levels they are tempered by complex growth conditions (Martens et al., 2002). Production improvements to commercial strains have mostly relied on rounds of mutagenesis and genetic engineering (Blanche et al., 1998; Pel & Hopper, 2010). Additional production improvements have been studied through dissolved oxygen control (Li et al., 2012), pH control (Li, Liu, Chu, et al., 2008) and continuous betaine feeding strategies (Li, Liu, Li, et al., 2008). Whilst these approaches have been relatively successful, the maximum production levels for these organisms appears to have been reached and, therefore, more recently attention has steered towards production via alternative organisms. One such study described the cloning of genes required for B_{12} production into *Bacillus megaterium* whereby a yield of 200 µg/L was achieved by growing on cheap media (Moore et al., 2014). However, in the era of synthetic biology the benefits of using *E. coli* cannot be overlooked. As a well-studied model organism, *E. coli* is a very convenient choice for metabolic engineering (Rosano & Ceccarelli, 2014). As such, it has been used as a host for the heterologous production of multiple products. These range from terpenoids (Martin et al., 2003; C. Zhang & Hong, 2020) to non-natural alcohols (K. Zhang et al., 2008) and biofuels (Fatma et al., 2018).

Previous work in the Warren group has led to an *E. coli* strain that has the capacity to produce vitamin B_{12} to around 40 mg/L of culture. This strain was named ED656-3B and contains the genes for the aerobic vitamin B_{12} pathway housed within two artificial operons that had been integrated into the host's genome. The two engineered operons within the *E. coli* genome are referred to as the "A" and "B" sites; the 'A-site' contains all the genes needed to produce adenosylcobalamin, whereas the 'B-site' contains an extra set of genes encoding the cobaltochelatase enzyme (Fig 3.1.1).

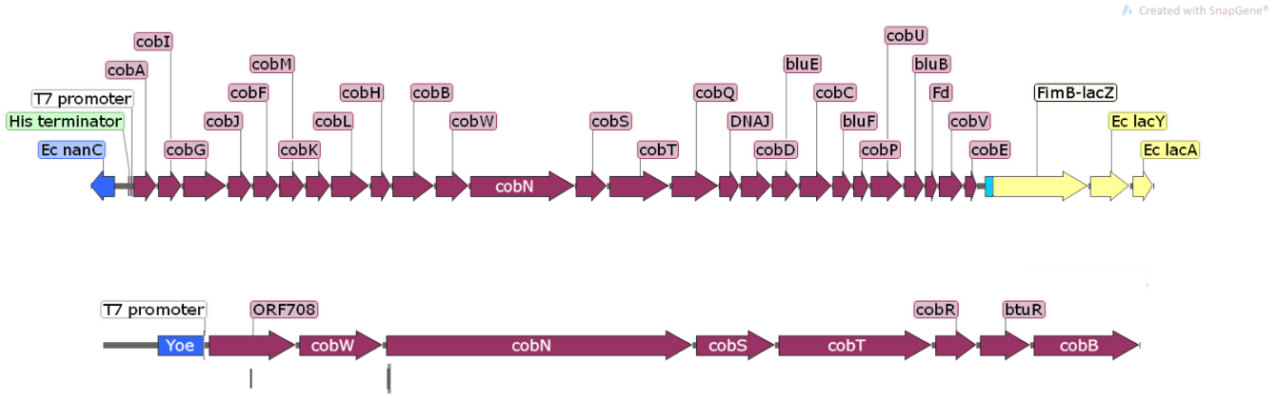


Figure 3.1 Cobalamin biosynthetic pathway operons in ED656-3B. Top: 'A-site' operon, containing all genes required to produce cobalamin. Bottom: 'B-site' operon, containing an extra set of cobaltochelatase genes.

The two operons are controlled by the T7 promoter, a powerful promoter that initiates high-level expression (Tabor, 1990), which, once induced, limits the lifespan of the cell (Bentley et al., 1990). Whilst this promoter system clearly induces a good level of B_{12} production, further B_{12} production could be achieved by moving towards a weaker constitutive promoter system. Such a constitute promoter would allow for continuous B_{12} production over a longer period of time. The benefit of a promoter library to find the optimal growth and product formation phenotypes has been previously investigated (Alper et al., 2005). In this study, the multifaceted characterisation of promoter strength was coupled with an investigation into the expression levels of specific genes and the effect on the phenotype in *E. coli*. An optimum is achieved in terms of growth and yield, after which an increase of promoter strength decreases the desired phenotype. This effect was shown to be remedied by increasing expression levels of downstream genes as a way to minimise toxic intermediate build-up. As the A-site operon of ED656-3B contains > 20 genes, upregulating each under a different promoter is not feasible. However, an attempt at phenotype optimisation can be made by placing the entire operon under the control of promoters of various strengths. Additionally, different combinations of promoter strengths in the A- and B-site operon might also have an effect on the amount of cobalamin production.

3.1.1 Project aim

The work presented in this chapter aims to place the cobalamin pathway in strain ED656-3B under constitutive promoters of varying strengths to find a phenotypical optimum represented in growth, cell lifespan and cobalamin yield.

3.2 Choosing constitutive promoters

To enable a range of expression levels of the two main recombinant cobalamin biosynthetic operons in *E. coli*, promoters were sought representing potentially strong, medium and weak initiators of transcription. The research reported by Alper et al. describes a promoter library that was used to assess gene expression levels and their effect on growth and product formation. Several promoters from this library were available in the lab. From these, promoters A2 and A3 are considered strong and medium promoters based on the relative mRNA when compared to their parent promoter, LTetO1. Their expression levels were also measured via GFP fluorescence (Alper et al., 2005). A third promoter, M2 (also available in the lab), was selected to represent a weak promoter, rounding out the selection of promoter strength representation. The 'weak' M2 promoter was selected from a promoters catalogue that had been reported as a part of an igem project and had similarly been quantified through the production of a fluorescent protein (*Promoters/Catalog/Anderson - Parts.Igem.Org*, n.d.).

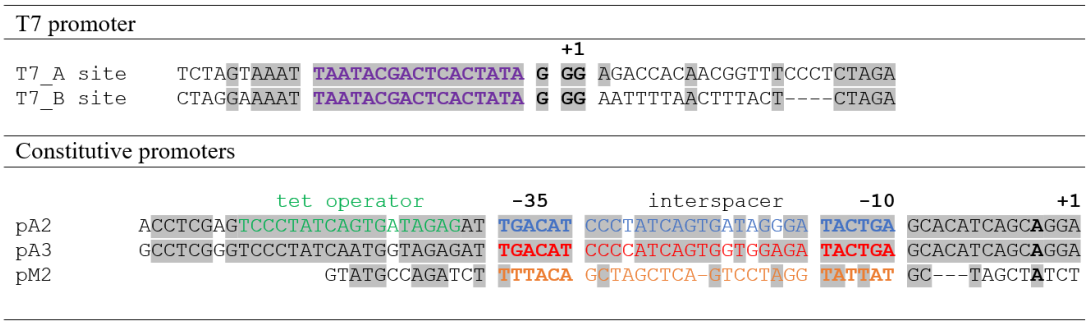


Figure 3.2 Promoter sequences of T7 (purple), A2 (blue), A3 (red) and M2 (orange). Conserved bases are highlighted in grey.

3.3 CRISPR-Cas9 cloning

The method of CRISPR-Cas9 cloning that is utilised here relies on two plasmids to provide the component parts of the complete editing mechanism to the strain. This mechanism utilises single guide DNA (sgRNA), which recognises a targeted sequence on the genome and guides the Cas9 endonuclease to the site for double-stranded cleavage. Donor DNA is integrated into the genome via homologous recombination. Figure 3.3 shows a schematic representation of this system.

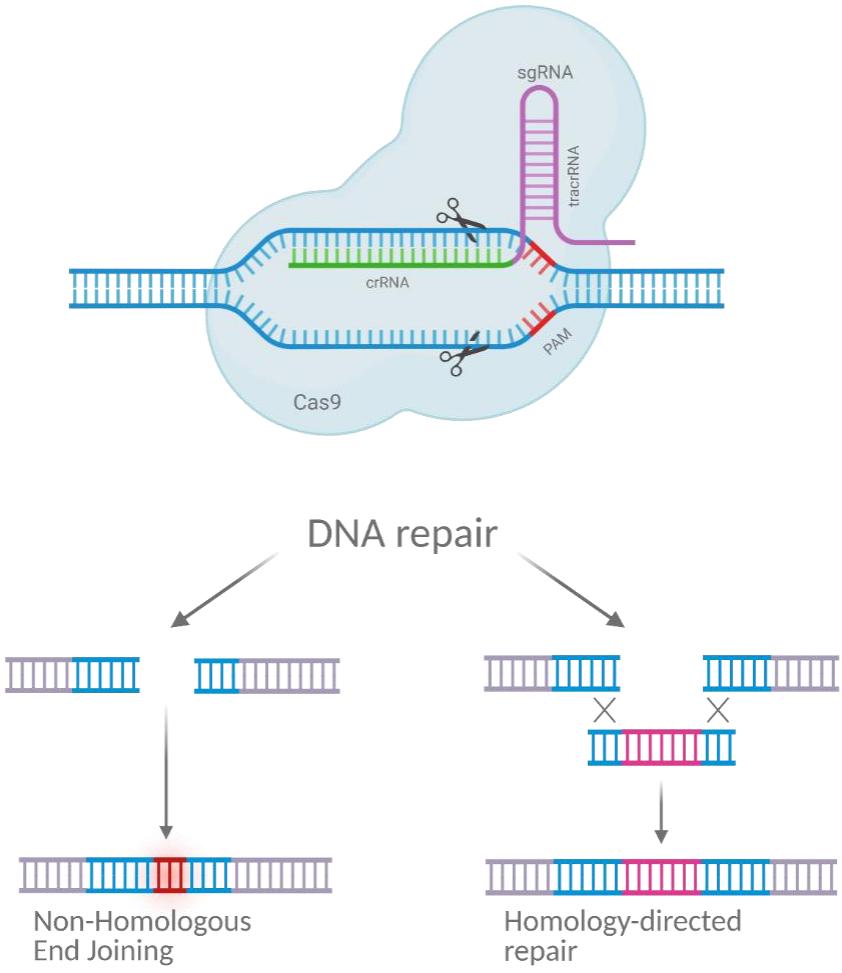


Figure 3.3 Schematic representation of CRISPR-Cas9-mediated DNA cleavage and subsequent repair. *The sgRNA is made up of the genome-specific CRISPR RNA (crRNA) (green) and transactivating CRISPR RNA (tracrRNA) (purple) which enables the Cas9 enzyme (blue form) to recognise and cleave the target DNA next to the protospacer adjacent motif (PAM) (red). DNA repair follows via non-homologous end joining or homology-directed repair.*

Genome editing in *E. coli* is initiated by transforming the desired strain with pCas9 (Jiang et al., 2015) (Figure 3.4), a plasmid that enables the cell to express the Cas9 endonuclease under its native promoter. The plasmid also allows the expression of the lambda red recombinase genes upon induction with L-Arabinose. The overexpression of the recombinase genes, prior to the introduction of the target- and donor DNA in the 'CRISPR transformation', is vital to editing success. Transformed cells are selected for pCas9 via kanamycin resistance and are grown at 30 °C to accommodate the temperature-sensitive origin of replication, which is subsequently employed to help remove the plasmid after successful selection of a desired mutant. To complete the removal of plasmids, pCas contains an sgRNA sequence which, when induced by IPTG, targets the replicon of pTarget-Afe.

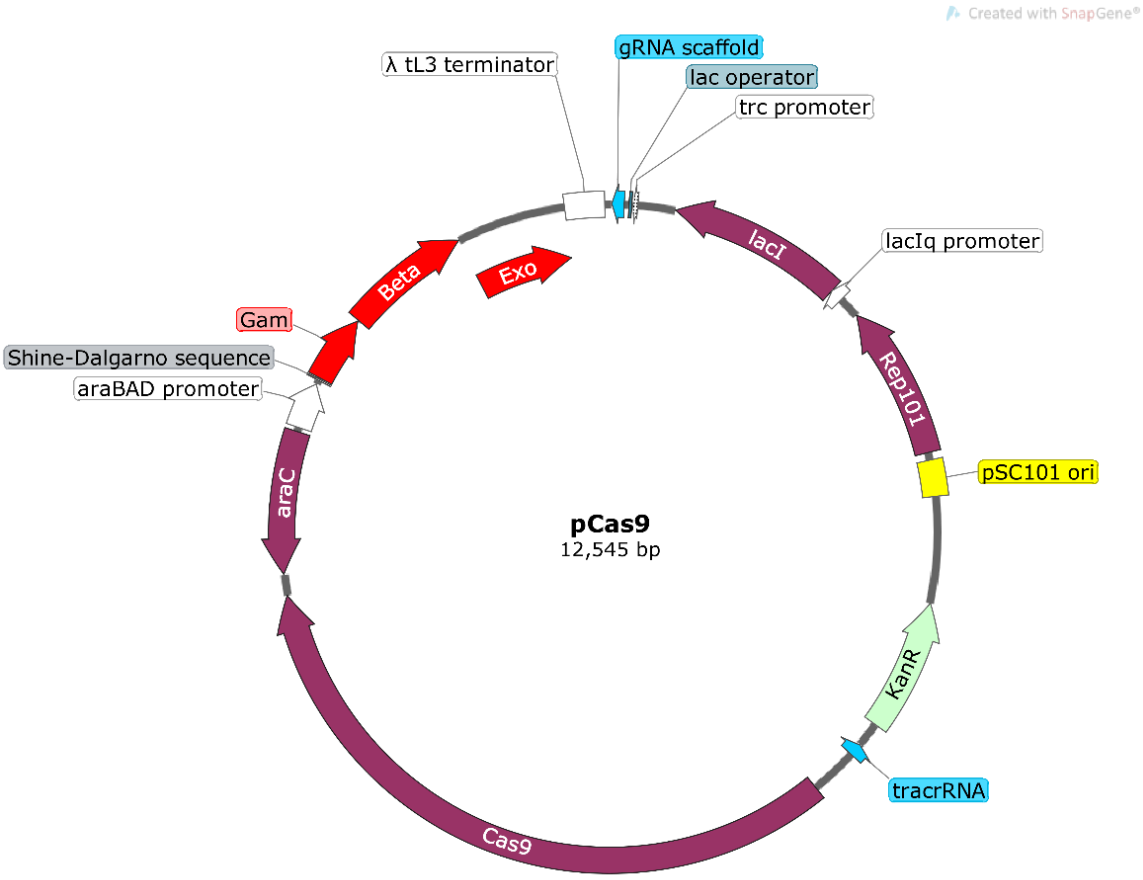


Figure 3.4 pCas9 plasmid. *Containing the Cas9 gene from S. pyogenes with its native promote and Lambda red recombinase genes (exo, bet, gam) under the control of the araB promoter inducible by L-arabinose. Origin of replication: repA101, temperature sensitive for self-curing. Kanamycin resistant. sgRNA containing an N20 sequence to target pMB1 replicon under the trc promoter inducible by IPTG.*

For the work described in this chapter, genome editing was performed on ED656-3B, the *E. coli* strain containing genes of the cobalamin pathway on two biosynthetic operons. When ED656-3B was transformed with pCas9 and induced with arabinose, it was provided with the target DNA as well as the donor DNA for homologous repair. The target DNA was introduced via pTarget-N20 (Figure 3.5). This plasmid contains the sgRNA which directs the Cas9 enzyme to the relevant promoter region in one of the two recombinant cobalamin-biosynthetic operons. Within the plasmid the constitutive promoter J23119 initiates transcription of the sgRNA. The plasmid also contains the *SmR* gene, which encodes an aminoglycoside adenylyltransferase that provides resistance to both spectinomycin and streptomycin (Murphy, 1985). Finally, an N20 sequence on the plasmid is the target for sgRNA transcribed on pCas. After a successful CRISPR clone is identified, this strain can be rid of pTarget by inducing the transcription of this sgRNA on pCas with IPTG.

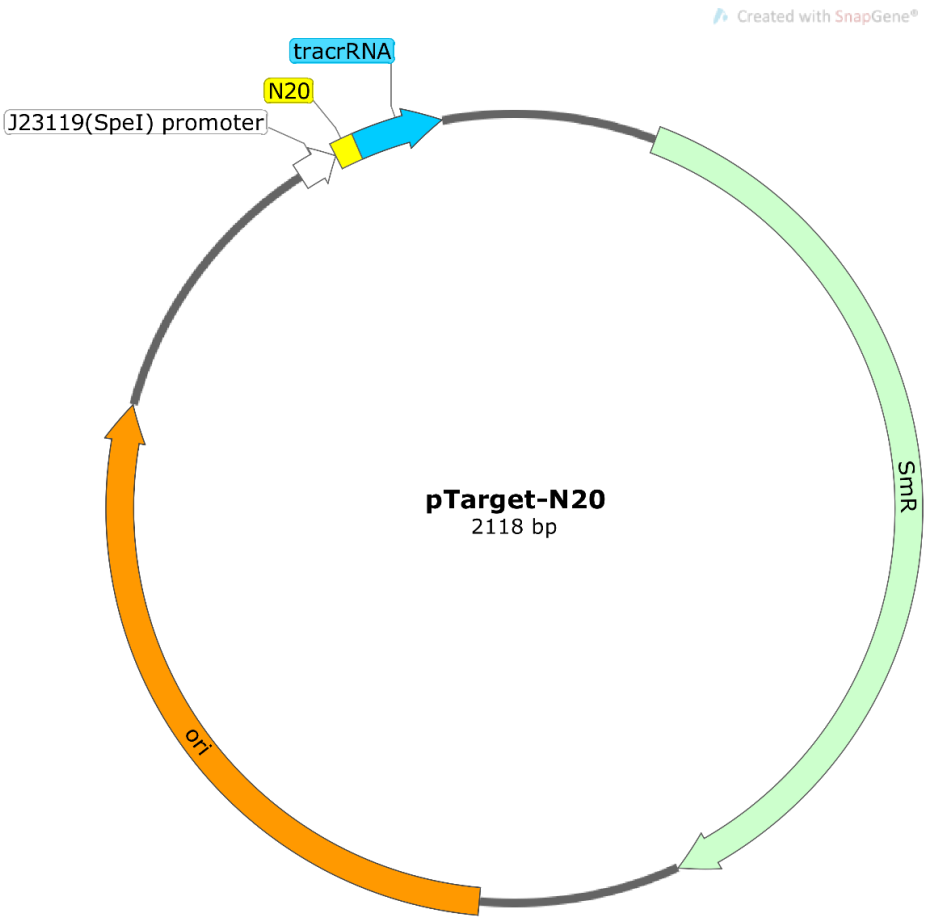


Figure 3.5 pTarget-N20. Containing the sgRNA to guide the Cas9 endonuclease toward the target DNA in the genome. *Spectinomycin and streptomycin resistant via SmR gene. Origin of replication: pMBI. Also contains an N20 which is targeted by pCas9 when induced with IPTG, to remove the plasmid.*

The second DNA requirement for the CRISPR-transformation is the donor DNA. This is restricted from plasmids and inserted as linearized DNA, flanked by 5' and 3' regions to enable homologous recombination with the genomic DNA. With these components the cell has all the tools to perform the CRISPR-Cas9 cut and subsequent desired genomic change (Figure 3.3). A full schematic representation of this process is outlined in Figure 3.6. After the CRISPR process colonies are selected for with kanamycin and spectinomycin resistance, and tested via PCR and subsequent sequencing to ensure the correct modification has been introduced. IPTG induction of the sgRNA on pCas ensured removal of the pTarget-N20 plasmid which was confirmed by spectinomycin sensitivity assays. The pCas plasmid is removed after heat-induced destruction of the temperature sensitive ori by growing the strains at 42°C for 6 h. The removal of pCas was confirmed via kanamycin sensitivity.

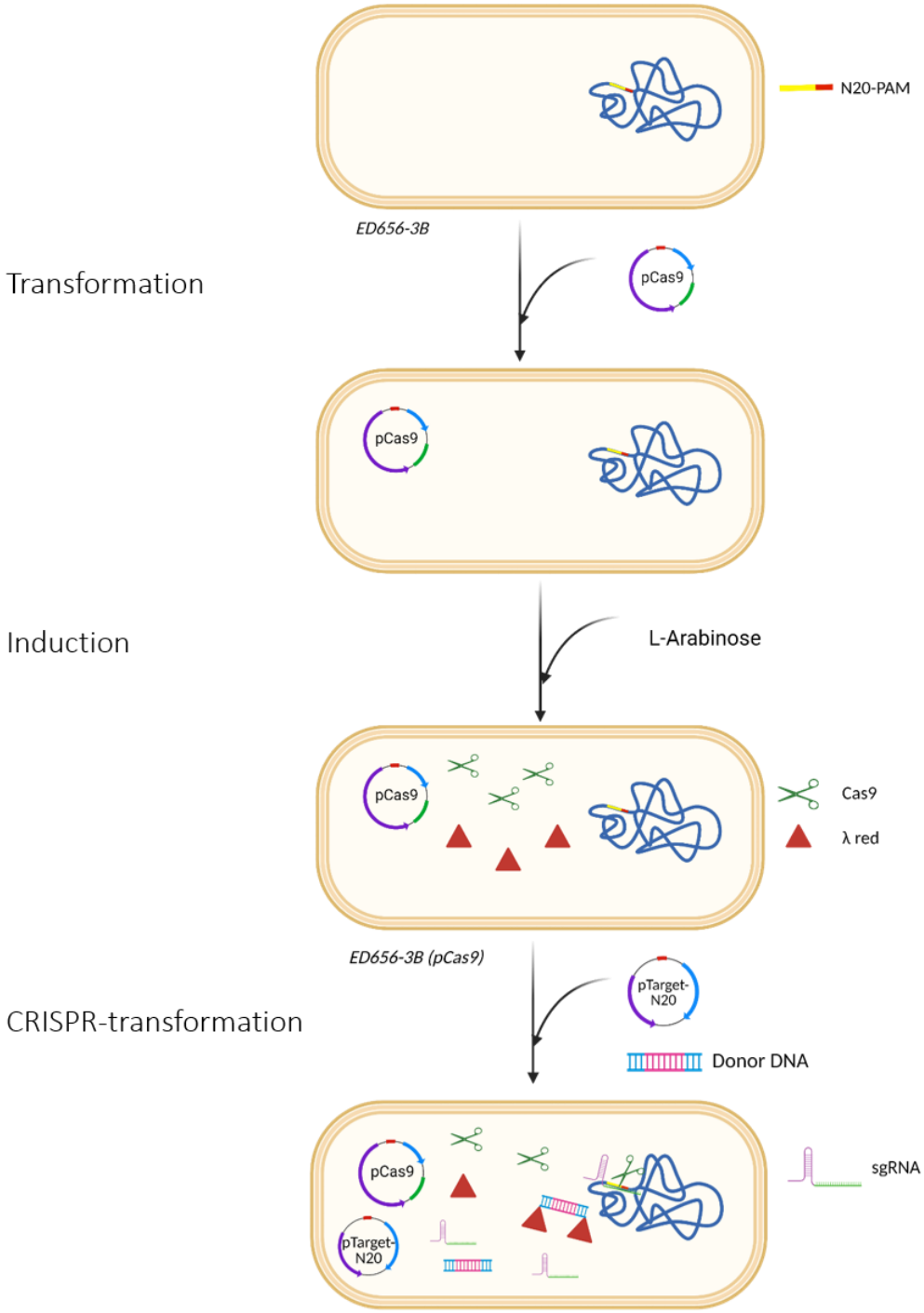


Figure 3.6 Process of introducing CRISPR into ED656-3B. Cell is transformed with pCas9. *Overexpression of Lambda red recombinase genes is induced by adding L-Arabinose. CRISPR transformation follows with pTarget-N20, introducing the sgRNA so Cas9 is directed to the N20 target sequence on the genomic DNA, as well as donor DNA which enables the lambda red system to perform homologous recombination.*

3.3.1 Cloning the target DNA

As described above, pTarget-N20 contains the full sgRNA and is generated by cloning the N20 sequence (the crRNA of the sgRNA) into pTarget-Afe which only contains the tracrRNA of the sgRNA. To target the new DNA components to the correct region of the ED656-3B genome, an N20 was selected inside - or close to - the T7 promoter. This was done using software from chopchop.cbu.uib.no/#. For the A-site, an N20 was selected inside of the promoter sequence. The selected B-site N20 was located 24 bp upstream from the T7 promoter (Figure 3.7).

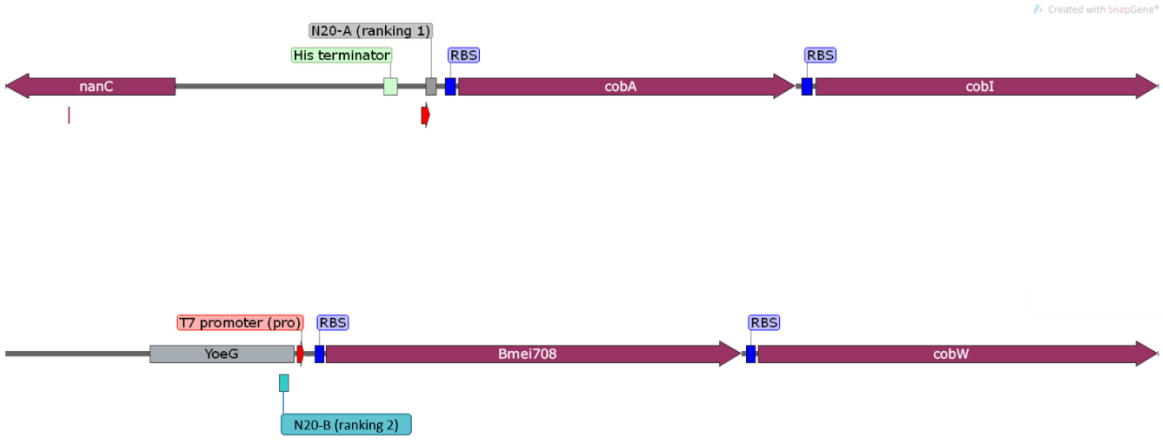


Figure 3.7: N20 target sequences in promoter region A-site (top) and B-site (bottom) in the genomic DNA of ED656-3B.

To synthesize the target inserts, forward and reverse primers were designed that contained the N20 sequences, flanked by linkers that matched the restricted ends of the SpeI and AfeI sites. These primers were self-annealed using a PCR block of 95°C for 2 minutes, followed by a slow ramp down to 4°C. This formed the N20 segment which was subsequently cloned into the pTarget-Afe vector using the restriction enzymes SpeI and AfeI (Figure 3.8). The resulting pTarget-N20-A contained the full sgRNA to lead Cas9 to cleave the T7 promoter in the A-site.

A-site
N20-PAM: ATTAATACGACTCACTATAGGGAGACCACAACGGTTTC T7 promoter in bold
N20

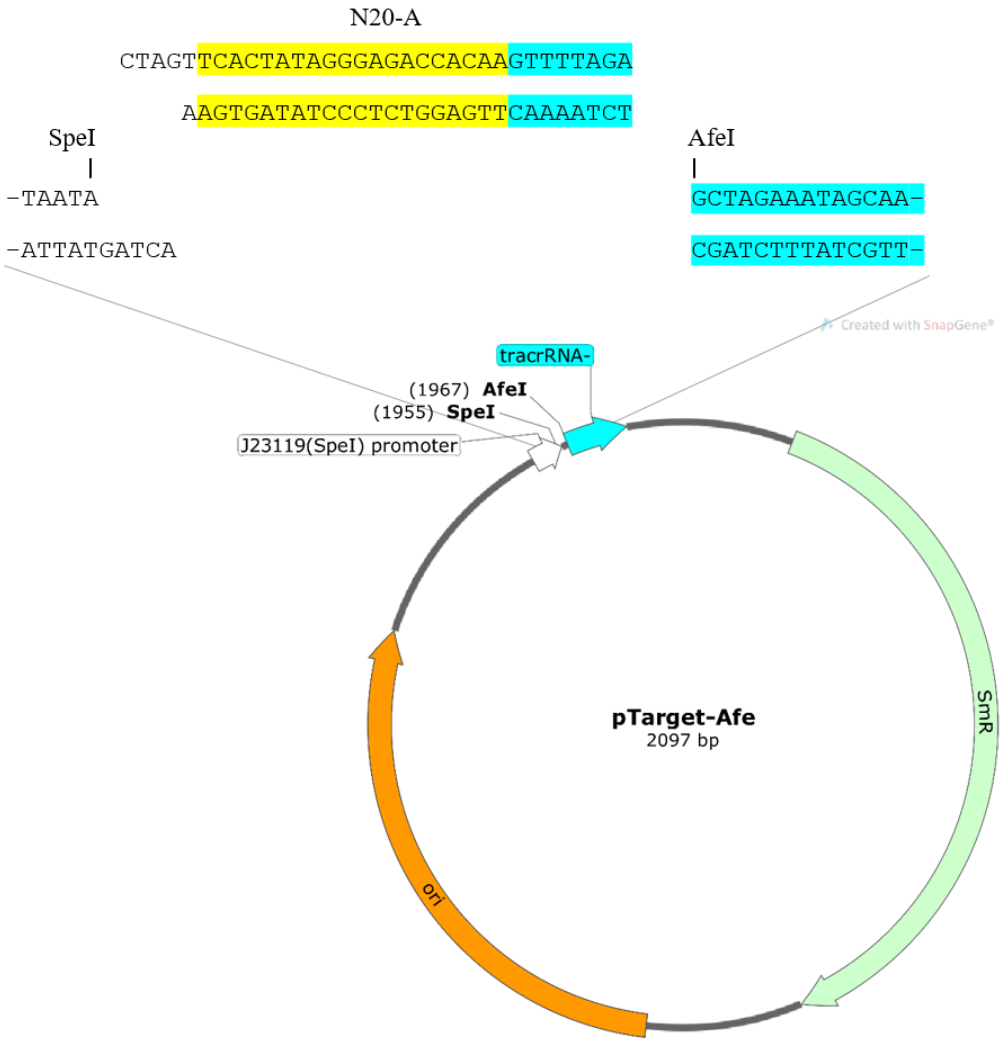


Figure 3.8. Cloning schematic of pTarget-N20-A plasmid. Promoter region of ED656-3B including the N20-PAM is shown ('A-site'). *Primers were designed to hold the N20 target sequence flanked by linkers. N20 insert was synthesized by self-annealing forward and reverse N20 primers and cloned into pTarget-Afe by restriction with SpeI and AfeI ('N20').*

The same method was followed to clone the B-site N20 into pTarget-Afe, as shown below in Figure 3.9.

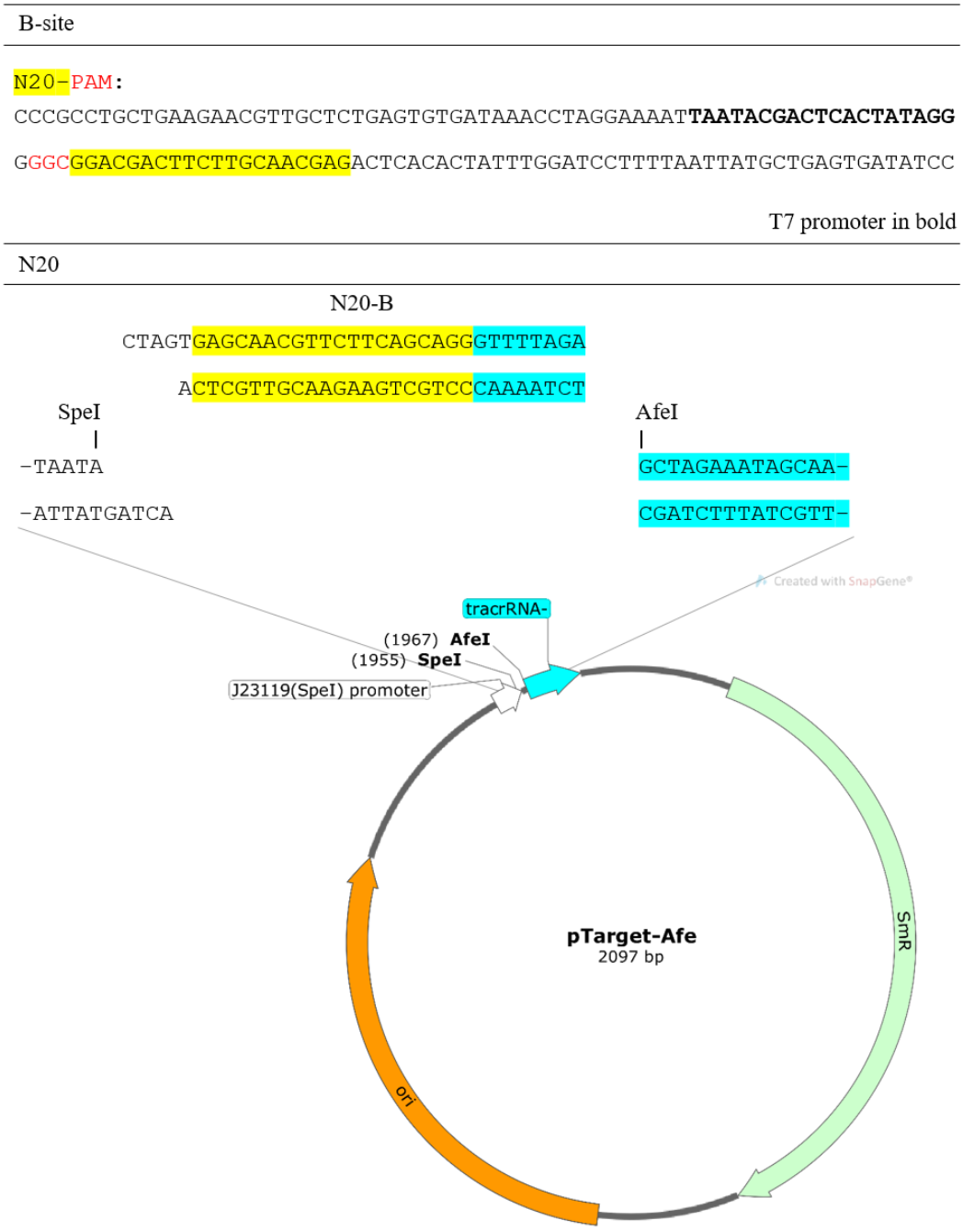


Figure 3.9 Cloning schematic of pTarget-N20-B plasmid. Promoter region of ED656-3B where the T7 promoter is indicated in bold, the N20-PAM is seen on the – strand ('B-site'). *Primers were designed to hold the N20 target sequence flanked by linkers. Insert was synthesized by self-annealing forward and reverse N20 primers and cloned into pTarget-Afe by restriction with SpeI and AfeI ('N20').*

3.3.2 Cloning the donor DNA

For homologous repair to take place after CRISPR-Cas9 cleavage, donor DNA was added that contained the desired change in DNA sequence. Additionally, the sequences contained 3' and 5' flanking regions that matched the genomic DNA, enabling homologous recombination. Here, the introduced DNA included the three constitutive promoters A2, A3 and M2 - representing high, medium and low expression levels, respectively. Plasmids containing these promoters were cloned to include the flanking regions. The 3' flanking regions were digested from plasmids pA2-*HmCherry*, pA3-*HmCherry* and pM2-*HmCherry* available in the lab. The 5'flanking regions were obtained via PCR, by amplification of the genomic DNA of ED656-3B.

3.3.2.1 A-site recovery DNA

Promoter vectors pA2-*HmCherry*, pA3-*HmCherry* and pM2-*HmCherry* were digested using XbaI and BamHI to remove *HmCherry*. The corresponding 3'flanking region was digested from plasmid pET3a-*RcCobA* using the same enzymes, and subsequently ligated to form p[promoter]-*RcCobA*. The 5' flanking region was amplified via colony PCR using primers Sal-nanC and BglII-nanC. The template for this reaction was strain ED656-3B. Donor DNA plasmids were completed by inserting the PCR products into the promoter-3' flanking region vectors using the restriction enzymes SalI and BglII followed by ligation. Figure 3.10 shows a schematic representation of the cloning process of p*NanC*-A2-*RcCobA*, the donor DNA containing the A2 constitutive promoter. The same process was followed for the other promoters, resulting in p*NanC*-A3-*RcCobA* and p*NanC*-M2-*RcCobA*.

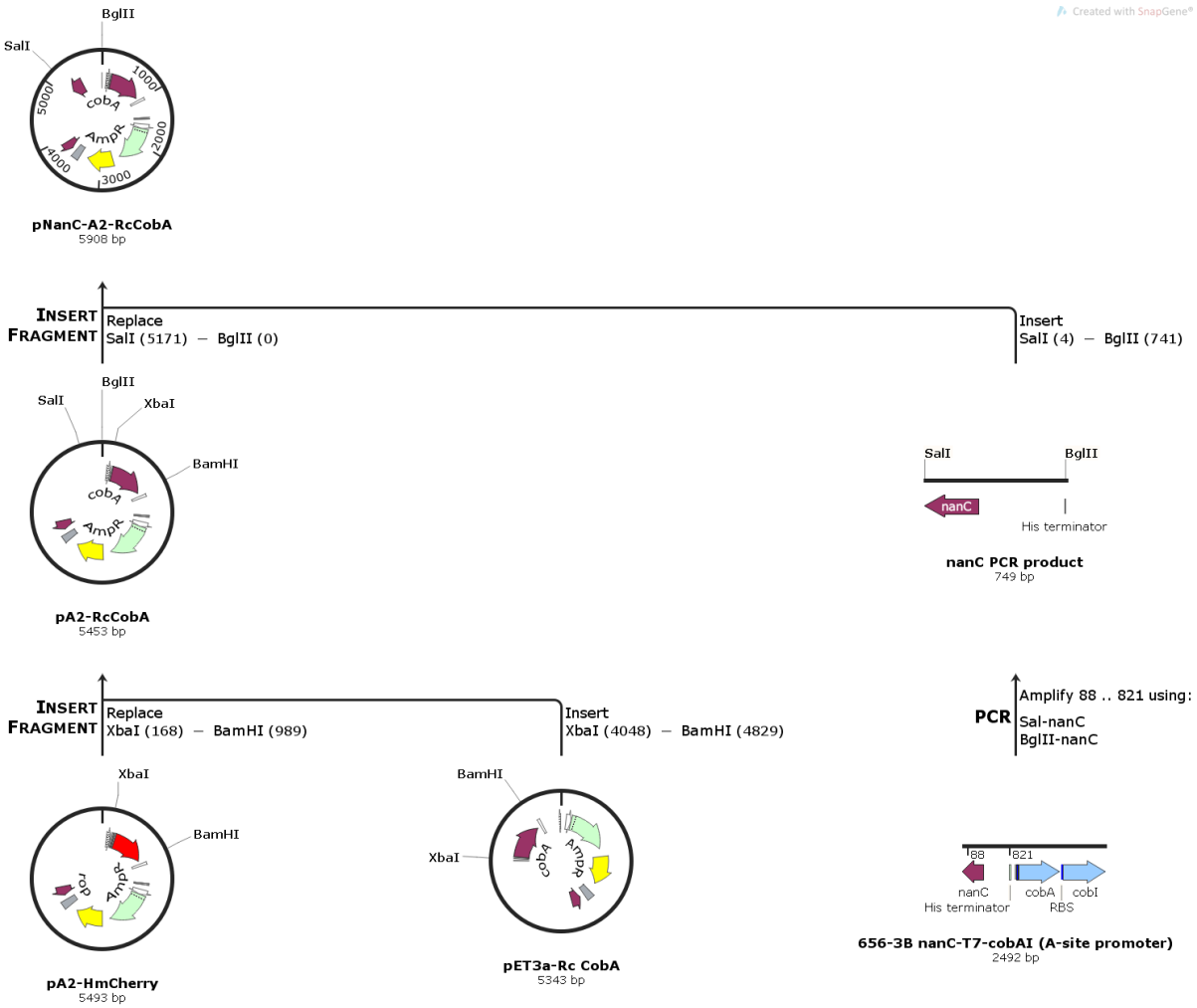


Figure 3.10 Schematic representation of the cloning process of donor DNA plasmid pNanC-A2-RcCobA. *pA2-HmCherry was adapted to replace mCherry with the 3'flanking region, Rc cobA, restricted from pET3a-Rc cobA with XbaI and BamHI forming pA2-RcCobA. The 5'flanking region was synthesized via colony PCR using primers Sal-nanC and BlgII-nanC with ED656-B acting as template. The PCR product was ligated into pA2-Rc cobA after restriction with SalI and BglII, forming final donor DNA plasmid pNanC-A2-RcCobA.*

3.3.2.2 B-site recovery DNA

For the cloning of the B-site donor DNA, the same procedure was followed as described in Section 3.3.2.1. The promoter vectors were digested with XbaI and BamHI to allow extraction of the corresponding 3' flanking region from plasmid pET3a-*Bmei708*. The resulting plasmids were called pA2-*Bmei708*, pA3-*Bmei708* and pM2-*Bmei708*. Using ED656-3B as template, the 5' flanking region was amplified via colony PCR using the primers SalI-Yoe2 and BglII-Yoe. Donor DNA plasmids were completed by inserting the PCR products into the p[promoter]-Bmei708 vectors using the SalI and BglII restriction enzymes followed by ligation. The cloning of pYoe-A2-Bmei708 is schematically represented in Figure 3.11. The other donor plasmids – pYoe-A3-Bmei708 and pYoe-M2-Bmei708 – were constructed in the same fashion.

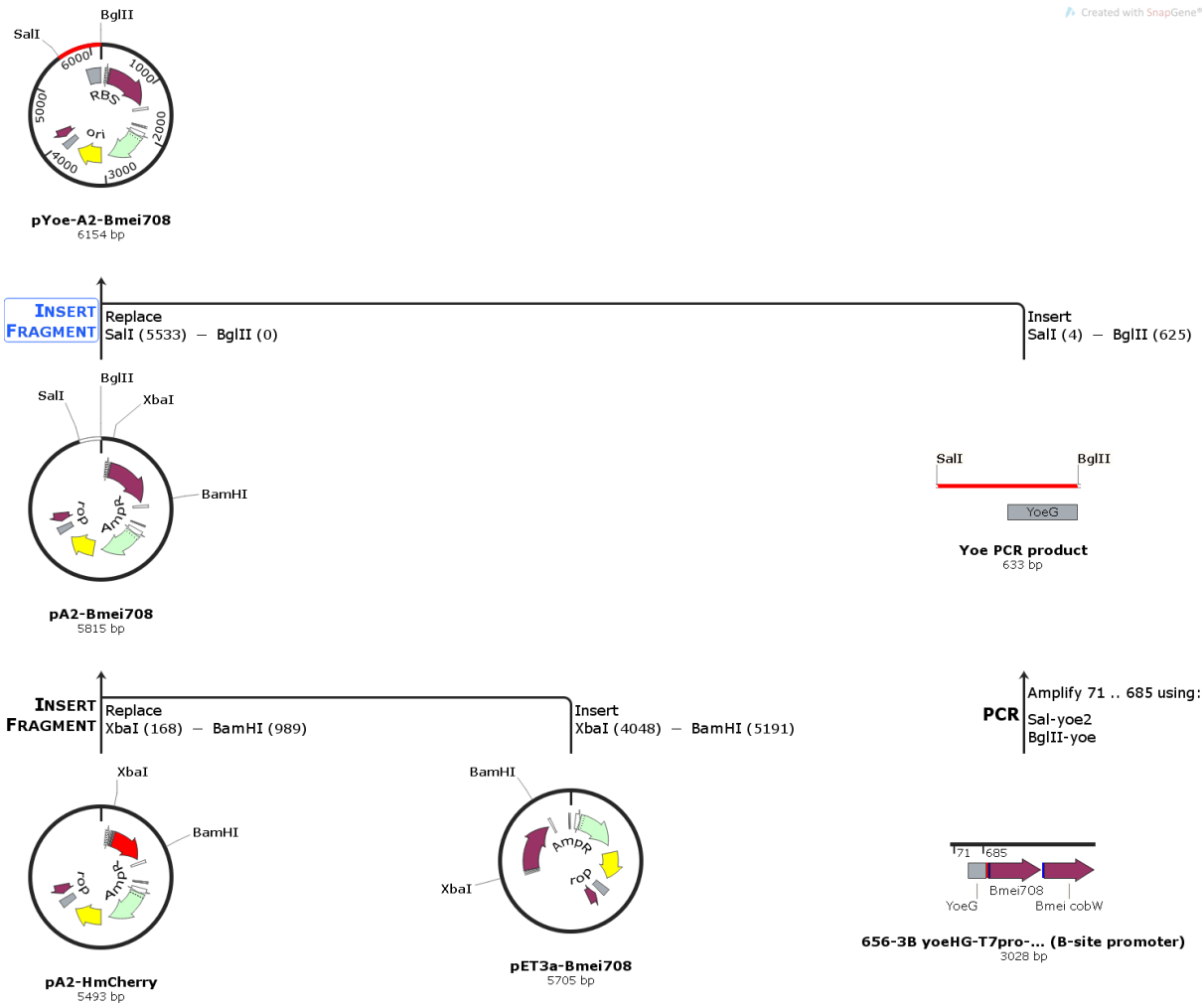


Figure 3.11 Schematic representation of the cloning process of donor DNA plasmid pYoe-A2-Bmei708. *pA2-HmCherry was adapted to replace mCherry with the 3'flanking region, Bmei708, restricted from pET3a-Bmei708 with XbaI and BamHI forming pA2-Bmei708. The 5'flanking region was synthesized via colony PCR using primers Sal-yoe2 and BlgII-yoe with ED656-B acting as template. The PCR product was ligated into pA2-Bmei708 after restriction with SalI and BglII, forming final donor DNA plasmid pYoe-A2-Bmei708.*

3.3.3 CRISPR transformation

ED656-3B (pCas) cells were prepared as described in Section 2.2.10 and transformed with the respective donor- and target DNA. The replacement of the T7 promoter by a constitutive one was tested in the resulting colonies via colony PCR. The forward primers were designed to bind upstream from the 5' flanking region to confirm the location of the PCR product. For the A-site, primers Site-A-F-check and cobA rev were used. Primers Yoe-check and Bmei708_rev were used to check the B-site constructs. ED656-3B was included as a control. The expected lengths of the PCR products are listed in Table 3.1. The resulting PCR products of the A and B site are seen in Figure 3.12 and Figure 3.13, respectively.

Table 3.1 Expected PCR product sizes of ED656-3B and all CRISPR-ed strains.

	A-site		*B-site*	
Strain	Promoter	Exp. Band size (bp)	Promoter	Exp. Band size (bp)
ED656-3B	T7	1076	T7	889
RS01	T7	1076	A2	940
RS02	T7	1076	M2	815
RS03	A2	1124	T7	889
RS05	M2	1000	T7	889
RS06_2	T7	1076	A3	939
RS07_1	A2	1124	M2	815
RS08	M2	1000	A2	940
RS09_1	M2	1000	A3	939
RS10_1	A3	1025	A2	940

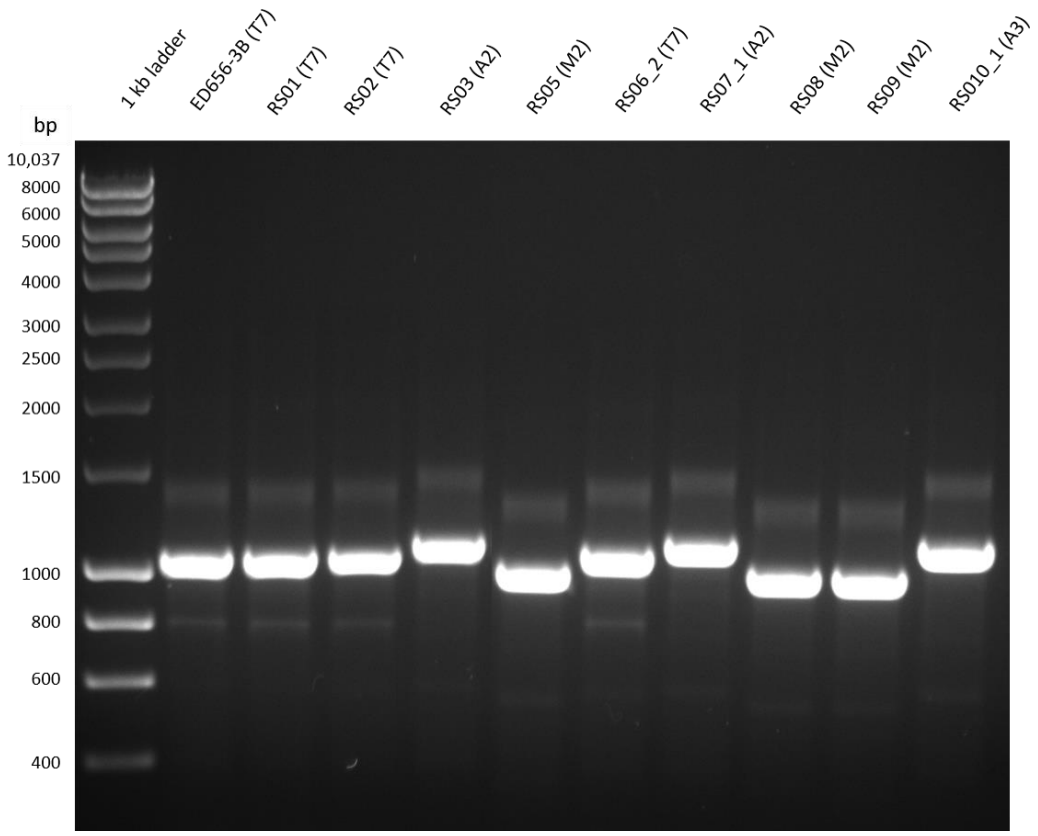


Figure 3.12. 1% Agarose gel of PCR products of A-site promoter region. *A colony PCR using primers Site-A-F-check and cobA rev was performed on all strains resulting the CRISPR-Cas9 transformation. The strains are denoted above the figure including which promoter was expected in brackets. ED656-3B was included as control for the T7 promoter. RS01, RS02 and RS06_2 show bands of equal size to that observed with ED656-3B (~1076 bp), indicating a T7 promoter in the A-site of these strains. Lanes of RS03, RS07_1 and RS10_1 show higher bands (~1124 bp) indicating either the A2 or the A3 promoter. Lanes RS05, RS08 and RS09 show bands around 1000 bp consistent with these strains containing the M2 promoter.*

Figure 3.12 shows the A-site promoter region PCR products are at a similar height when ED656-3B, RS01, RS02 and RS06_2 are used as template DNA, indicating these all contain a T7 promoter. The PCR product bands obtained when RS05, RS08 and RS09 DNA was used as template are observed to run lower, around 1000 bp, consistent with these strains containing an M2 promoter in this site as this is smaller. The bands which ran higher – as seen for RS03, RS07_1 and RS10_1 – point toward a successful change of T7 promoter to either the A2 or the A3 promoter, as these have run at a 1 bp difference.

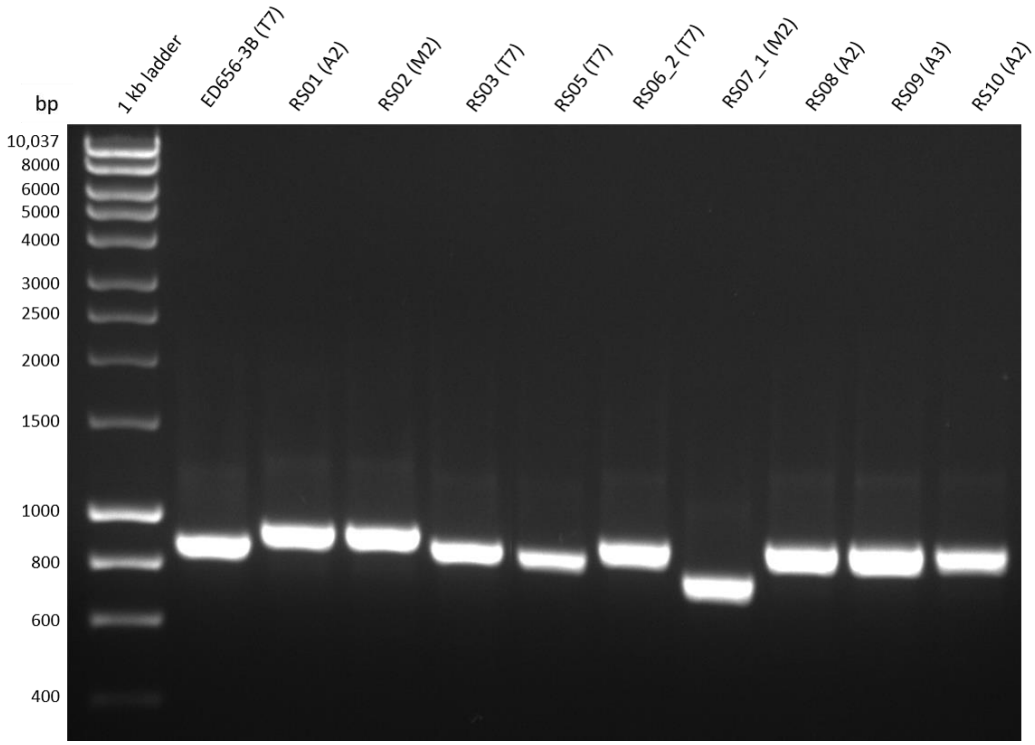


Figure 3.13. 1% Agarose gel of PCR products of B-site promoter region. *A colony PCR using primers Yoe-check and Bmei708 was performed on all strains resulting from the CRISPR-Cas9 transformation. The strains are denoted above the figure including which promoter was expected in brackets. ED656-3B was included as a control for the T7 promoter. RS03 and RS05 show bands at equal heights as that of ED656-3B (~889 bp), indicating a T7 promoter in the B-site of these strains. Lanes of RS01, RS06_2 and RS08-RS10_1 show higher bands (~940 bp) indicating the presence of either the A2 or the A3 promoter. Lanes RS02 and RS07_1 showed bands around 815 bp consistent with these strains containing the M2 promoter.*

The results of the colony PCR of the B-site promoter region are shown in Figure 3.13. The bands derived from strains ED656-3B, RS03 and RS05 all ran with a similar size of ~889, indicating that the T7 promoter had been retained in this site. The bands observed with strains RS01, RS06_2 and RS08 - RS10_1 migrated with a larger size, which again points towards a successful change of T7 promoter to either the A2 or the A3 promoter. The bands derived from strains RS02 and RS07_1 ran with a lower size at ~815, consistent with a successful change from T7 to the M2 promoter.

3.3.4 Sequence alignments

The PCR results indicated that all desired sites had been changed from the original T7 promoter to one of the constitutive ones. To confirm the correct changes, the strains were sequenced using primers cobA_rev and Bmei708_rev for the A- and the B-site, respectively. The resulting sequences were aligned with the sequence from the original strain ED656-3B, to determine if the T7 promoter has been changed. Figure 3.14 shows these alignments for the A-site of ED656-3B whereas Figure 3.15 shows the B-site alignments. In these figures the promoters are colour-coded; purple for T7, blue for A2, red for A3 and orange for M2.

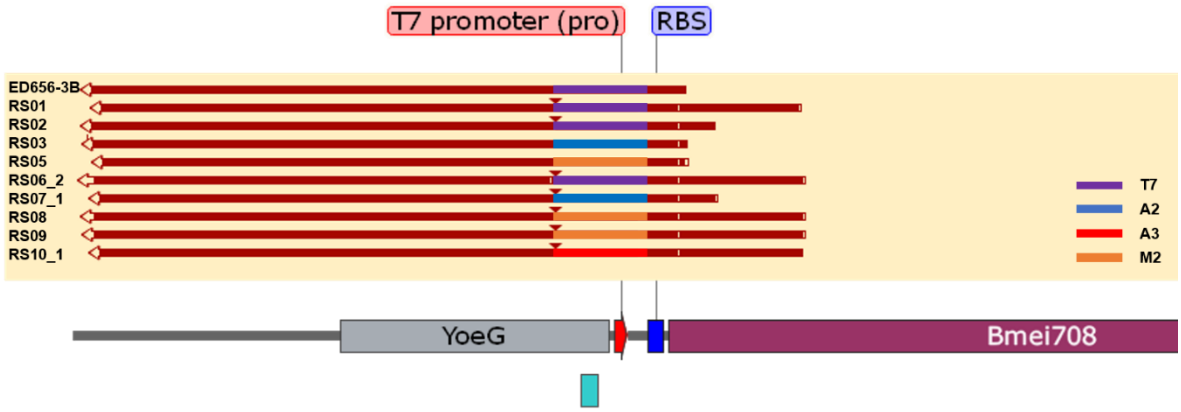


Figure 3.14: A-site sequence alignment. *All strains were sequenced and aligned with ED656-3B. ED656-3B was included as a control. Strains are listed to the left of the sequence arrows. Promoters are colour-coded purple for T7, blue for A2, red for A3, and orange for M2.*

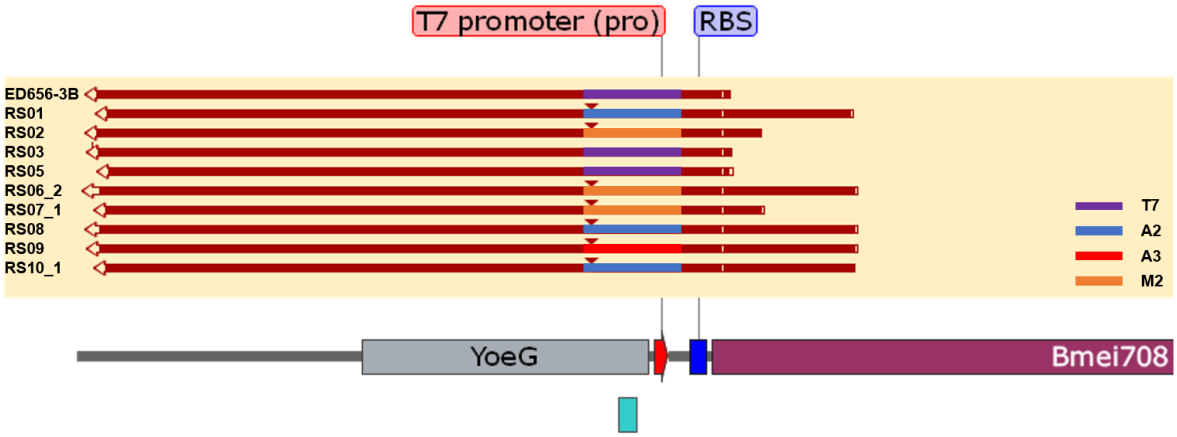


Figure 3.15 B-site sequence alignment. *All strains were sequenced and aligned with the original strain in SnapGene. ED656-3B was included as a control. Strains are listed to the left of the sequence arrows. Promoters are colour-coded purple for T7, blue for A2, red for A3, and orange for M2.*

Based on the alignments, the engineering of nine strains have been confirmed. These strains are listed in Table 3.2.

Table 3.2 Promoter combinations in the various *E. coli* strains as confirmed by sequencing.

Strain	*A-site promoter*	*B-site promoter*
RS01	T7	A2
RS02	T7	M2
RS03	A2	T7
RS05	M2	T7
RS06_2	T7	A3
RS07_1	A2	M2
RS08	M2	A2
RS09_1	M2	A3
RS10_1	A3	A2

3.4 Cobalamin production tests

To investigate the effect of the alternative promoters on the phenotype of the engineered *E. coli* strains, the cells were cultured in shake flasks containing YE-M9-G media (Section 2.1.5). This media was previously optimised to grow ED656-3B. Some media components have been shown to have metal binding properties (Ramamoorthy & Kushner, 1975), and in previous optimization studies addition of tryptone has been detrimental to B_{12} production, presumably by binding cobalt. Tryptone is therefore excluded from the media used herein.

Figure 3.16 shows a schematic representation of the process of growing the cultures and the analyses done to characterise the strains' phenotypes. This multifaceted characterisation includes monitoring the cell growth, measured via the optical density at 600 nm, and the determination of cobalamin production, measured via bioassay plates (Raux et al., 1996). The bioassay plates contain *Salmonella typhimurium* strain AR3612 (*metE*$^-$, *cysG*$^-$) or AR2680 (*metE*$^-$, *cbiB*$^-$). Wild-type *S. typhimurium* is able to synthesize cobalamin via the anaerobic pathway. The *cysG*$^-$ mutation in AR3612 blocks the beginning of the cobalamin pathway between uroporphyrinogen III and precorrin-2. The rest of the pathway remains functional when intermediates are scavenged. As *Salmonella* is unable to absorb intermediate prior to adenosylcobyric acid, AR3612 is able to grow on the intermediates between adenosylcobyric acid to adenosylcobalamin (B_{12}). The *cysG*$^-$ mutation in also leaves this strain unable to synthesize cysteine as the strain cannot produce siroheme, the prosthetic group associated with sulphite reductase, and hence cysteine is supplemented in the media. In the case of AR2680, the *cbiB*$^-$ mutation causes a block later in the cobalamin pathway, which means that the first intermediate detected by this strain is cobinamide. The *metE*$^-$ mutation disables the cobalamin-independent methionine synthase, thereby making the strain dependent upon the presence of MetH, the cobalamin-dependent methionine synthase. Consequently, the strain is dependent on the presence of exogenous cobalamin for its survival.

Additionally, expression from the A-site was analysed using β-gal assays and SDS-PAGE, whereas the B-site expression was monitored by western blot. The A-site operon ends with a *lacZ* gene, enabling quantification of expression levels in this site. Visualization via western blot of the B-site CobN enzyme was made possible due to the presence of a hexa-histidine tag. These analyses were performed at 24 hours and 48 hours after inoculation, to test for endurance differences between the stronger and weaker promoters.

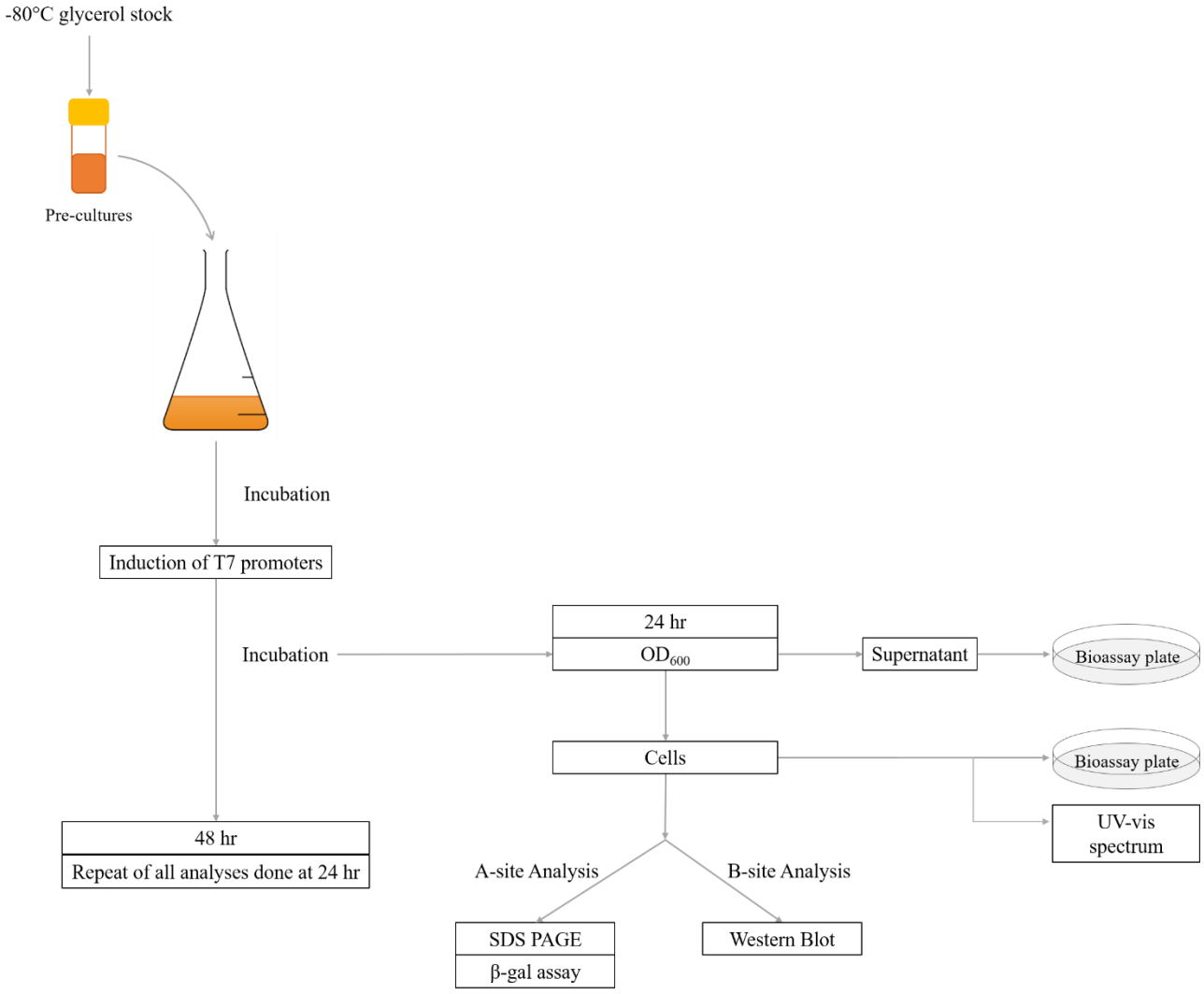


Figure 3.16. Schematic representation of multifaceted strain analysis. *B_{12}-producing strains were cultured on YE-M9-G media and growth and cobalamin production were measured after 24 h and 48 h. A- and B-site protein expression was measured by a combination of B-gal assay, SDS PAGE and western blot.*

3.4.1 Cell growth

The cell growth – and lifespan – of the engineered strains were monitored by measuring the optical density spectrophotometrically at 600 nm at 24 h and 48 h post inoculation (Figure 3.17). The engineered strains were cultured on M9-YE-G media in flasks which were shaken to enhance oxygenation. The cultures were grown at 28 °C to accommodate the genes of the biosynthetic B_{12} pathway which originate from *Rhodobacter capsulatus*. After 6 hours, expression under the T7 promoters was induced by adding IPTG. To test the differences in cell growth, a mixed analysis of variance was conducted in open source statistical software JASP, with cell growth (OD_{600}) as a dependent variable, time (24h vs 48h) as repeated measure, and strain as the 'between subject' - variable. No significant interaction effect between strain and time was found, indicating the change in promoters have no effect on cell growth.

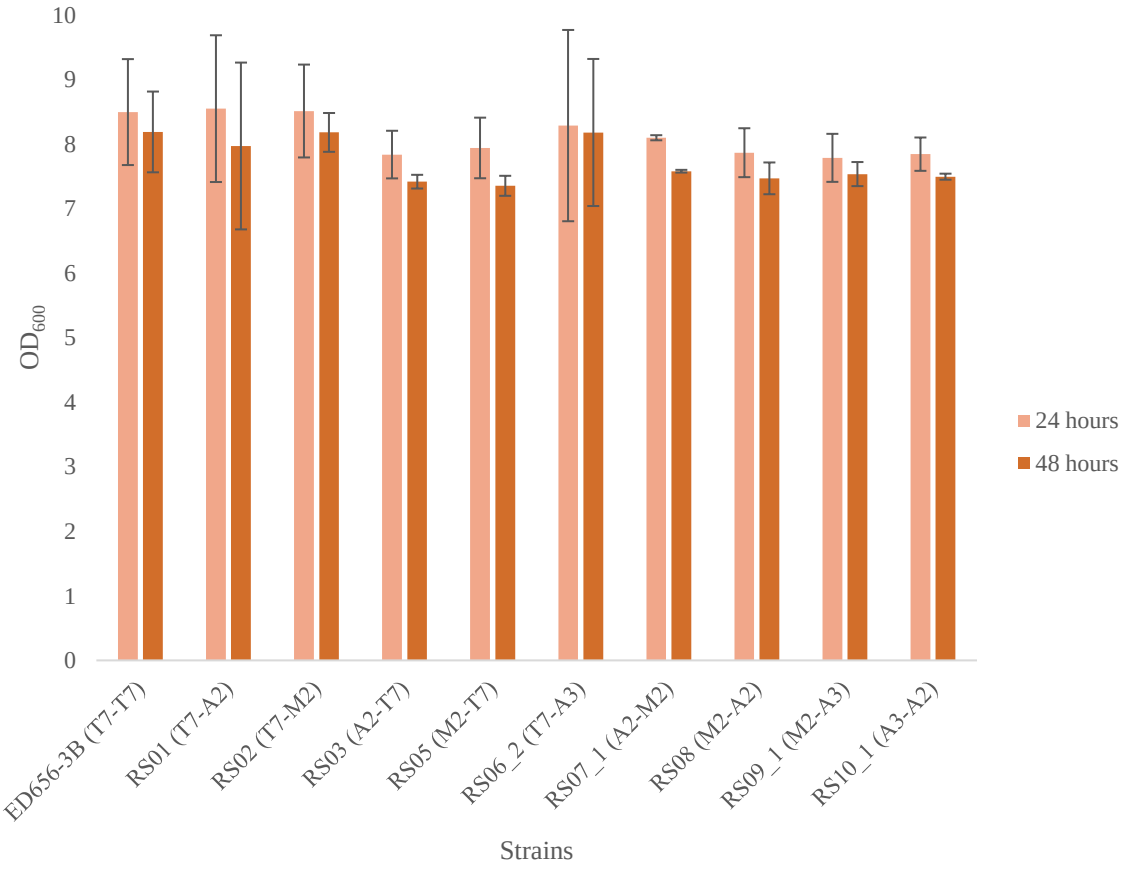


Figure 3.17 Optical cell density of cobalamin-producing cultures at 24 and 48 h post inoculation. *Strains were grown in duplicate on YE-M9-G media and the T7 promoters were induced with 0.4 mM IPTG 6 h after inoculation. Promoters controlling the A- and B-site are indicated in brackets, at position 1 and 2, respectively. Error bars indicate the SD of the values from the biological duplicates.*

The pellets of the harvested cultures after 48 hours are seen below in Figure 3.18. There was a clear colour difference between the cultures that contained a T7 promoter in their A-site, and those where the A-site promoter was changed to a constitutive one. This is visible as a slight pink colouration in the pellets of the A-site T7 cultures, whereas the remaining strains had a deeper brown colour. This change in colour could be explained by a difference in cobalamin-concentration, or the accumulation of one of the intermediates of the pathway.

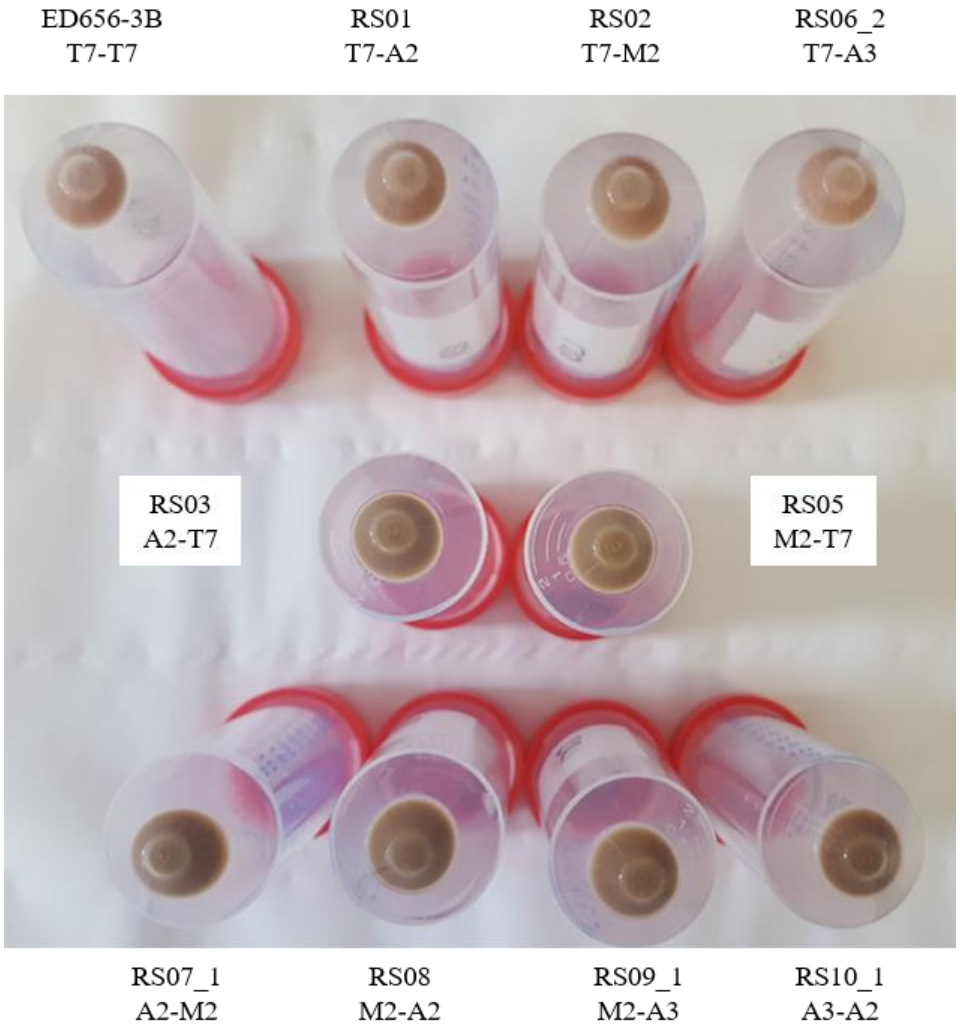


Figure 3.18 Cell pellets of B_{12} - producing cultures, harvested after 48 hours. *The strains were cultured in 25 mL YE-M9-G media in 250 mL shake flasks, at 28°C. Promoters in the A-and B-site are denoted underneath each strain, respectively.*

3.4.2 Cobalamin production

The cobalamin yield of the engineered strains was determined by way of bioassay plates. The cultures were harvested by centrifugation to allow the media to be separated from the cells. The cell pellets were resuspended in buffer and lysed by boiling. Supernatant samples were applied directly onto *Salmonella* bioassay plates which were cultured at 37° C overnight. Cell-extracted samples were diluted to fall within the range of the standard curve, which was generated from the application of known amounts of cobalamin to the bioassay plate. The resulting diameters of the growth patches were measured and values were converted to concentration using a standard curve derived from a standard plate (Figure 3.19). The control plate is essential as the results will vary day to day depending on the age and state of dehydration of the plates.

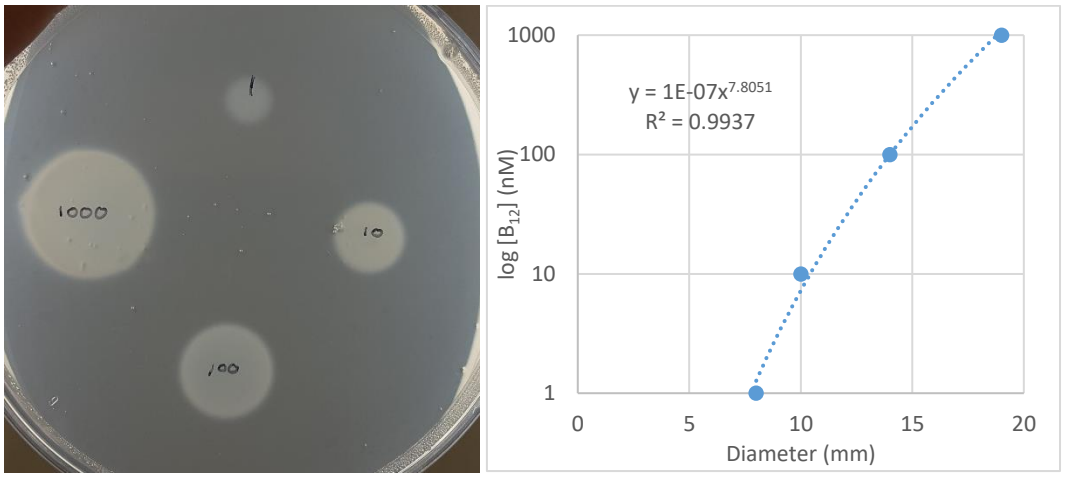


Figure 3.19. *Salmonella* bioassay standard plate (left) and derived standard curve (right). *10 µL of 1 nM, 10 nM, 100 nM and 1000 nM B_{12} standard solutions were dropped on an M9 minimal media agar plate with S. typhimurium AR2680. Diameters of resulting patches were measured and plotted to give a standard curve.*

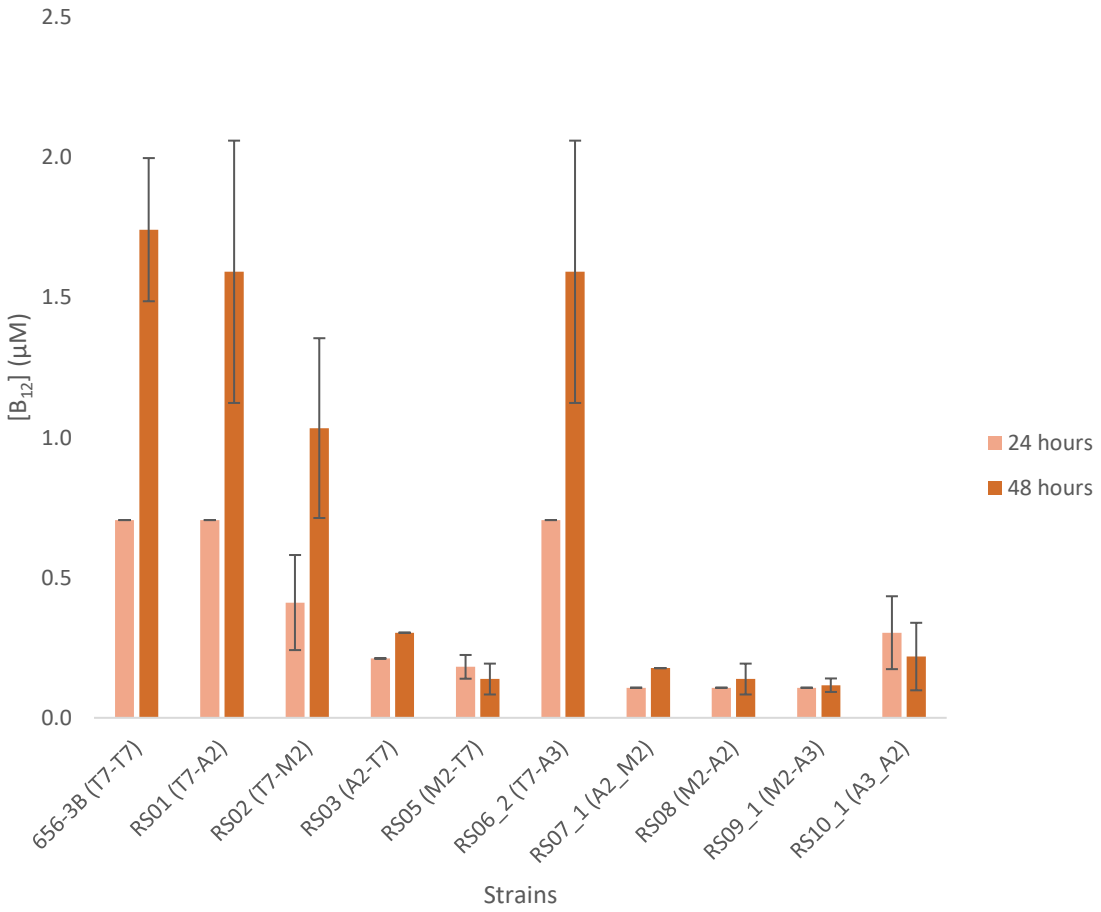


Figure 3.20. Cobalamin production of engineered strains, cultured on YE-M9-G media. *Yield measured in duplicate at 24 h and 48 h after inoculation via Salmonella bioassay plates. Error bars indicate the SD of the values from the biological duplicates.*

Strains ED656-3B, RS01, and RS06_2 produced very similar levels of cobalamin. RS02 produced intermediary level of B_{12}, somewhat lower than the top producers but significantly higher than the poorest producers. It is notable that any strain in which the A-site T7 promoter was changed to a constitutive one the cobalamin production was observed to be very low. To test the differences, a mixed analysis of variance was conducted in open source statistical software JASP, with [cobalamin] (μM) as a dependent variable, time (24 h vs 48 h) as repeated measure, and strain as the 'between subject' - variable. A significant interaction effect between strain and time was found; $F = 6.850$, $p = 0.003$, which warranted for further inspection of post-hoc comparisons. Post-hoc comparisons (with Bonferroni correction) showed that at 24 hours none of the strains significantly differed from the original ED656-3B strain, however, at 48 hours, all but RS01, RS06_2 showed a significant decrease of cobalamin production when compared to ED656-3B. All p values were < 0.001.

The size of the error bars must be mentioned here. There is an exponential relationship between the size of the growth circle and the concentration of cobalamin that is added. Consequently, a small difference in the measurement of the growth circle correlates with a large change in concentration. The diameter of the growth circles is visualised in Figure 3.21. These diameters are measured manually and are therefore prone to human error/inaccuracy. The errors are largest after 48 hours. This difference in results may be explained by the strains' variance in T7 'shut-off' time point. High expression under a T7 promoter causes high levels of metabolic stress. *E. coli* is known to mutate to end T7-controlled expression to avoid cell death (Rosano & Ceccarelli, 2014). This theory is strengthened by the fact that the largest errors are seen in strains in which the T7 promoters control the A-site – where it causes the highest metabolic strain. It is evident this method of establishing a true cobalamin concentration is flawed, however for the purposes of comparative production yields between the different engineered strains, it is sufficient.

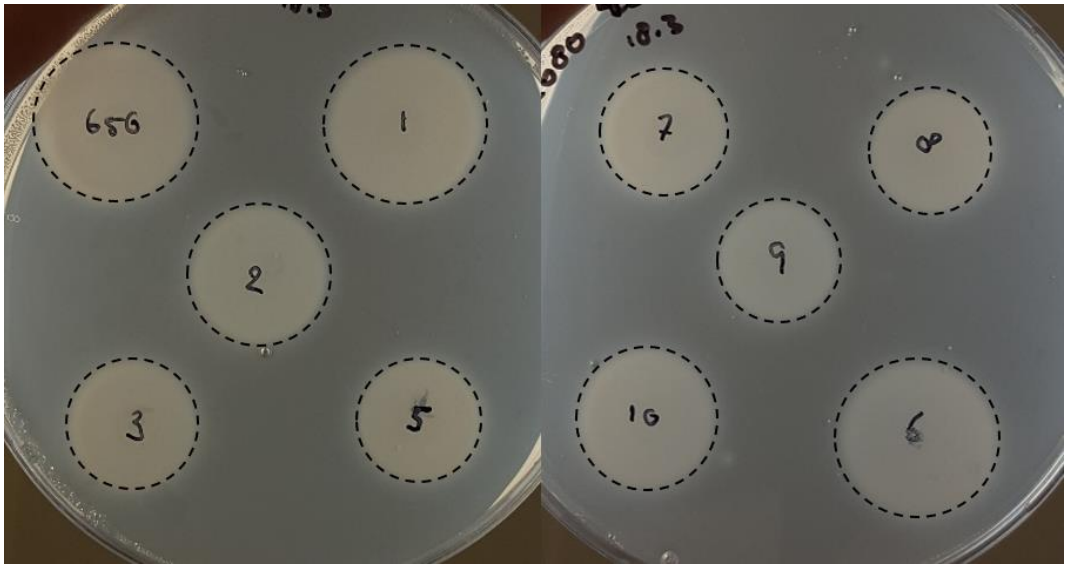


Figure 3.21 *Salmonella* AR2680 bioassay plates with B_{12}-producting strain samples after 48 h of inoculation. *Engineered B_{12}-producing strains were cultured for 48 hours, cells lysed by boiling and samples dropped on the plates. Incubation at 37° C revealed growth circles relative to the production of cobalamin of these strains.*

3.4.3 A-site analysis

The initial characterisation focused on cell growth and B_{12} production. This section zooms in on the expression levels displayed in the A-site as a result of the change in promoter controlling the A-site operon.

3.3.2.3 SDS-PAGE

Characterisation of protein expression patterns was performed via SDS PAGE. Culture samples of the 24 h and 48 h time points were OD corrected and lysed by boiling before being run on a 10% denaturing gel. The protein bands were visualised with coomassie blue stain. During the loading of the 24-hour gel, high viscosity - indicating a high concentration of DNA – complicated the loading process. This resulted in some streaking on the gel image. Incubating the 48-hour time point samples with benzonase relieved this issue, as is seen in the difference of band clarity between the gels in Figure 3.22 and Figure 3.23.

The gels show a difference in expression levels between those strains whose A-sites are under the control of a T7 promoter and those with a constitutive promoter. This difference in band pattern is visualised as thicker bands around 80-85 kDa, 44 kDa and 30 kDa as seen in lanes 2, 3, 4 and 7. The 30 kDa band – denoted with the red number 3 in the gel – likely

represents the methylases of the B_{12} aerobic pathway which are all around this size and therefore indistinguishable from one another on a gel (Schubert et al., 2003). The 44 kDa protein (number 2) may be CobG, the oxidoreductase that catalyses the ring contraction step in the aerobic pathway (Debussche et al., 1993). Another candidate is CobB, at 45 kDa, the enzyme that is responsible amidating hydrogenobyrinic acid and the *a* and *c* side chains (Debussche et al., 1990). At 51 kDa, CobQ is another possible candidate for this band, representing cobyrinic acid synthase, which catalyzes the amidations of side-chains *b*, *d*, *e* and *g* at the final stage of the aerobic pathway (Leeper et al., 2012). The band at 80 kDa (number 1) cannot be identified as one of the B_{12}- pathway enzymes. This band may represent an *E. coli* stress response protein such as GroEL. Though described as a heat-shock protein, the expression has been shown to be essential to *E. coli* even at temperatures as low as 17° C (Fayet et al., 1989). These experiments were duplicated with the same results.

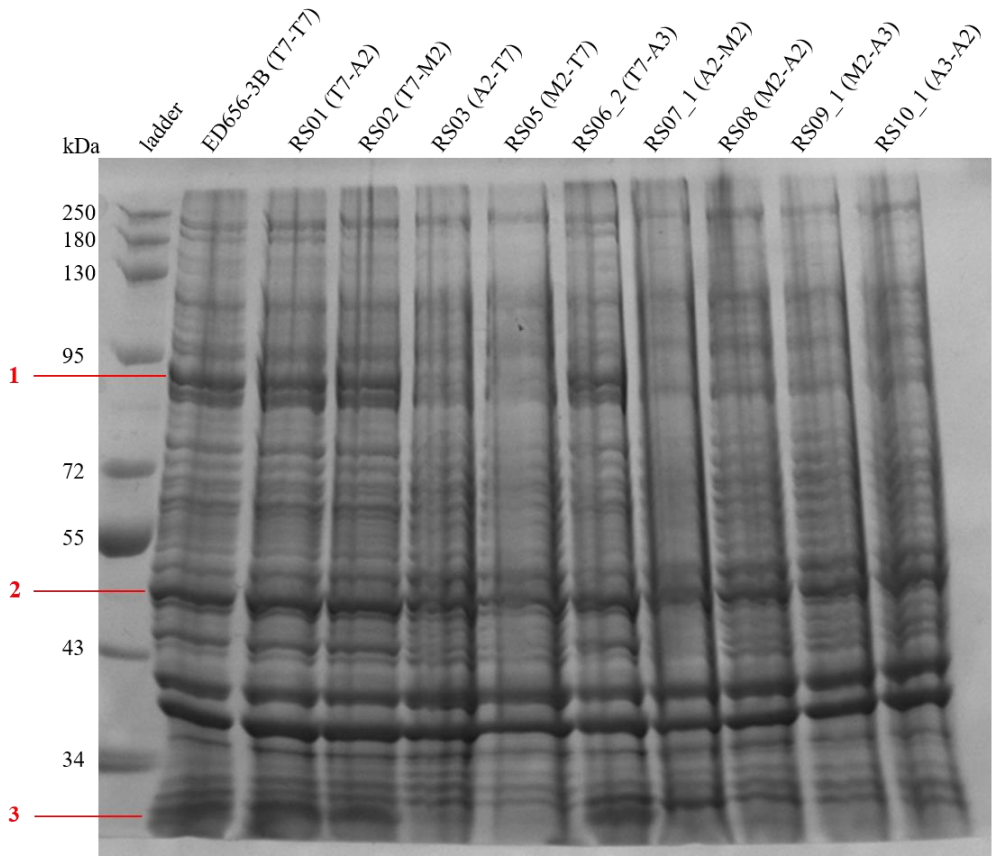


Figure 3.22 10% SDS PAGE gel of B_{12}- producing strain cultures after 24 h. *Strains are denoted above with the promoters controlling that strain's A-and B-site in brackets in the first and second position, resp. Culture samples were OD corrected and heated to 95°C in 1:1 SDS sample buffer. Protein bands were visualized with coomassie blue stain. Red numbers indicate areas of expression differences between different lanes.*

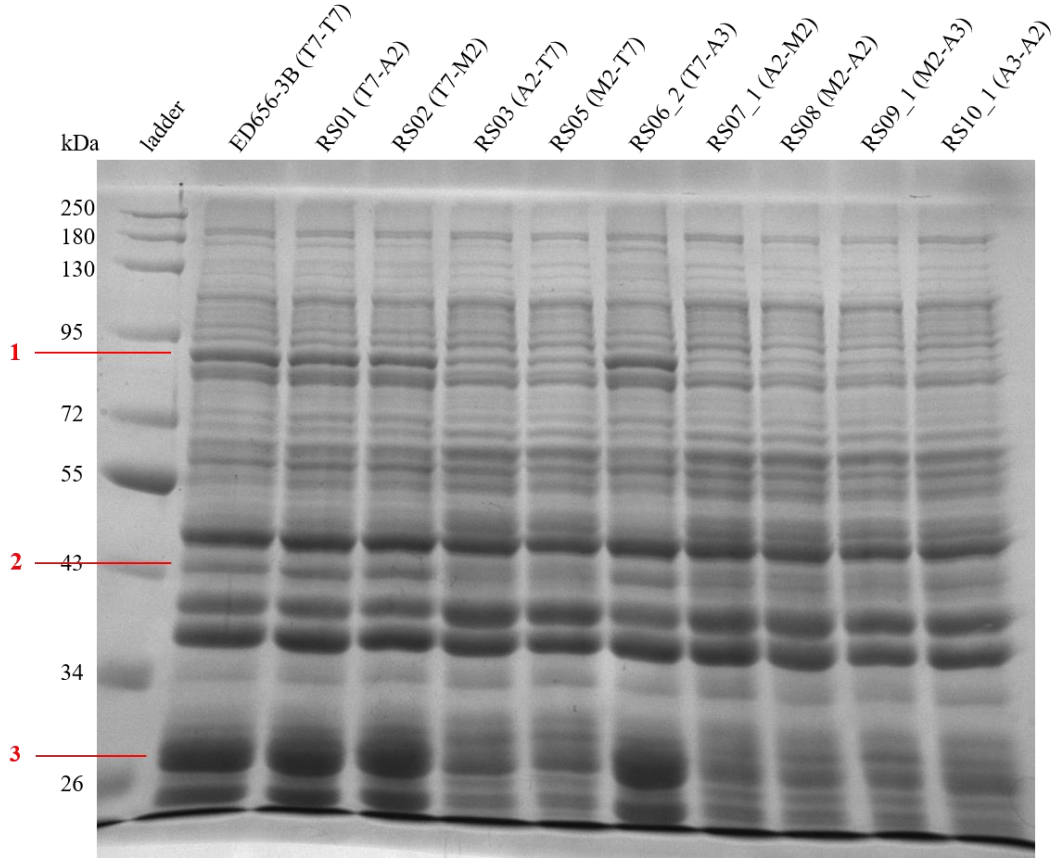


Figure 3.23 10% SDS PAGE gel of B_{12}- producing strain cultures after 48 h. *Strains are denoted above with the promoters controlling that strain's A-and B-site in brackets. Samples were OD-corrected and treated with benzonase to combat DNA aggregation. Culture samples were heated to 95°C in 1:1 SDS sample buffer. Protein bands were visualized with coomassie blue stain. Red numbers indicate areas of expression differences between different lanes.*

3.3.2.4 β-galactosidase assays

The A-site operon contains a *lacZ* gene at the end of the operon. This allows an opportunity for a quantitative measurement of expression levels of the operon. *lacZ* encodes β-galactosidase, an enzyme that cleaves lactose, a disaccharide, into glucose and galactose. When using the synthetic substrate o-nitrophenyl-β-galactoside, the cleavage releases the chromogenic nitrophenol as well as galactose. The product formation can therefore be followed spectrophotometrically and will be directly in line with the amount of β-galactosidase being expressed, allowing differences in expression levels of the A-site between the different strains to be determined.

The assay samples were taken from cell cultures at 24 h and 48 h after inoculation of the B_{12}-producing strains, as described in Section 3.4.1. β-galactosidase assays were executed via an adapted Miller protocol (X. Zhang & Bremer, 1995). Centrifugation of the reaction mix before measuring the A_{420} allowed for a discard of the A_{550} cell debris factor of the original equation. The Miller units were calculated using Equation 1, where v represents the volume of culture added in mL and t the reaction time in minutes. The absorbance at 420 measures the amount of product formation of o-nitrophenol as the substrate, o -nitrophenyl-β- D – galactopyranoside (ONPG), is cleaved by β-galactosidase.

Equation 1

$$Miller\ Units = 1000 * \frac{A_{420}}{A_{600} * v * t}$$

Figure 3.24 shows the expression levels varied most after 24 hours, where a distinct difference is seen between those strains which still contain a T7 promoter and those which were changed to constitutive promoters. ED656-3B, RS01, RS02 and RS06_2 demonstrated the highest expression levels at the A-site. To test the differences observed, a mixed analysis of variance was, with β-gal activity (Miller Units) as a dependent variable, time (24h vs 48h) as repeated measure, and strain as the 'between subject' - variable. A significant interaction effect between strain and time was found; $F = 10.331$, $p < 0.001$, which warranted further inspection. Post-hoc comparisons (with Bonferroni correction) showed that at 24 hours, all but strains RS01, RS02 and RS06_2 showed a significantly lower β-gal activity than the original strain, ED656-3B ($p < 0.001$). After 48 hours, however, no significant difference in activity was measured between the strains.

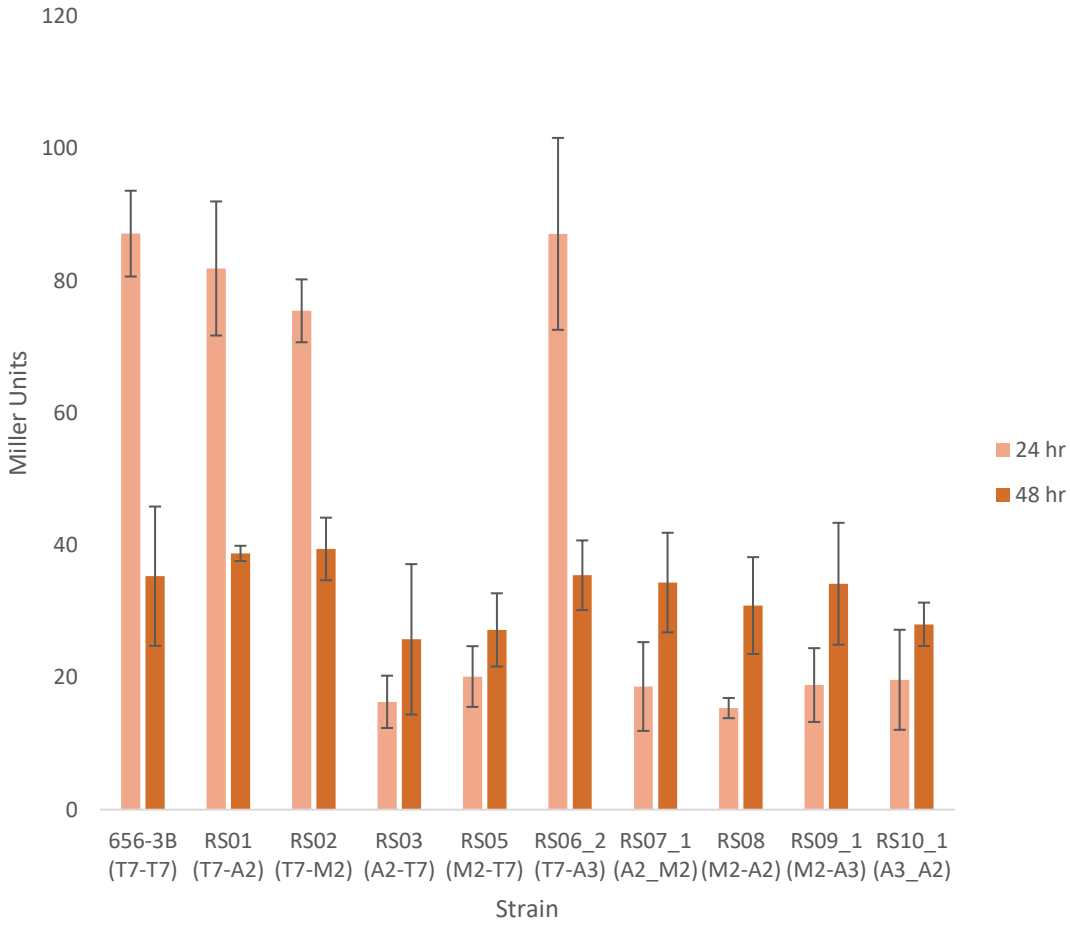


Figure 3.24 B galactosidase activity measured in the A-site operon of the engineered B_{12}-producing strains. *Samples were taken at 24 h and 48 h after inoculation of cultures. Levels of O-nitrophenyl were measured at 420 nm, indicating activity of β-galactosidase. Measurements done as biological duplicates, with error bars indicating the SD of the two values. The A-site expression of those operons still under the control of a T7 promoter seem to have the highest β-gal activity after 24 hours, as seen by strains ED656-3B, RS01 RS02 and RS06_2. After 48 hours, no significant difference was measured between all strains.*

3.4.4 B-site analysis

The B-site residing Bmei CobN protein (139 kDa) contains an N-terminal hexahistidine tag. This enabled the visualization of protein expression levels via western blot using an anti-his tag antibody. Samples were OD-corrected and run on SDS PAGE as described in Section 3.4.3 before being transferred to a nitrocellulose membrane. Incubation with the primary antibody, mouse α-His IgG, was followed with the secondary rabbit α-mouse IgG antibody – peroxidase conjugate. Bmei ^{6xhis}CobN was visualised by incubation with 5-bromo-4-chloro-3-indolyl phosphate. The resulting western blots are seen in Figure 3.25 (24 h) and Figure 3.26(48 h).

Figure 3.25. Western Blot of B_{12}- producing strain cultures after 24 h. *Culture samples 24 h after inoculation were normalised to OD_{600} 1.0 and cells were lysed via boiling 1:1 in SDS PAGE sample buffer. A 10% SDS PAGE gel was run on SDS PAGE and proteins were transferred to nitrocellulose membrane. Hexahis-tagged Bmei cobN (139 kDa) in the B-site was visualised via incubation with primary mouse α-his IgG, subsequent α- mouse IgG-AP and chromogenic BCIP substrate.*

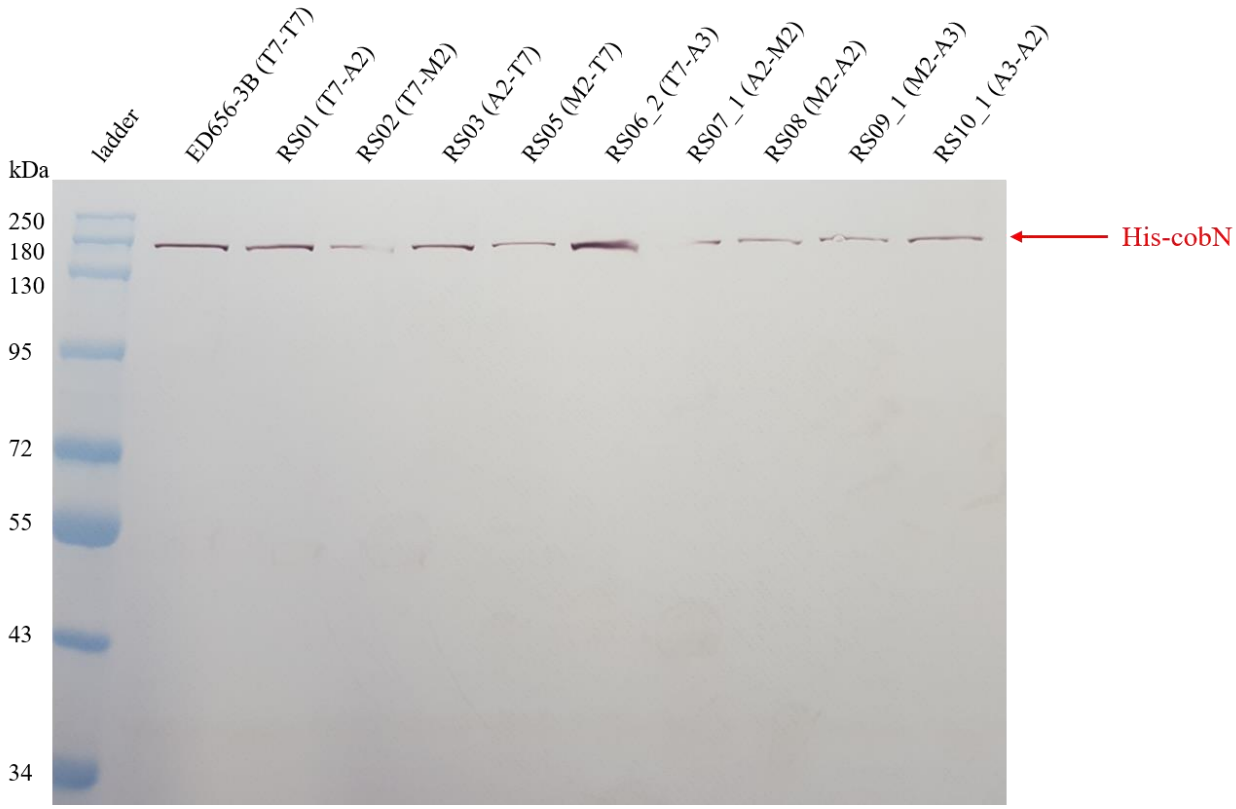


Figure 3.26. Western Blot of B_{12}- producing strain cultures after 24 h. *Culture samples were taken 48 h after inoculation and normalised to OD_{600} 1.0. Sample viscosity was relieved via benzonase incubation before lysing the cells via boiling 1:1 in SDS PAGE sample buffer. A 10% SDS PAGE gel was run and proteins were transferred to nitrocellulose membrane. Hexa-His-tagged Bmei CobN (139 kDa) in the B-site was visualised via incubation with primary mouse α-his IgG, subsequent α- mouse IgG-AP and chromogenic BCIP substrate.*

The band patterns are similar after 24 h to that seen on the western blot of the 48-h samples, indicating no more CobN has been expressed after 24 hours. One exception to this seems to be RS03, where the band is more distinct after 48 hours than is seen on the 24-hour western blot. Overall, however, the constitutive promoters do not seem to have elongated this expression time frame.

There does seem to be a slight difference in expression strength between some of the strains, however. ED656-3B, RS01 and RS06_2 have stronger bands than the other strains, indicating higher expression of the B-site CobN. This site is under the constitutive A2 promoter in RS01 which is the strongest of the constitutive promoters, but weaker than the T7 promoter. Even more unexpected is this result in RS06_2, which contains the weakest constitutive promoter in its B-site; A3.

3.4.5 Pathway flow analysis

Previous work has indicated a build-up of the intermediate hydrogenobyrinic acid *a,c*-diamide (HBAD) in ED656-3B. One of the main reasons of varying the promoter strength in the A- and B site operons was to enhance pathway flow, thereby relieving this constriction effect within the pathway. The samples were grown as described in Section 3.4, and were tested for HBAD build-up. Harvested cells were resuspended in buffer and lysed by boiling. The spectra were then taken of the soluble fraction of lysed cell extract samples (Figure 3.27).

The spectra reveal a distinct difference in those strains which contain an A-site T7 promoter and those that harbour an A-site operon controlled under a constitutive promoter. Those with constitutive promoters show a higher overall absorbance between 300 and 495 nm whereas the A-site T7-strains show an overall lower absorbance in this area, with a more distinct peak around 400 nm. This peak is consistent with porphyrins. However, the data does not show a clear spectrum for HBAD, which would be visible with peaks at 328, 495 and 522 nm. This is likely due to the presence of many impurities within the crude sample.

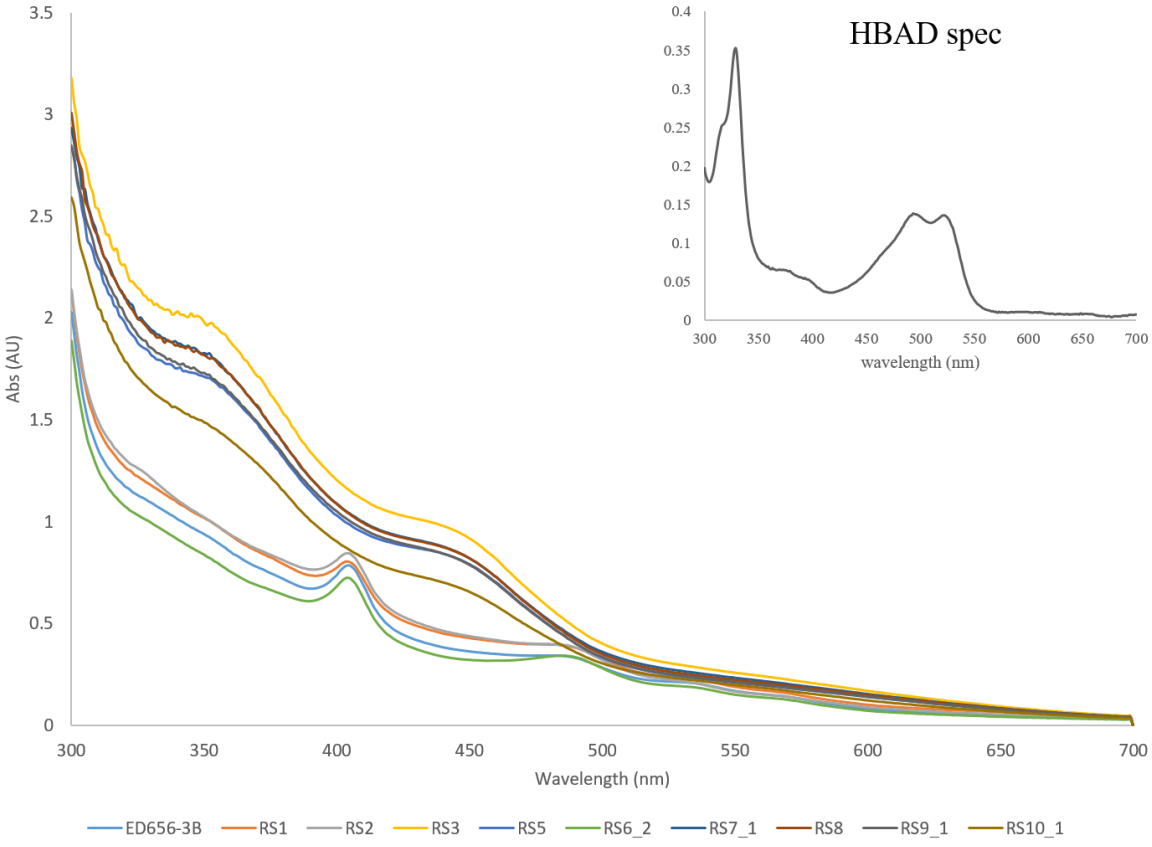


Figure 3.27 Spectra of soluble fraction of lysed B_{12}-producing strain cultures after 48 h. *Example of HBAD spectrum included. B_{12}-producing strains were cultured and cells were harvested and lysed by boiling in buffer. Spectra were taken of the soluble fraction. The absorbance spectrum of purified HBAD is given in frame.*

3.4.6 Fermentation cultures

Whilst shake flask cultures give a good indication of whether any one strain has developed the ability of producing a relatively higher yield of cobalamin, there are limits to this way of culturing. One such limitation on the cells life span and productivity is pH. During time-course experiments, the pH of the cultures is seen to decrease from 7.0 to 6.0 after 23 hours, and further plummets to 5.0 - 5.5 after 31 hours. This decrease is explained by the build-up of metabolites. Furthermore, the aeration is limited in flasks, even when shaking. Industrially, strains would be cultured in fermenters, where both factors would be regulated. A higher OD_{600} can be obtained, and production can be sustained for a longer period of time. It was therefore interesting to see whether this method of culturing makes a difference in the yield between ED656-3B and strains with alternative promoters. The set-up allowed for 3 strains to be cultured simultaneously. RS03 (A2-T7) was chosen to represent medium-strong constitutive expression on the A-site and T7-controlled expression of the B-site. RS08 was included to represent a strain where both sites were altered to constitutive promoters; the weaker M2 in the A-site and stronger A2 in the B-site. ED656-3B was included for comparative purposes.

The media used for fermentation culturing was YE-M9-GG, differing from the shake flask media by having added glucose. Whilst T7 promoter systems are known to be ‘leaky’, this addition ensures the T7 promoters are not expressing whilst the cultures obtain a higher cell density before being induced. The cultures were prepared as described in Section 2.2.11 where the pH was maintained at 7.0 and oxygen levels were maximised. The OD_{600} was taken periodically over the time span of 70 hours.

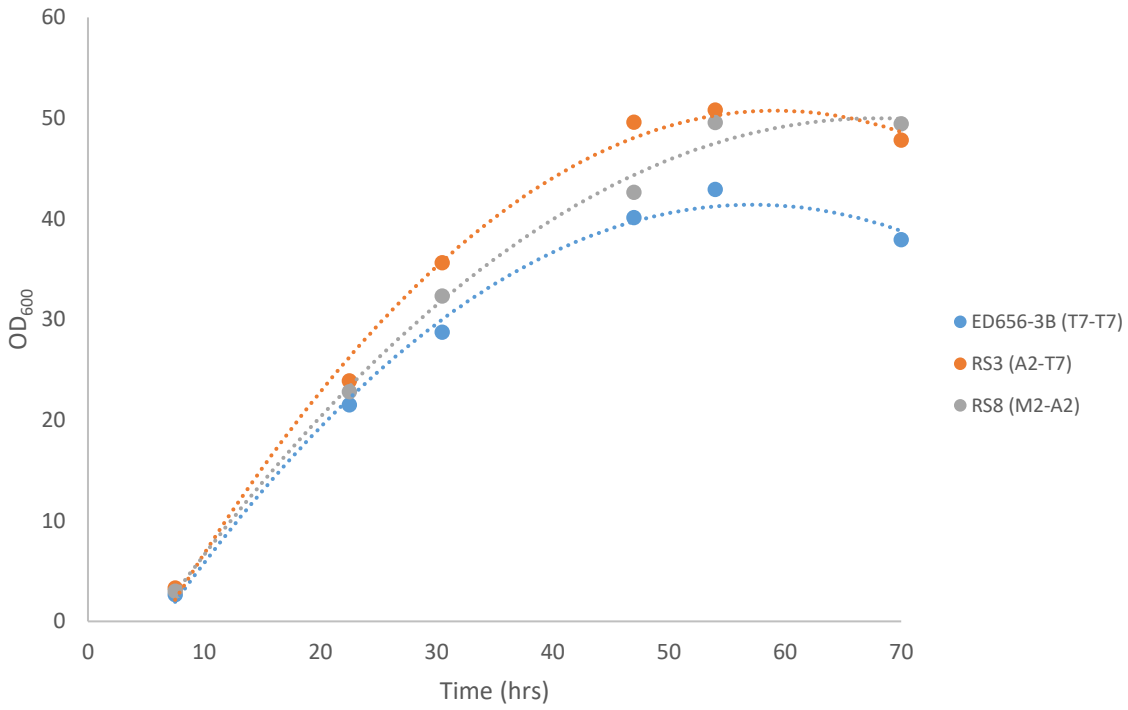


Figure 3.28 Cell density of cobalamin-producing cultures ED656-3B (blue), RS03 (orange) and RS08 (grey) over 70 hours of fermenter culturing. *Strains were grown on YE-M9-GG media, T7 promoters were induced with 0.4mM IPTG 6 h after inoculation.*

As seen in Figure 3.28, the strains all grew similarly to one another in the controlled conditions of the fermenters. ED656-3B came out at a slightly lower cell density, showing slower growth from the first time point after induction (at 22.5 h), to a higher reduction in OD_{600} - indicating cell death - at 70 h.

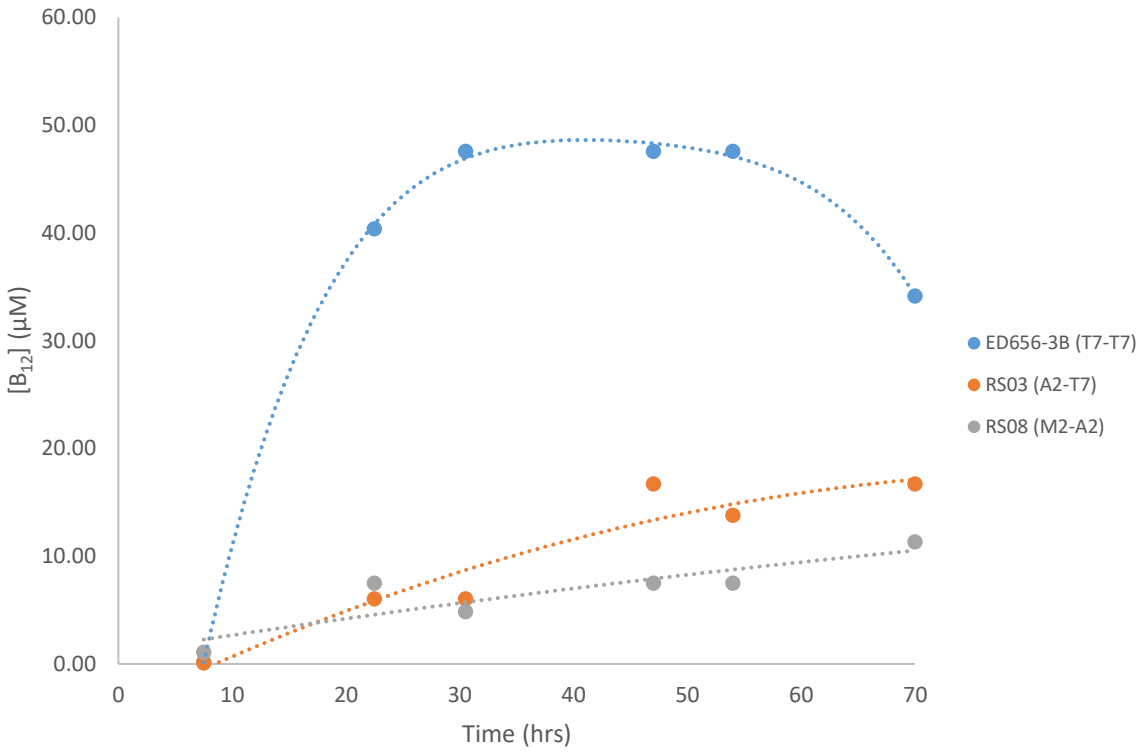


Figure 3.29 Cobalamin production of strains ED656-3B (blue), RS03 (orange) and RS08 (grey) over 70 hours of fermenter culturing. *Strains were grown on YE-M9-GG media, T7 promoters were induced with 0.4mM IPTG 6 h after inoculation. B_{12} yields were measured via Salmonella bioassay plates.*

The cobalamin yield was measured via bioassay plates as described in Section 3.4.2. The strains containing constitutive promoters showed a continuous increase in yield, even after 3 days, but do not come close to approaching that of ED656-3B. The weaker constitutive promoters found in RS03 (A2-T7) and RS08 (M2-A2) do not produce enough of the enzymes found in the A-site operon, which could explain this effect. Adding glucose ensures the T7 promoters are shut off and therefore, the culture is able to obtain a healthier state before inducing production.

3.5 Discussion

The successful engineering of 9 B_{12}-producing *E. coli* strains has been described. Adapted from ED656-3B, the resulting strains varied in the promoters controlling either the A – or B – site, or both. Constitutive promoters A2, A3, and M2 were chosen to represent high, medium and weak expression levels, respectively, to investigate the effect of expression levels on the phenotype as pertains to overall cell health, yield, and intermediate build-up.

The biggest impact on cobalamin yield is shown to be the A-site promoter. When the T7 promoter was changed to any of the constitutive promoters tested here, B_{12} production is lowered to ~25% after 24 hours, and dropping further to ~10% after 48 hours. The protein expression in this site is higher under a T7 promoter as is reflected in Figure 3.22 and Figure 3.23, where overexpression of several A-site proteins is seen in strains ED6563B (T7-T7), RS01 (T7-A2), RS02 (T7-M2) and RS-6_2 (T7-A2). Higher protein expression in the A-site of these strains is also reflected by the β-galactosidase assays, indicating higher amounts of mRNA being produced. As with the cobalamin yield, the β-galactosidase activity after 24 hours drops with 75% when the A-site activity controlled under a constitutive promoter when compared to T7-expression. After 48 hours however, this difference is levelled out, indicating the T7 promoter was shut off between these time-points. That the difference in B_{12} levels still increases in this timeframe indicates the proteins of the pathway are still present and active whilst the β-galactosidase is not. Whilst this could reflect a difference in half-life between the proteins, it is more likely explained by degradation of the relevant mRNA. As the *lacZ* gene resides at the very end of the operon, it stands to reason this section of the mRNA would be degraded before the section which holds the B_{12} pathway transcript. This question could be further explored by RT-qPCR experiments, allowing for insight into the relative mRNA levels of the different strains at various time points. Additionally, this method would allow for insight into the stability of the transcripts when designed with primers at the beginning, middle as well as the end of the operon. This could shed light on potential additional bottlenecks within the pathway. With this in mind, mass spectrometry could complement this data.

The B-site expression is reflected in the expression of the cobaltochelatase from *Bruscella melitensis*, which is visualised on western blot. The trend shown in Figure 3.25 andFigure 3.26 does not reflect the difference in promoter strength as is seen in the A-site. The most pronounced bands are seen in stains ED656-3B (T7-T7), RS01 (T7-A2) and RS06_2 (T7-A3). Whilst ED656-3B does contain a T7 promoter in its B-site, the others do not. The results reflect a correlation between the T7 promoter in the A-site and the B-site expression more so than the promoter strength of the B-site. It may be that a feedback system exists, allowing enhancement of the expression of the B-site cobaltochelatase when high A-site expression has resulted in intermediate build-up.

One aim of the approach of weaker promoters was increased cell health as result of reduced metabolic strain. Whilst a change is seen in reduced protein expression - indicating less metabolic strain - as a result of weaker, constitutive promoters in the A-site, only a slight positive effect on phenotype was observed. The cell density of strains containing constitutive promoters in the A-site was generally lower than that observed with the T7 promoter. It is very likely that the strain which is put on the cell's metabolism from the point of inoculation as a result of constitutive promoters outweighs any positive impact the promoter strength might when compared to the T7 system.

The method used to quantify the cobalamin yield of the engineered strains was determined by use of *Salmonella* bioassay plates. This method relies on a specific *Salmonella* strain's inability to produce cobalamin due to *cysG* or *cbiB* mutations and is therefore dependent on the uptake of cobalamin or its intermediates from cobyric acid or cobinamide - depending on the mutant strain – to cobalamin (Raux et al., 1996). Although quick and useful, the method leaves room for some human error during preparation as well as analysis. Live *Salmonella* is added to the agar before being poured out to the plates. Whilst a control plate containing standard solutions is applied to account for daily changes in the plates, related to age and state of dehydration of the plates, there may well be discrepancies between each individual plates caused by difference in agar volume. This would affect *Salmonella* concentration, differences in oxygen exposure, all able to cause variances in results. Additionally, the manual measuring of the growth circle diameters allows room for human error. A recent paper by Young et al. describes an automated analysis of growth areas, circumventing this issue (Young et al., 2021). Alternatively, liquid culture assays have been described which are measured via optical density (Bunbury et al., 2022). The agar plate method therefore provides only an estimate of the B_{12} concentrations, but for the purpose of comparative yields between the various strains, the method suffices.

Figure 3.24 shows the same trend in expression level differences between the strains at 24 hours. A higher expression is seen in strains ED656-3B (T7-T7), RS01 (T7-A2), RS02 (T7-M2), and RS06_2 (T7-A3) when compared to the other strains at this time point. After 48 hours, however this no significant difference is observed anymore. This could be explained by the T7 promoter system having turned off and activity reducing to that of remaining enzymes. It must be noted that the Miller units measured for all samples are quite low when compared to the range generally described in the literature for a standard β-galactosidase assay. In his book, 'Experiments in Molecular Genetics', Dr. Miller explains that 1 Miller Unit represents uninduced *E. coli* (low β-Gal production) and that a fully induced culture will be represented by approximately 1000 units (Miller, 1972). This would put the activity of the strains in the range of uninduced cultures. Since the data shows a clear trend supporting the differences between the T7-controlled expression levels and the constitutive ones, these results cannot be disregarded. Literature often describes β-galactosidase assay of plasmid-located expression systems, one contributing factor of lower values can be the location of the lacZ gene in the genomic DNA. A high-copy plasmid will allow for much higher expression that that of a single gene on the genomic DNA. A high-copy number expression vector would generally not exceed 20-50. Factoring this in, the data would be more in range of the Miller Units expected. Another contributing factor of the low values could be the age of the samples. As multiple analyses were performed on samples from the same cultures, time prevented them to be done simultaneously. Freezing of the samples could have had an effect on the activity measured.

The data presented shows higher expression levels in the A-site of ED656-3B, RS01, RS02 and RS06_2 when compared to the others, a result in line with the promoter strengths described (Alper et al., 2005; *Promoters/Catalog/Anderson - Parts.Igem.Org*, n.d.). A distinct difference between the various constitutive promoter strengths is not represented by the data. The B-site expression was not quantified to any high degree of accuracy as the western blots were not clear enough for densitometry analysis. There seems to be some difference in expression levels between ED656-3B, RS01, RS06_2 and the rest – the three mentioned strains showing a slightly thicker band on these blots. This might indicate higher expression levels of the B-site in comparison with the other strains. This would be surprising as strains RS01 and RS06_2 contain a weaker, constitutive promoter in this site, whereas the B-site operon of RS03 and RS05 are controlled by the stronger T7 promoter. This leads to the expectation that the differences as seen in Figure 3.25 and Figure 3.26 are not significant. True promoter strength is measured not only by expression levels, but also by investigating the relationship between mRNA synthesis and protein expression. Additionally, there is the stability – and therefore turnover – of transcripts and proteins to

be considered. The conclusions drawn in this chapter are based on a simplified system which assumes all mRNA products of the same operon are equally stable. The machinery of expression then is equal in all strains, regardless of the promoter which controls translation. Expression should therefore represent translational differences between the strains' operons. All mRNA products of the same operon being equal – and equally stable – is a big assumption to make, however. and would have done with backing up via qPCR data. Primers were designed in preparation to investigate this experimentally, however, lack of time prevented the execution.

Chapter 4

Identifying alternative targets for improving vitamin B_{12} production in an *E. coli* strain.

4.1 Introduction

Rhodobacter capsulatus has the ability to produce cobalamin when grown in aerobic as well as anaerobic conditions, although the genes of the biosynthetic operon bear more homology to the aerobic pathway (McGoldrick et al., 2002). Strain ED656-3B is a modified *E. coli* where the B_{12} biosynthetic genes are integrated into its genome. This strain is able to produce a respectable 40 mg/L of B_{12}, however, to make the strain commercially viable production levels closer to 1 g/L need to be approached.

Looking at the cobalamin biosynthetic pathway, there are a number of steps that present possible targets for cobalamin production improvement of strain ED656-3B. One approach would be to help the provision of specific cellular resources that are explicit to the B_{12} pathway. Increased levels of substrates could push the reaction equilibria towards products and so increase the final B_{12} yield. The first committed precursor in the tetrapyrrole pathway is 5-aminolevulinic acid (ALA), which can be produced via one of two ways; the C4 pathway, and the C5 pathway. In the C4 pathway, also known as the Shemin pathway, ALA is produced by the enzymatic condensation of glycine with succinyl-CoA (Shemin et al., 1956). This route is mainly utilised by animals, fungi and α-proteobacteria (Leeper et al., 2012) and is catalysed by ALA synthase (ALAS), also denoted as HemA from its involvement in heme synthesis. Plants and many bacteria - including *E. coli* - instead make use of the C5 pathway which biosynthesizes ALA from glutamic acid (Beale, 1990; O'Neill et al., 1989). This conversion requires 3 steps which are catalysed by glutamyl-tRNA synthetase (GluRS), glutamyl-tRNA reductase (GluTR) and glutamate-1-semialdehyde-2,1-aminomutase (GSAM). Confusingly, GluTR is also referred to as HemA. For this reason, the HemAs are here referred to according to the pathway in which they belong; $HemA^{C4}$ and $HemA^{C5}$ (Figure 1.5.2). *E. coli* produces 5-ALA via the C5 pathway, and as it is deemed a rate limiting step for tetrapyrrole biosynthesis (Lascelles, 1978), several attempts at increasing ALA levels have been made by overexpression of these genes. As there are three steps to this pathway, including a dependence on cofactor $tRNA^{Glu}$, the C4 ALAS from *R. sphaeroides* has been introduced by way of metabolic engineering for increased ALA production in *E. coli* resulting in a 5-fold increase in yield (Werf & Zeikus, 1996).

The first committed step to direct intermediates along the cobalamin pathway is the conversion of uroporphyrinogen III to precorrin-2. This step separates the cobalamin pathway from the biosynthetic pathways leading to (bacterio-)chlorophyll and heme, and is facilitated by the enzyme called S-Adenosyl- L-methionine-dependent

uroporphyrinogen-III C- methyltransferase (SUMT). This enzyme, also denoted as CobA, catalyses two methylations of the tetrapyrrole ring at positions C2 and C7 (Sattler et al., 1995; Vévodová et al., 2004). The CobA enzyme was first isolated from *Pseudomonas denitrificans* and was found to be kinetically inefficient with a turnover number of only 38 h^{-1}. It shows inhibition not only by one of its products, S-adenosyl-homocysteine (SAH), but also by its substrate, uroporphyrinogen III (Crouzet et al., 1990). This high level of inhibition implies a regulatory role for the enzyme, controlling the metabolic fluxes through the pathways that branch off at this point. In *E. coli*, the reaction catalysed by CobA is also the final enzymatic step before the biosynthetic pathway towards cobalamin branches off from that of siroheme. From here, another methyltransferase, CobI, is responsible for adding the methyl group at the C20 position (Thibaut et al., 1990). Improving the activity of these first two committed steps could encourage more flow into the cobalamin pathway – and away from the other pathways for which uroporphyrinogen III is also a precursor.

Another possible target for production improvement is the availability of S-adenosyl-methionine (SAM) as this provides the source for the 8 methyl groups that are added to the tetrapyrrole macrocycle along the cobalamin pathway. SAM is a major metabolic product of the methionine biosynthetic pathway as well as a key component in the regulation of the methionine pathway (Figure 4.1). SAM synthesis can be considered the rate limiting step and is catalysed by S-adenolsyl-L-methionine synthase (SAMS) (Markham & Pajares, 2009), also referred to as methionine adenosyltransferase (MAT), and MetK – derived from the gene which encodes the enzyme: *metK*.

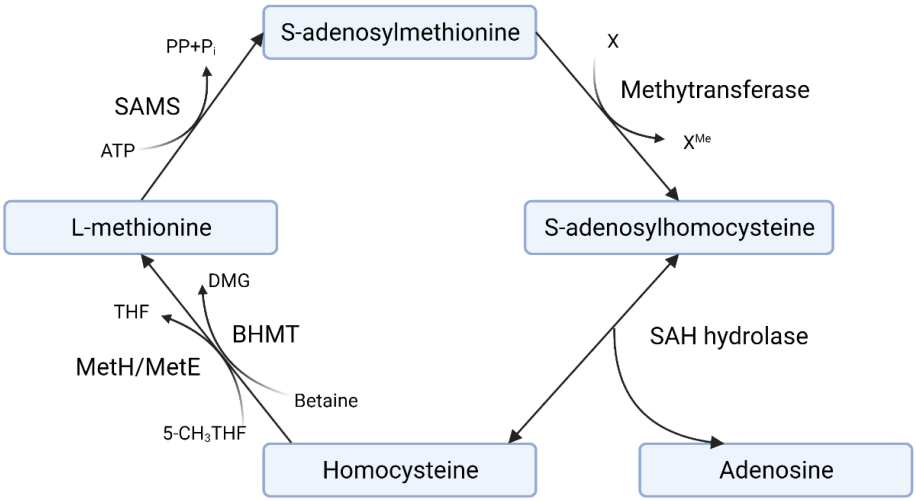


Figure 4.1 Methionine recovery pathway. *The X stands for any substrate which is recognised - and can be methylated - by methyltransferases. Within the cobalamin pathways SAM_derived methyl groups are added to precorrin-n intermediates, where the n refers to the number of methyl groups that have been added to the framework. SAMS is the S-adenosyl methionine synthase enzyme, also denoted as MetK. MetH and MetE are methionine synthases which are cobalamin-dependent and independent, respectively. BHMT is the betaine – homocysteine methyltransferase which catalyses the methylation of homocysteine to methionine using betain instead of the more commonly used CH_3-H_4folate* **(Matthews et al., 2008)**.

Additionally, S-adenosylhomocysteine hydrolase (SAHH) is an enzyme that is also worth looking at when attempting to increase the amount of available SAM. SAHH hydrolyzes S-adenosylhomocysteine (SAH) to homocysteine and adenosine, a reaction that is reversible and favours SAH synthesis. Promoting the hydrolysis of SAH would encourage flow into the methionine recovery pathway, thereby increasing SAM availability. SAH is a known powerful inhibitor of many of the cobalamin biosynthetic methyltransferases.

Additional copies of these target genes on the genome of ED656-3B could result in a beneficial effect on the yield reached by our cobalamin-producing strain. To avoid recombination, however, these genes cannot be sourced from *E. coli*. The genes were therefore selected from alternative organisms.

The work described in this chapter investigates whether overproduction of some of the key enzymes outlined in Figure 4.2 results in any improvement in cobalamin production . This was approached by employing plasmids containing the genes of these enzymes from various sources under the control of a T7 promoter. The effect of (over-)expression of the enzymes on B_{12} production was monitored in order to identify a significant change. This was followed by placing the genes under the control of a more regulated *tet*-operator system. Whilst still a medium-copy number plasmid, this system is much more tightly regulated, as expression levels correspond to the amount of inducing agent tetracyclin. This allows investigation of conditions that more closely resemble genomic expression levels and monitoring of the B_{12} production under these conditions.

4.1.1 Project aim

The work presented in this chapter aims to increase the yield of B_{12}-producing *E. coli* strain ED656-3B by testing several possible targets within – or feeding into – the B_{12} biosynthetic pathway.

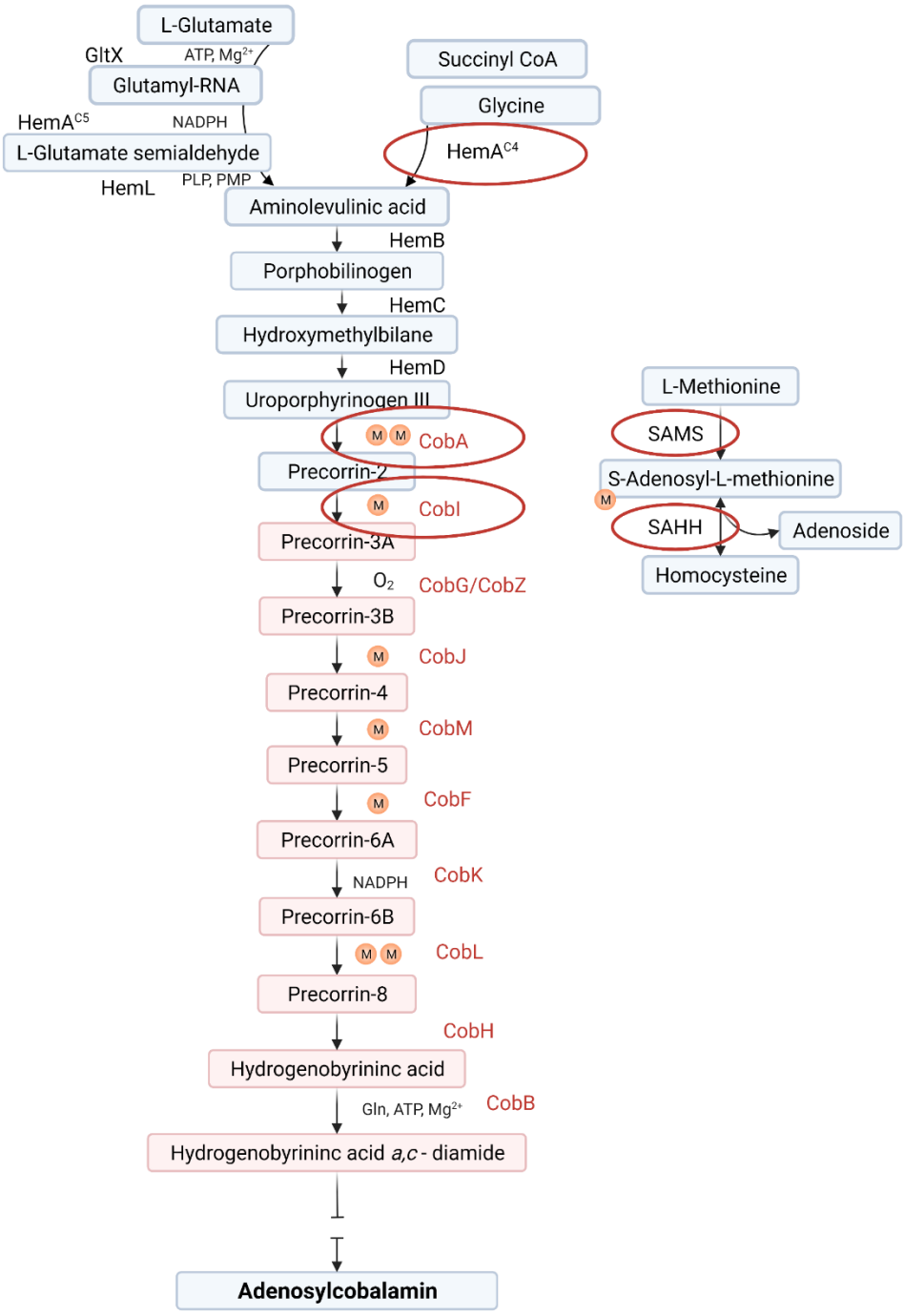


Figure 4.2 Cobalamin biosynthetic pathway including production improvement targets (circled red).

4.2 Introducing target genes to ED656-3B

To investigate the proposed enzymes mentioned in Section 4.1 as possible production improvement targets, genes for these enzymes were sourced from various genomic DNAs available in the lab.

The choice of genes was subject to availability as well as rational design. All genes selected were available on vectors in the lab. The HemA which is used by strain ED656-3B is its inherent $HemA^{C5}$ enzyme, facilitating ALA synthesis via the C5 pathway. To increase ALA levels, $HemA^{C4}$ was selected from *Rhodobacter capsulatus* with the aim of adding this to ED656-3B, thereby enabling the strain to synthetize ALA via both pathways. The *cobA* alternatives that were selected for the cobalamin-producing strain originated from *Methanosarcina barkeri (Mbar)* as well as *Pseudomonas denitrificans (Pd)*. In protein expression studies of the cobalamin pathway, it was found that expression of CobA from *Rhodobacter capsulatus* was extremely low (Delobel, 2008). *Mbar* CobA has been reported to be highly overexpressed (Schroeder et al., 2009) and was therefore included in these studies. *Mbar* CobI was selected for the same reason. As previously mentioned, *Pd* CobA is subject to substrate as well as product inhibition. The *Pd* CobA used in this study was mutated at position 187 to replace a Lysine for an Alanine by a previous researcher in the Warren lab. This mutation cured the enzyme of all substrate inhibition (based on unpublished work done in the Warren lab), and *Pd* $CobA^{K187A}$ was therefore included in these production improvement studies. The *cobI* gene that was selected for the strain stemmed from *Brucella melitensis.* Furthermore, to improve the B_{12} pathway by providing higher concentrations of key components feeding into the pathway, human as well as yeast SAM synthetases were selected, as well as SAH hydrolase from rat.

It was initially decided to study the effect of overexpression of these genes from plasmids in ED656-3B. If a positive effect was observed, the selected gene would then be subjected to CRISPR-Cas9 cloning to integrate it into the genome. ED656-3B was therefore transformed with the plasmids containing the genes of interest (Table 4.1) as described in Section 2.2.6. Vectors pET3a and pET14b were selected for by ampicillin resistance, whereas pET28 was selected on the basis of kanamycin resistance. With plasmids pET28-*h*SAMS, pET28-*rat*SAHH, pET-*rat*SAHH - *h*SAMS and pET14b-*y*SAMS, the transformation of ED656-3B was successful, however, with pET14b-*Rc hemA* and pET14b – *Pd* $cobA^{K187A}$, the incubation temperature needed to be lowered to 28°C to yield colonies containing the plasmid. Colonies of ED656-3B containing the pET3a -*Mbar cobA* were not obtained. This could be explained by the high expression level of the enzyme encoded by

the gene, even when the presence of 0.2% glucose lowers the basal transcription of the *lac* promoter controlling the T7RNA polymerase, as indicated by previous gene expression profiles done in the Warren lab. Such a high level of CobA proved detrimental to cell health as is evident by the lack of growth.

Table 4.1. List of plasmids to be introduced to ED656-3B containing genes of interest for potential B_{12} production improvement.

Plasmid	Gene(s) of interest
pET3a – *Bmei cobI*	*cobI* from *Brucella melitensis*
pET3a – *Rc HemA*	$hemA^{C4}$ from *Rhodobacter capsulatus*
pET14b – *Pd* $cobA^{K187A}$	$cobA^{K187A}$ from *Pseudomonas denitrificans*
pET14b – *ySAMS*	SAM synthetase from yeast
pET28 – *rat SAHH*	SAH hydrolase from rat
pET28 – *hSAMS*	Human SAM synthetase, originally denoted as MAT2A
pET28 – *rat SAHH - hSAMS*	SAH hydrolase from rat and human SAM synthetase

4.2.1 Cell growth

The growth of ED656-3B, containing the plasmids listed in Table 4.1 – as well as empty vectors – was monitored by measuring the optical density spectrophotometrically at 600 nm over the course of 30 hours. The cultures were grown in triplicate on M9-YE-G media in flasks, which were shaken to enhance oxygenation. Incubation temperatures were set to 28°C to accommodate the *R. capsulatus* genes of the B_{12} biosynthetic pathway. After 6 hours, expression from the T7 promoters was induced by adding IPTG. The results, outlined in Figure 4.3, show that most of the expressed proteins have no significant difference on cell growth. The exception to this observation was with pET14b-*Pd cobA*^{K187A}, indicating again that the overexpression of CobA can be toxic for the cells. This is similar to what was seen with the expression of the *Mbar* CobA, which also appears to be toxic to the cells as expression levels are much higher for the latter enzyme. Expression of Rc HemA and rat SAHH seem to cause a longer lag phase, but the OD values are caught up after 12 and 24 hours, respectively.

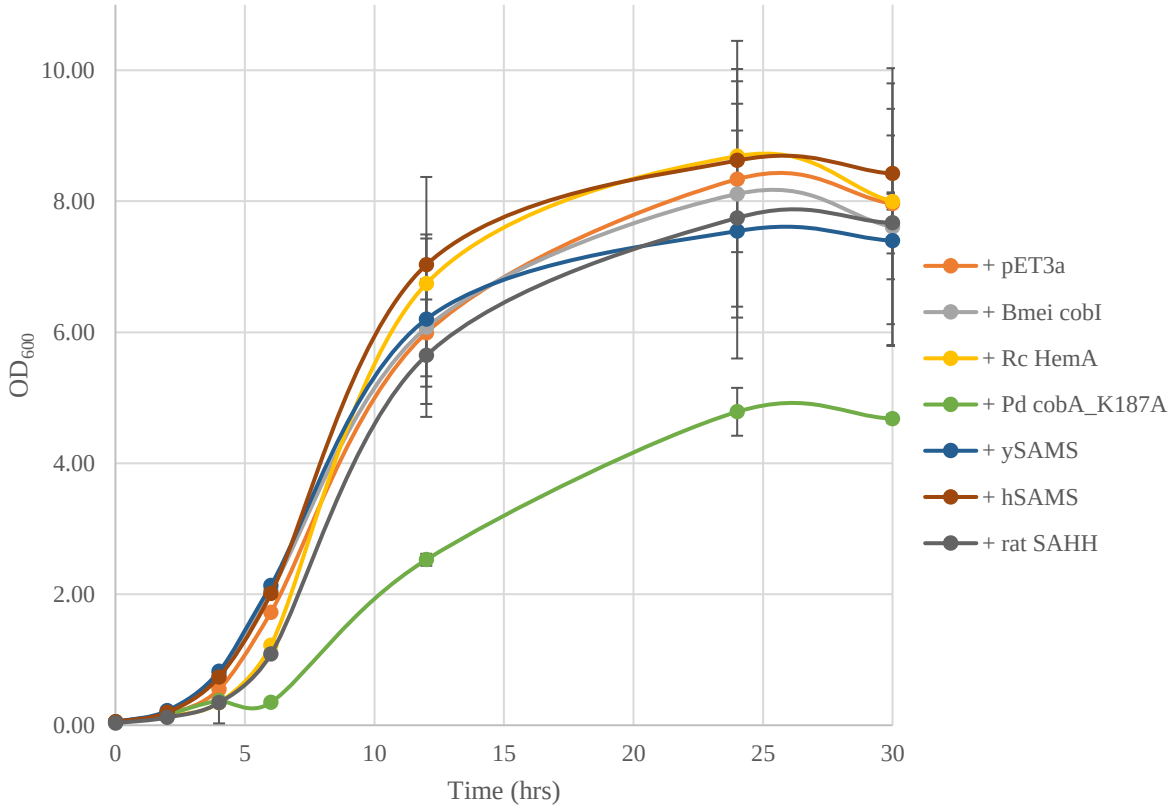


Figure 4.3 Growth of ED656-3B with plasmids carrying genes of interest. *Cultures were grown in triplicate on M9-YE-G media. Empty vectors were included as controls. Optical density was measured spectrophotometrically at 600 nm over the course of 30 hours from the moment of inoculation. Error bars indicate the SD of the values from the biological replicates. All cultures display similar growth curves except for those carrying pET14b and pET14b-Pd cobA^{K187A}, which show significantly lower cell growth from 12 hours on.*

4.2.2 Cobalamin production

The cobalamin yield of the cultures was determined by way of bioassay plates as previously described (Section 3.4.2). The cultures were harvested and supernatant was separated from the cells. The cell pellets were resuspended in buffer and lysed by boiling. Supernatant samples were applied directly onto *Salmonella* bioassay plates which were cultured at 37 °C overnight. Cell-extracted samples were diluted to fall within the range of known standard concentrations of B_{12} that had been applied to the plate. The resulting diameters of the growth circles were measured and values were converted to concentration using a standard curve derived from the known concentrations of B_{12}. The resulting cobalamin concentrations were OD corrected as these values differed significantly for some of the cultures (Figure 4.3). From the results reported in Figure 4.4, it is evident that any additional gene expression at this level has a negative impact on the B_{12} production of the strain. ED656-3B with either *hSAMS*, *Bmei cobI*, *Rc HemA* or *rSAHH* has reduced B_{12} production that represents only around 20% of that observed with the parent strain. Expression of ySAMS and *Pd* $CobA^{K187A}$ resulted in a decrease of over 40 fold. Interestingly, the strain which contained pET14b-*Pd* $cobA^{K187A}$ produced some cobalamin up until the point of induction, but at this stage B_{12} production is quickly stopped altogether, which coincides with a restoration of cell growth (Figure 4.3). This points to the cell 'saving itself' by reducing the high metabolic stress associated with induction of a T7 promoter. This is likely caused by mutations that stop the B_{12} biosynthetic pathway. Expression of *Pd* $cobA^{K187A}$ is shown to increase this effect.

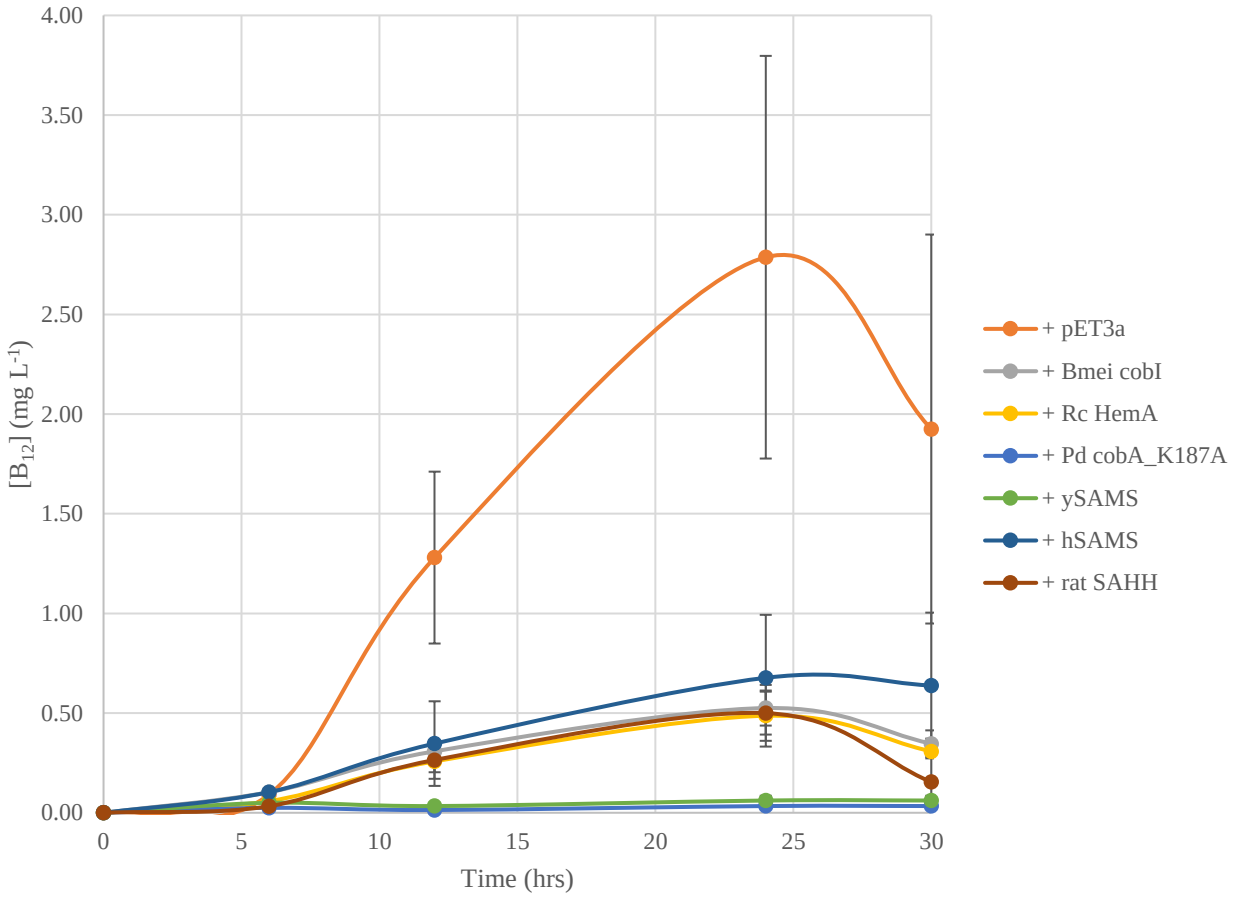


Figure 4.4 Cobalamin production of ED656-3B cultures expressing additional genes in the B_{12} biosynthetic pathway via multi copy vectors, *cultured on YE-M9-G media. Yield measured in triplicate over the course of 30 hours from the point of inoculation, as calculated via Salmonella bioassay plates. [B_{12}] values were OD-corrected to account for the significant differences in cell growth between the cultures. Error bars indicate the SD of the values from the biological replicates. ED656-3B with and without an empty pET3a vector show the highest B_{12} levels, with most other cultures producing less than half these amounts. Complete lack of B_{12} production is seen for cultures with expression of Pd CobA^{K187A} and yeast SAM synthetase.*

4.2.3 Protein expression profile

Investigating the correlation of the reduced B_{12} yield with the overexpression of the additional genes was done via SDS PAGE. Culture samples were taken at 12 and 24 hours after inducing the protein expression under the T7 promoter. The expected protein sizes encoded by the added genes are listed in Table 4.2 In the resulting SDS PAGE gel shown in Figure 4.5 we see a clear overexpression of *Rc* HemA (45.7 kDa), *Pd* CobA^{K187A} (29.2 kDa) and *h*SAMS (46.2 kDa). This could account for the reduced B_{12} yield, as the energy cost of overexpression of a single gene under a T7 promoter on a multi-copy plasmid is likely to take away from the biosynthetic pathway itself. This does not seem the obvious reason for *Bmei* CobI (26.4 kDa), *y*SAMS (44.5 kDa) or *r*SAHH (51.4 kDa), as these are not visibly overexpressed. The drop in cobalamin production in these cases could be attributed to the enzymes' activities. Should these enzymes be low in number, but high in activity, the metabolic cost could similarly be diverted away from the cobalamin biosynthetic pathway.

Table 4.2 Expected protein sizes of target genes

Protein	Size (kDa)
Bmei CobI	26.4
Rc HemA	45.7
Pd CobA^{K187A}	29.2
Yeast SAMS	44.5
Human SAMS	46.2
Rat SAHH	54.1

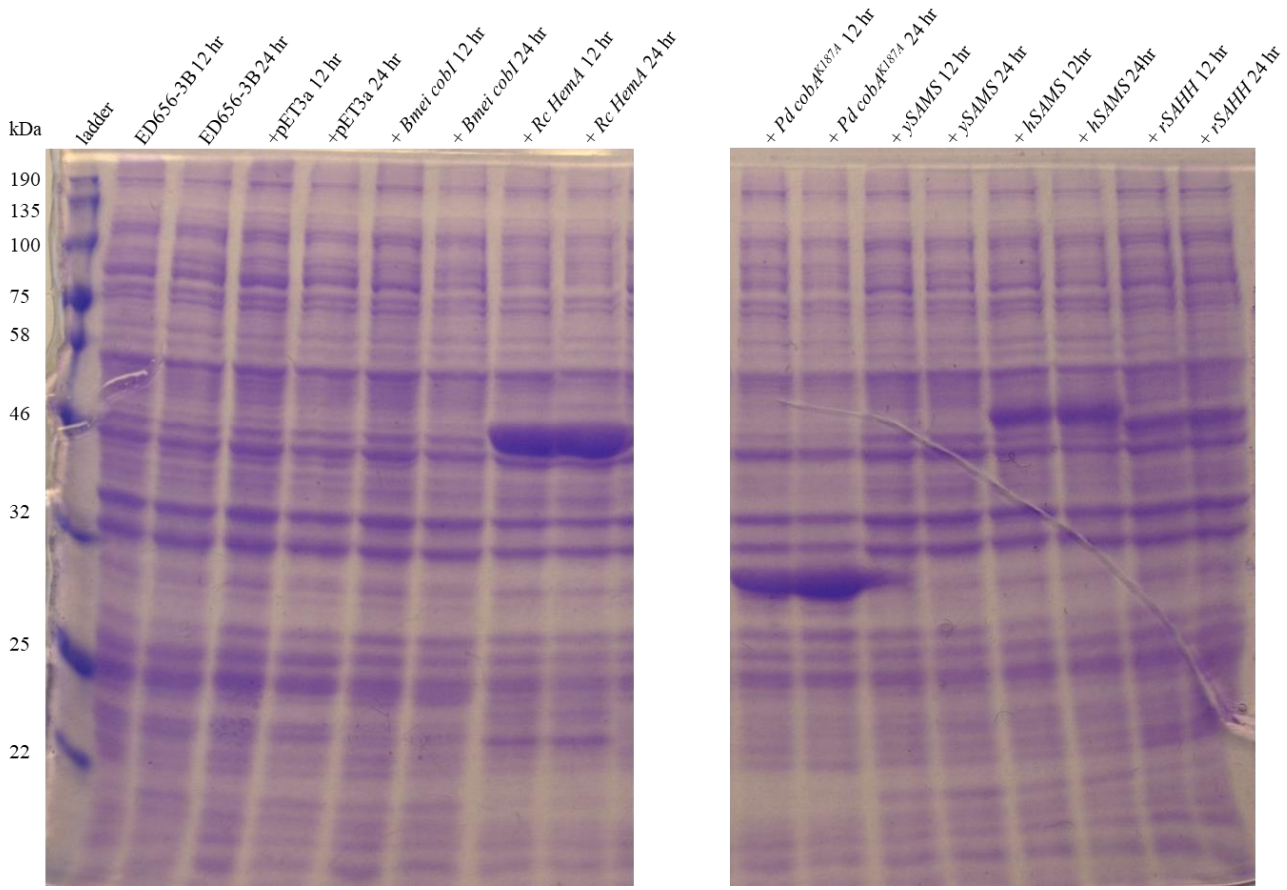


Figure 4.5 Analysis of ED656-3B culture samples containing various genes on multicopy plasmids. *Protein expression under the T7 promoter was induced after 6 hours. Samples were taken 12 and 24 hours after this point, boiled for 10'in SDS PAGE buffer and loaded on a 12.5% SDS PAGE gel. Overexpression of Rc HemA is visible at 12 and 24 hours as distinct bands at the expected size of 45.7 kDa. Pd cobAK187A is visibly overexpressed with bands at the expected size 29.2 kDa after 12 and 24 hours, as well as the human SAM synthetase; hMAT2A, which is seen at the expected size of 46.2 kDa.*

4.3 Effect of expression of target genes under a *TetR-PtetA* promoter system on B_{12} production by ED656-3B.

The overexpression of any gene under the control of the T7 promoter from a plasmid appears to have an adverse effect on B_{12} biosynthesis in ED656-3B. As previously mentioned it is likely caused by the metabolic cost that takes away from the cobalamin biosynthetic pathway. Tighter regulation of the expression levels could remedy this, and give a better representation of the effects of the selected genes on B_{12} production if these genes were expressed from the genome. As the T7 promoter system does not allow for tight regulation, attention was turned to an alternative promoter system; that of the tetracycline inducible Tet repressor-operator system. Stemming from the resistance mechanism in gram-negative bacteria, the tetracycline repressor (TetR) controls the expression of tetracycline-exporter protein, TetA. TetR achieves this by binding to the Tet operators *tetO1* and *tetO2*, thereby repressing the promoter system P_{tet}. When tetracycline (Tc, bound to Mg; $[MgTc]^+$) is present, it binds to TetR, causing a conformational change that results in the dissociation of TetR from *tetO*, enabling transcription of *tetA* to occur (Evans & Mizrahi, 2015). Facilitating nonspecific cation export, high levels of TetA expression can cause a collapse in membrane potential. The feedback system is therefore very sensitive. Tetracycline's bacteriostatic properties arise from its binding to ribosomes, thereby repressing protein expression (Chopra & Roberts, 2001). The binding constant (k_{on}/k_{off}) of TetR to *tetO* is 10^{11} M^{-1}, whereas for Tc, the binding constant to ribosomes is $K_a \sim 10^6$ M^{-1}. Binding of $[MgTc]^+$ to TetR ($K_a \sim 10^9$ M^{-1}), lowers the association constant of TetR to *tetO* by 9 orders of magnitude (Orth et al., 2000).

4.3.1 Cloning genes under TetR-PtetA

Cloning the aforementioned genes under the control of this TetR-tetO system allows for a tight regulation of the expression levels reflecting more closely the levels of expression that would be observed by genome-encoded enzyme production. Additionally, the tight regulation of expression levels could overcome some of the transformation issues associated with the *Mbar cobA* compatibility with ED656-3B. Cloning of the relevant genes into a TetR-controlled plasmid was achieved by using restriction enzymes to extract the genes of interests and then re-ligating them in pET-tetR-PtetA; a plasmid containing the *tet* promoter as well as the *tetR* repressor gene. This section (promoter and repressor gene) was flanked by BglII and SpeI, which was used to digest the operator system before ligating them into the plasmids containing the genes of interest. In the case of *Mbar cobA*, the digested fragment was ligated into the pET3a-*Mbar cobA* vector, digested with BglII and SpeI-complimentary restriction enzyme; XbaI (Figure 4.6). All final plasmids were checked via restriction digests and subsequent sequencing, and transformation of ED656-3B was achieved. With gene expression tightly repressed, all transformants yielded colonies.

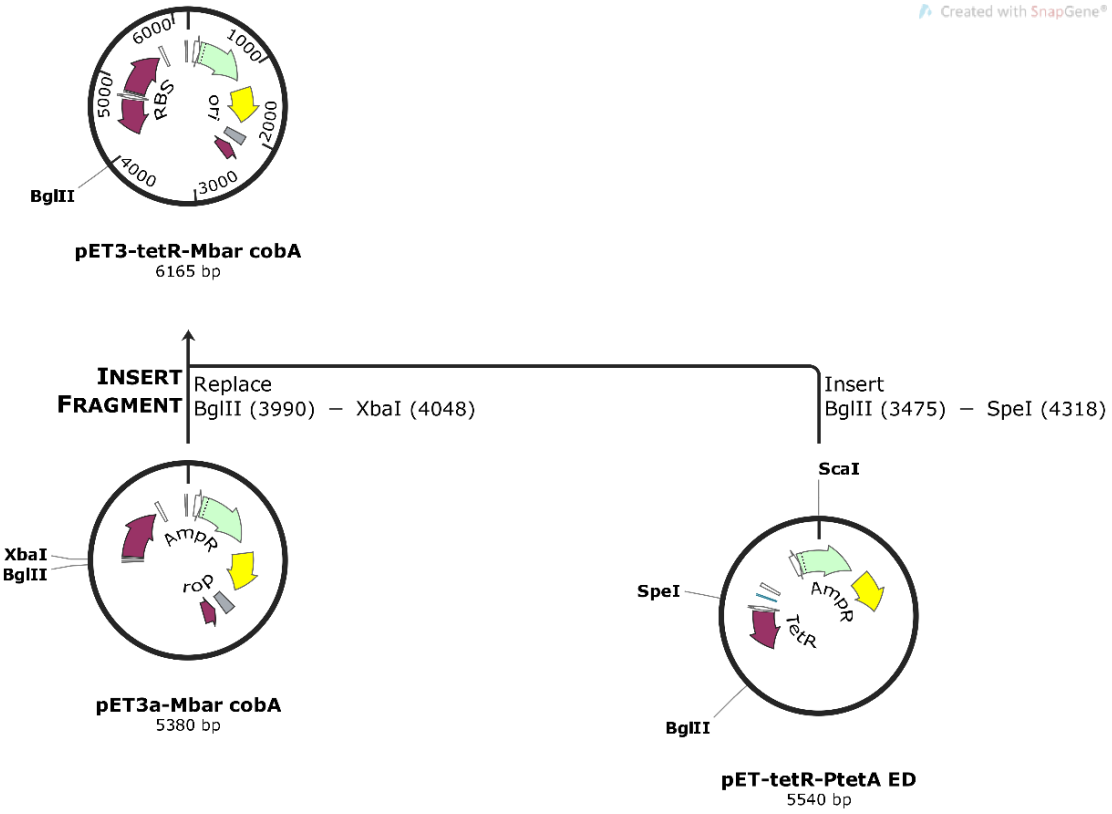


Figure 4.6 Cloning schematic of Mbar cobA to be expressed by the tetR-Ptet promoter system. *The tetR-Ptet promoter system was extracted from pET-tetR-PtetA by digesting with restriction enzymes BglII and SpeI. Complementary restriction sites of XbaI and SpeI allowed for ligation of the digested gene into pET3a – Mbar cobA that was digested with BglII and XbaI.*

4.3.2 Cell growth

Cobalamin production of ED656-3B expressing additional target genes was tested as described previously, where cultures were grown in M9-YE-G at 28˚C in duplicate for 6 hours at which point expression of the B_{12}-biosynthetic pathway was induced by adding IPTG and the expression of genes under control of the *tet*-promoter was induced by adding 2.5 µg/L tetracycline. The cell growth was measured at 6 and 24 hours after inoculation and as Figure 4.7 (as well as statistical analysis) shows, there were no major differences between any of the cultures. The change of the T7 promoter to the more tightly regulated tetracycline promoter permitted the successful transformation of ED656-3B with all genes, including *Mbar cobA*.

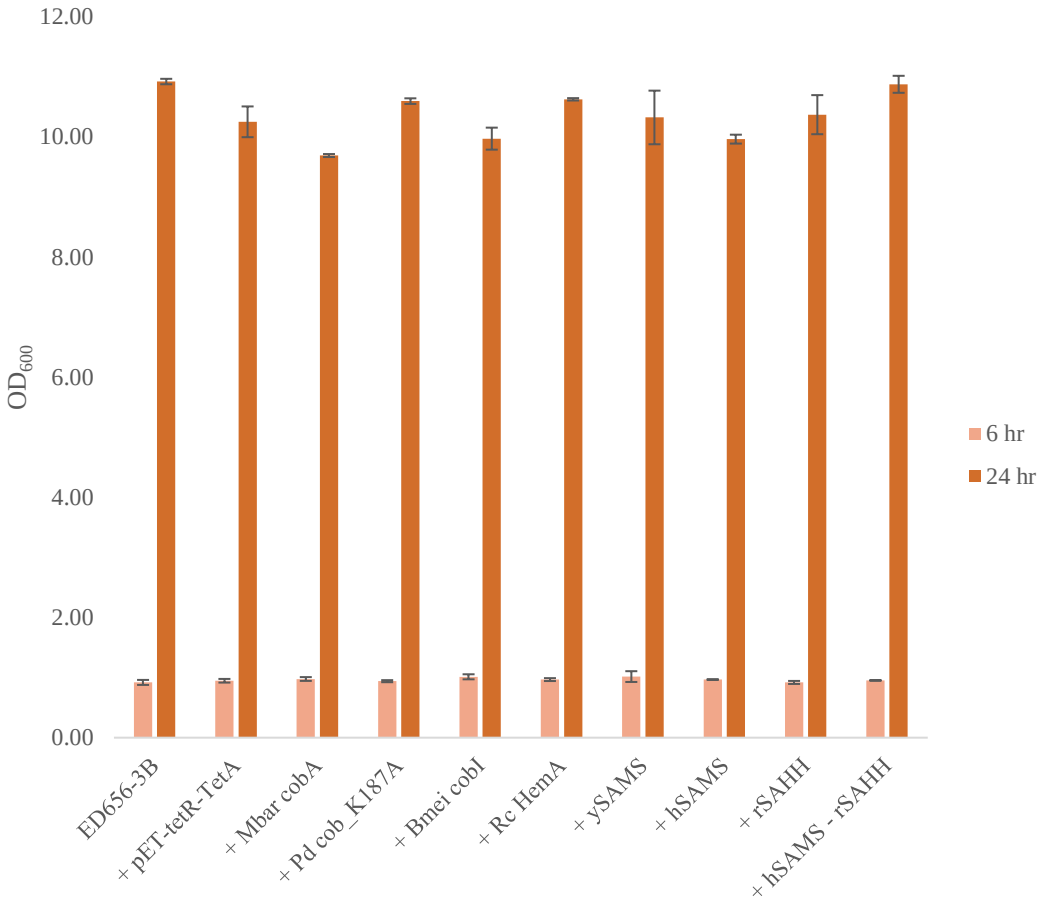


Figure 4.7 Cell density of cobalamin-producing cultures expressing additional target genes, 6 and 12 h after inoculation. *Strains were grown in duplicate on YE-M9-G media, T7 and tet promoters were induced 6 h after inoculation with 0.4mM IPTG and 2.5 mg/L tetracycline, respectively. ED656-3B with and without an empty pET-tetR-PtetA vector were included as control samples. Error bars indicate the SD of the values from the replicates.*

4.3.3 Cobalamin production

The cobalamin yield was tested 24 hours post inoculation of the cultures. This was done via the *Salmonella metE*$^{-}$ *cysG*$^{-}$ bioassay plates as previously described (Section 3.4.2). Figure 4.8 shows the cobalamin yield per strain. To test the differences, an analysis of variance was conducted, with [cobalamin] (uM) as a dependent variable and strain as the 'between subject' - variable. A significant main effect of strain was found; $F = 109.038$, $p < 0.001$, which warranted for further inspection of post-hoc comparisons. Post-hoc comparisons (with Bonferroni correction) showed that introduction of *M. barkeri cobA*, *R. capsulatus hemA*, *hSAMS* or *rSAHH* to ED656-3B increased the yield significantly ($p < 0.001$). The introduction of *P. denitrificans cobA*^{K187A} resulted in a significant decrease in cobalamin yield ($p < 0.001$), when compared to the 'empty' ED656-3B. This indicated that the additional expression of CobA from *M. barkeri* resulted in a two-fold increase of B_{12} production when compared to the empty strain, whilst the additional expression of CobA^{K187A} from *P. denitrificans* appears to significantly lower the cobalamin yield. Expression of HemA from *Rhodobacter capsulatus* and rat SAH hydrolase resulted in a 30% increase in yield, however, the remaining proteins have not attributed any significant change in cobalamin production.

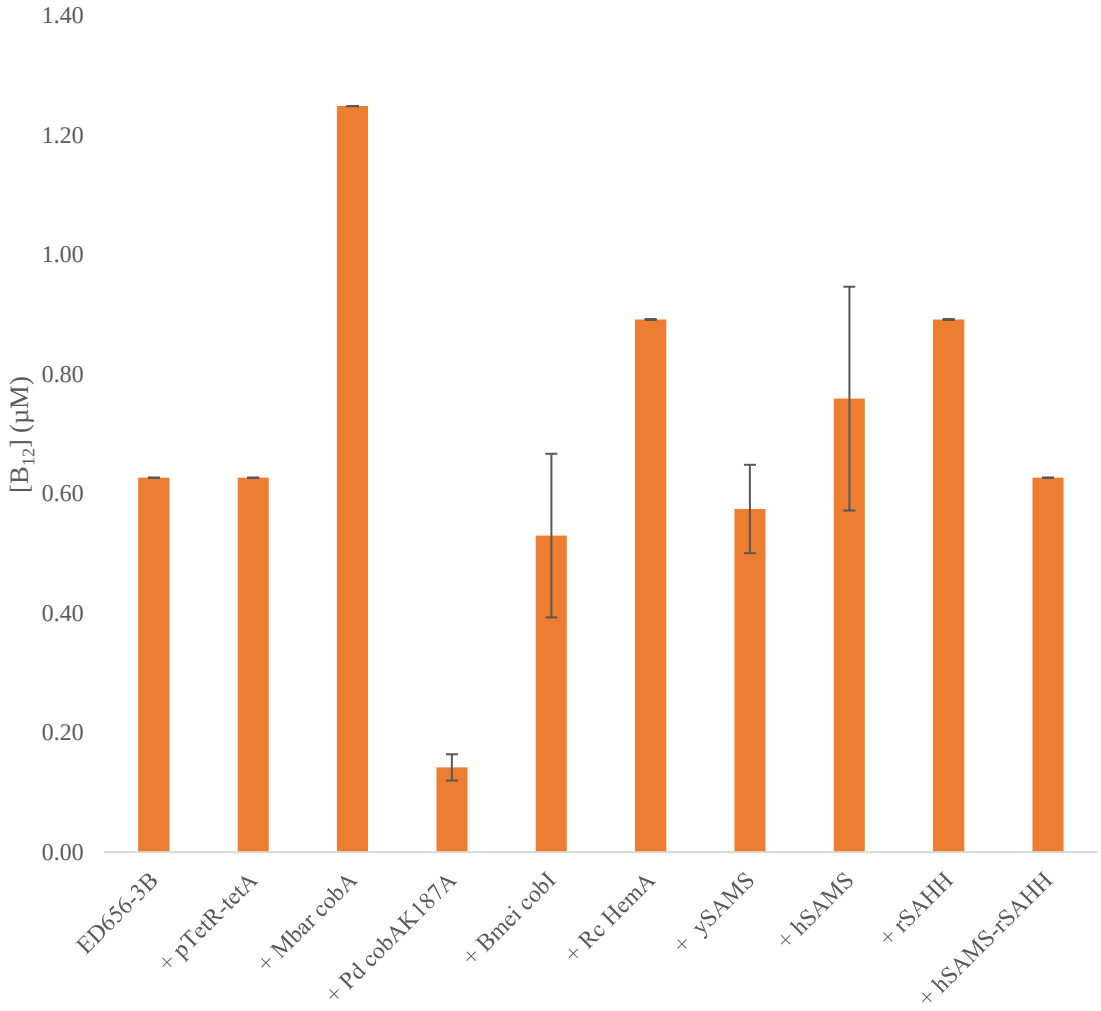


Figure 4.8. Cobalamin production of ED656-3B cultures expressing additional genes in the B_{12} biosynthetic pathway via multicopy vectors, under control of the tetO system, *cultured in duplicate on YE-M9-G media. Yield measured at 24 hours after inoculation, as calculated via Salmonella bioassay plates. Error bars indicate the SD of the values of the biological replicates.*

4.3.4 Gene expression under TetR-PtetA

Protein expression within the culture samples was analysed after 24 hours via SDS PAGE (Figure 4.9). The expected protein sizes are listed in Table 4.3. The results indicate overexpression of yeast SAM synthetase as the lane which contains culture sample ED656-3B + pET-tetR- ySAM synthetase shows a distinct band around 50 kDa. Whilst this is slightly higher than the expected size of yeast SAM synthetase (44.5 kDa), it is around the 10% error margin expected of SDS PAGE gels, it seems likely that the overproduced band represents the protein. None of the remaining proteins that are encoded by the genes on the pET-TetR-PtetA vectors are visibly overexpressed.

Table 4.3 Expected protein sizes of target genes

Protein	Size (kDa)
Mbar CobA	27.2
Pd CobA^{K187A}	29.2
Bmei CobI	26.4
Rc HemA	45.7
Yeast SAMS	44.5
Human SAMS	46.2
Rat SAHH	51.1

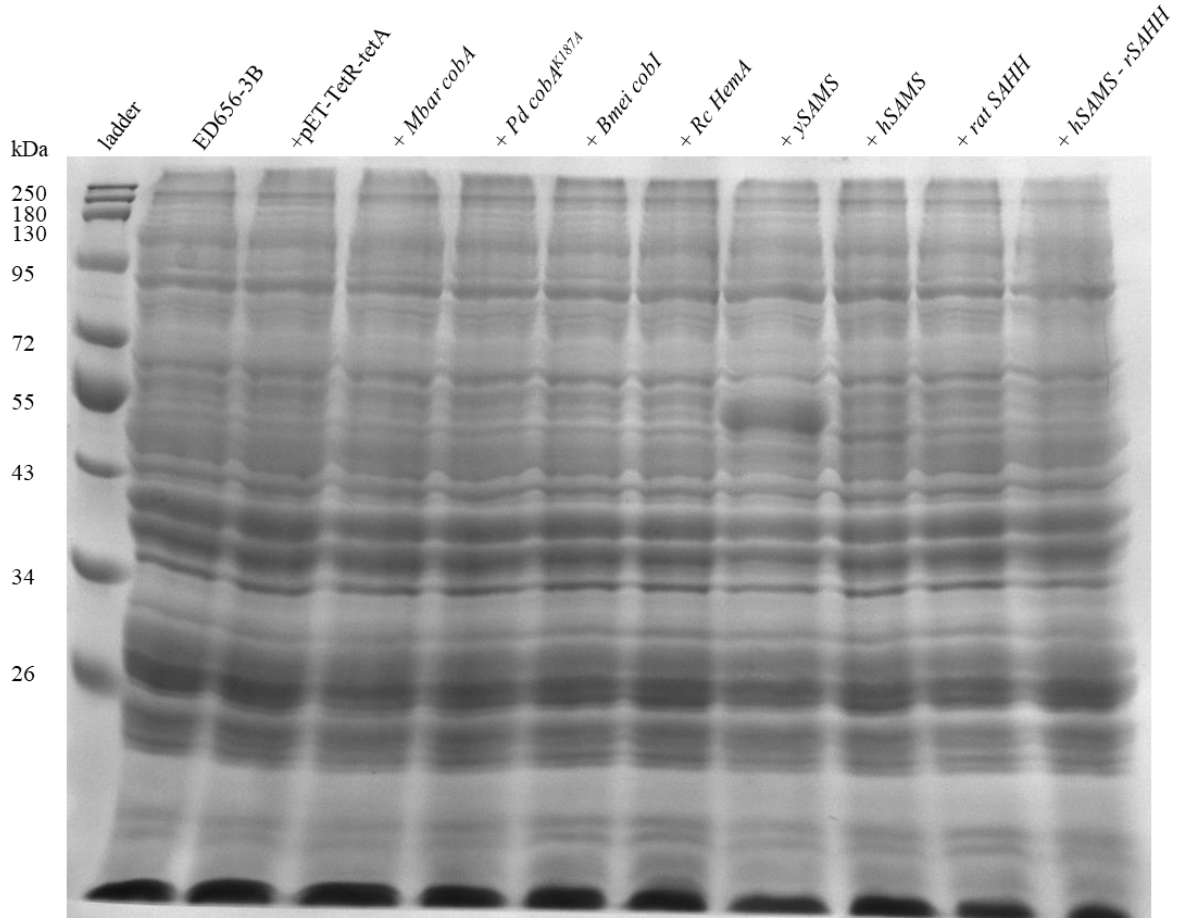


Figure 4.9 SDS PAGE analysis of cultures 24 hours after inoculation. *ED656-3B, expressing the enzymes of the B_{12}– biosynthetic pathway, as well as the additional enzymes encoded by genes provided by vectors (denoted above each lane). Cultures samples were OF corrected and heated to 95 °C in 1:1 SDS sample buffer. Protein bands were visualized with coomassie blue stain.*

4.4 CRISPR – Cas9 cloning

Low level expression of *Mbar* CobA in ED656-3B resulted in a significant increase in cobalamin production by the strain. The natural next step was to see if this increased yield was also achieved when the *Mbar cobA* gene was placed on the genome of ED656-3B. As the difference in *Mbar cobA* versus *Rc cobA* seems to be expression levels (Delobel, 2008), the approach taken here was the replacement of *Rc cobA* in operon A with *Mbar cobA*.

4.4.1 Cloning the target DNA

As previously described in Section 3.3.1, the target DNA is designed by choosing an N20 sequence at the place of the genome where an insert or change is desired. This sequence is then cloned into pTarget-Afe, creating pTarget-N20X. Figure 4.10 shows the target N20 sequence in ED656-3B, which was located in the *Rc cobA* gene in the A-site operon. To synthesize the target sequence to be inserted into pTarget-Afe, forward and reverse primers were designed that contained these N20 sequences. The N20 was flanked by linkers that matched SpeI and AfeI allowing restriction cloning into pTarget Afe as previously described in Section 3.3.1.

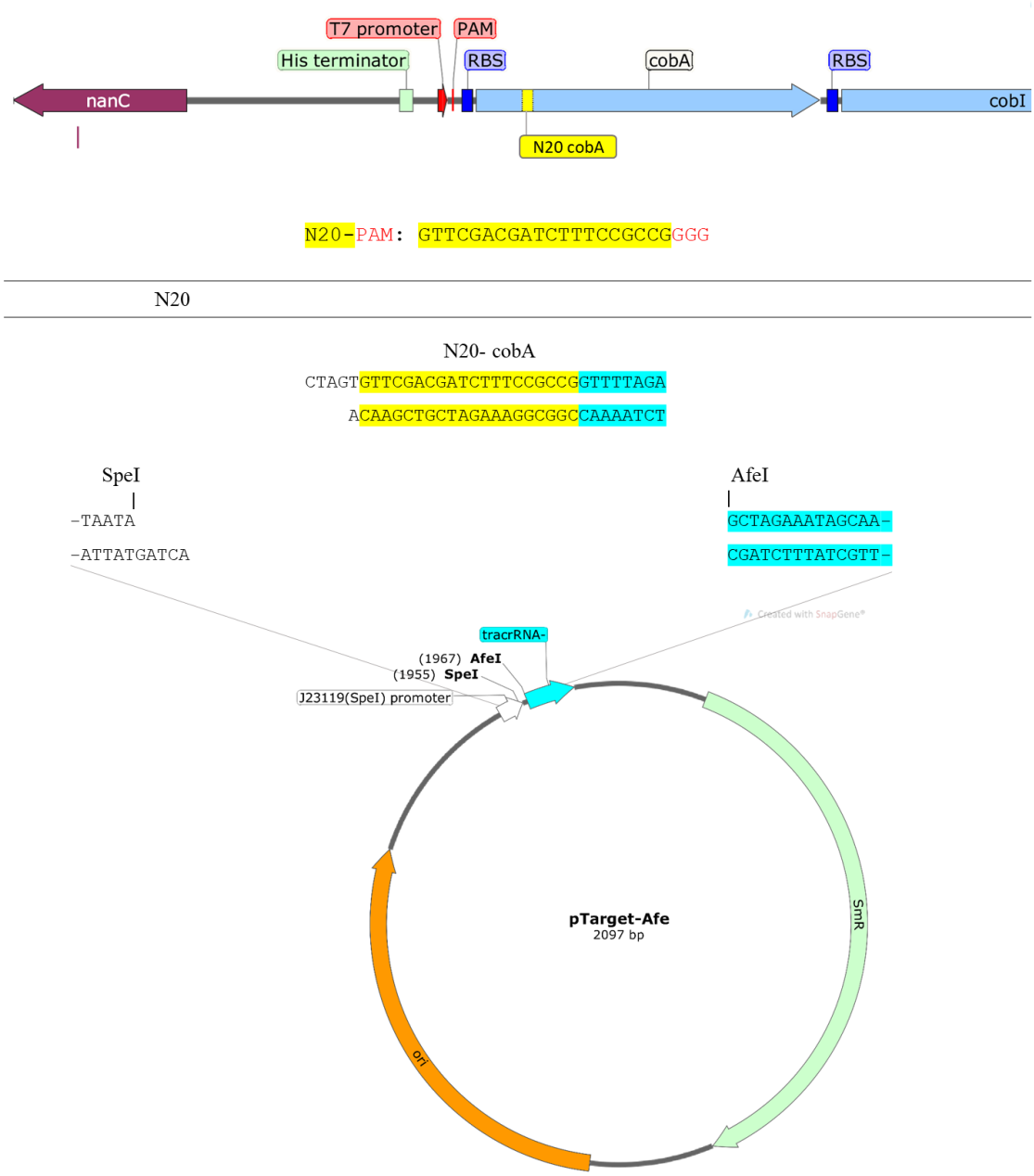


Figure 4.10 Cloning schematic of pTarget-N20 cobA plasmid. *A-site operon of ED656-3B where the N20-PAM is seen inside of the cobA gene. Sequence is shown below. Primers were designed to hold the N20 sequence flanked by linkers. Insert was synthesized by self-annealing forward and reverse N20 primers and cloned into pTarget-Afe by restriction with SpeI and AfeI.*

4.4.2 Cloning the donor DNA

For *Rc cobA* to be replaced by *Mbar cobA*, the latter gene needed to be inserted between 3'and 5' flanking regions, matching the genomic DNA. As described in Section 3.3.2, this allows for homologous recombination repair to take place after CRISPR-Cas9 cleavage. The 5' flanking region was digested from plasmid p*NanC*-A3-*Rc cobA* (Section 3.3.2.1) using restriction enzymes SalI and BglII, and was inserted into pET3a-*Mbar cobA*, upstream from the T7 promoter which controls transcription of *Mbar cobA*. The 3' flanking region was then digested from pET3a-*Rc cobI* using the restriction enzymes XbaI and HindIII and inserted into the SpeI-HindIII site of pNanC-T7-*Mbar cobA*, making use of the complementary restriction sites of XbaI and SpeI. This cloning process resulting in pET-*nanC-Mbar cobA-Rc cobI* is represented schematically in Figure 4.11.

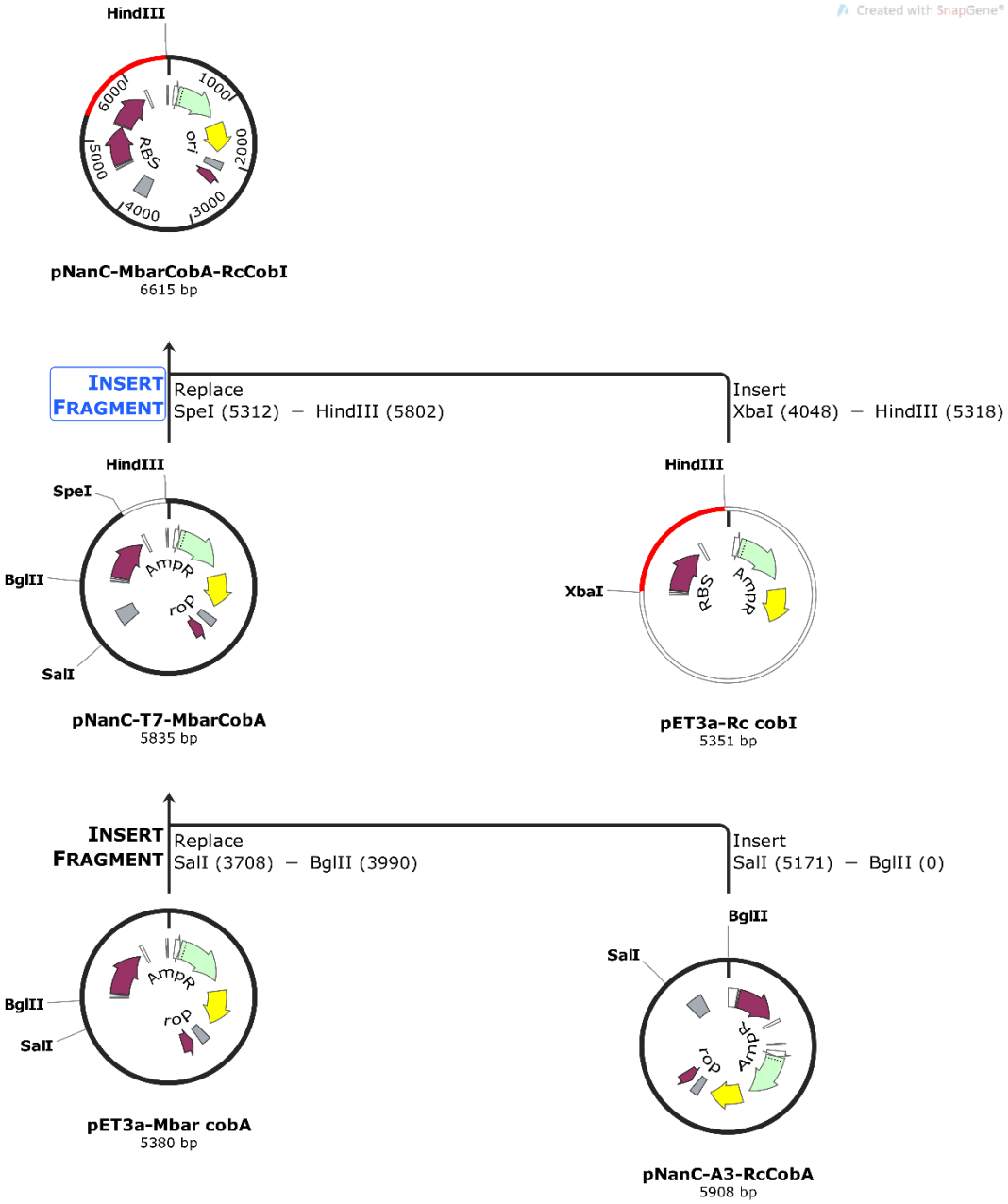


Figure 4.11 Schematic representation of the cloning process of donor DNA plasmid pNanC-*Mbar cobA-Rc cobI.* *pET3a-Mbar cobA was adapted to include 3'flanking region pNanC, digested from pNanC-A3-RcCobA by SalI and BglII, forming pNanC-T7-Mbar cobA. The 3'flanking region was then taken from pET3a-Rc cobI by restriction enzymes XbaI and HindIII and inserted into the SpeI-HindIII fragment of pNanC-T7-Mbar cobA, forming pNanC-Mbar cobA-Rc cobI.*

4.4.3 CRISPR transformation

ED656-3B was transformed with pCas and subsequently with pTarget-N20 cobA as well as the linear donor DNA; Mbar cobA, digested with SalI-HindIII from pET-nanC-Mbar cobA-Rc cobI. The transformation plates resulted in >50 colonies. 10 colonies were tested by colony PCR using primers T7P and MSB-CobA STOP (Figure 4.12). Unfortunately, none of the colonies tested contained the integrated *Mbar cobA* gene. Despite more attempts, it was not possible to achieve this CRISPR integration with the lab time constraints.

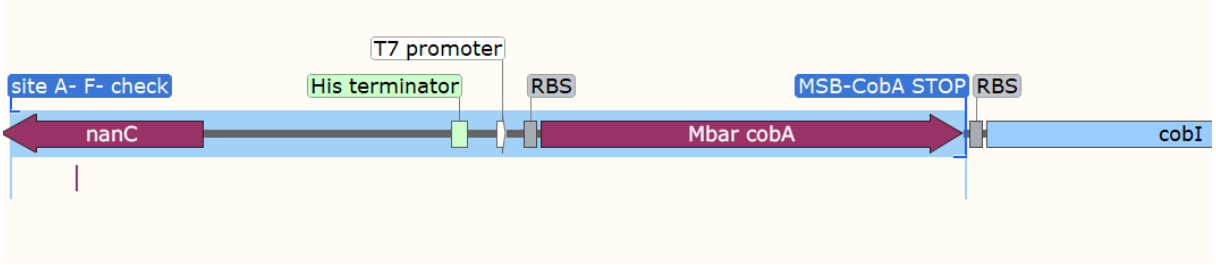


Figure 4.12 Beginning of the B_{12} biosynthetic pathway with the integrated *Mbar cobA* gene. *Highlighted section indicates the PCR fragment that would be amplified via colony-PCR of successful CRISPR strain by using primers site A-F-check and MSB-CobA STOP.*

4.5 Discussion

Early B_{12} biosynthetic pathway genes *Rc hemA*, *Pd cobA*^{K187A} and *Bmei cobI* were added to B_{12}-producing *E. coli* strain ED656-3B via transformation with multicopy plasmids containing these genes. Additional transformations of ED656-3B provided genes *ySAMS*, *hSAMS* and *rSAHH*, which encode enzymes involved in steps feeding into the B_{12} biosynthetic pathway. Under the control of the T7 promoter, induction of gene transcription resulted in overexpression of *Rc* HemA, *Pd* CobA^{K187A} and hSAMS as is seen on the SDS PAGE gel in Figure 4.5. While the bands of *Bmei* CobI and *rat* SAH hydrolase are not apparent, the altered growth and B_{12} production of these cultures showed that expression of these proteins, even to lesser levels, affected the metabolic system of ED656-3B. The effect of any added protein expression on cell growth is made evident in Figure 4.3. The growth curve of ED656-3B containing pET14b- *Pd cobA*^{K187A} shows a lag after protein expression is induced at 6 hours ultimately only reaching 50 % of the OD that ED656-3B reaches at the 24-hour time point. Interestingly, the strain containing pET14b- *Pd cobA*^{K187A} shows the presence of some B_{12} at the point of induction, but continued production is not observed. This coincides with the moment when the growth increases, supporting the idea that the cells shut off B_{12} production by mutating either in the pathway or the T7 promoter in order to survive. The other transformants show a 50 – 75% drop in B_{12} production. The *Mbar cobA* plasmid, which was known to have a huge protein overexpression proved lethal for the B_{12} producing *E. coli* strain. Evidently, the added metabolic strain of overexpressing proteins from a multi copy vector results in a significant decrease of cobalamin production. To avoid this effect, and simulate truer conditions to added copies of the target genes on the genomic DNA, protein expression was placed under the control of a more tightly regulated *tetO* operator.

A difference is seen in the ability to transform ED656-3B successfully with pET-TetR-PtetA – *Mbar cobA*. CobA from *Methanosarcina barkeri* on its original pET3a vector resulted in toxicity for the cell, as was evident from the inability of ED656-3B to be transformed with this plasmid and form colonies. The expression of *Mbar* CobA is seen to result in a doubling of the amount of B_{12} produced by ED656-3B when grown in culture as can be seen in Figure 4.8. These results also show a 50% increase from cultures containing the added *Rc* HemA and rat SAH hydrolase enzymes. The only enzyme which shows a negative impact on B_{12} production results from expressing *Pd* CobA^{K187A}, which has a yield that is ~25% that of the control ED656-3B culture. This difference between *Mbar* CobA and *Pd* CobA is surprising, especially since the K187A mutation of the *Pseudomonas* variant removes its substrate inhibition and should therefore allow a higher productivity.

This, however, could provide the reason of the decrease as a more active enzyme could also result in a higher metabolic cost, taking away from the pathway as a whole. This is not reflected in the cell growth, which Figure 4.7 shows is similar for all cultures grown, but could be again due to the strain mutating to shut off B_{12} production in order to survive.

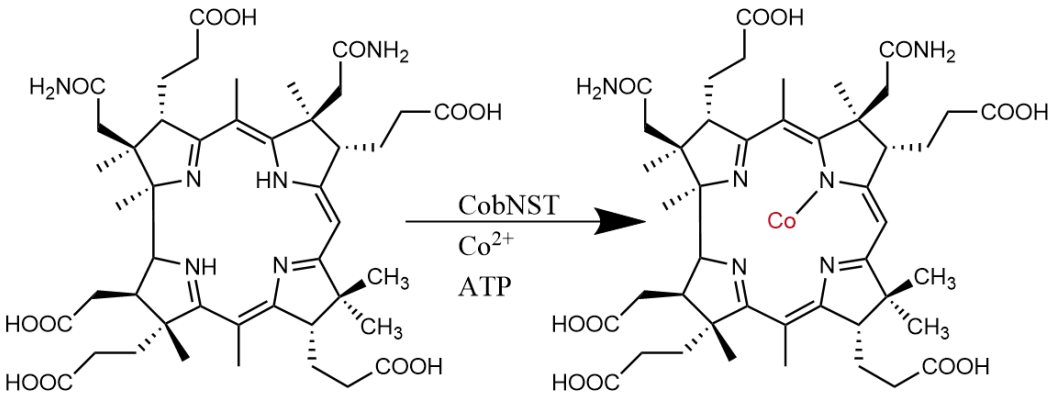
The natural next steps in this project is to investigate the effect of these target genes when inserted into the genome. CRISPR-Cas9 to exchange the *Rc cobA* for *Mbar cobA* on the genome of ED656-3B was attempted, but a successful clone has not yet been obtained. The identified problem lies with the target DNA, which was presented on pTarget-N20 *cobA*. With more time in the lab, a different N20 within the *Rc cobA* would have been selected as well as an alternative method for CRISPR. The CRISPR protocol which was followed here introduces the pCas plasmid and subsequently the pTarget-N20 as well as the donor DNA as a linear digest. This second and double transformation has always been quite inefficient. A new approach, which was developed after my last attempt, has shown to be more efficient. In this protocol, the order in which the DNA components are introduced to ED656-3B, has been modified. The donor DNA is presented to the strain first via transformation with the donor DNA plasmid giving all bacteria the fragment to integrate into the genome. Subsequently, transformation with pCas followed by induction provides the Cas9 endonuclease and lambda red recombinases. Finally, the strain is transformed with pTarget-N20 to allow for the site-specific cut in the genomic DNA. In this order, the double transformation without antibiotic pressure is avoided and the integration rate is much higher.

Chapter 5

Characterisation of the cobaltochelatase CobNST from *B. melitensis* - Determination of wild-type and mutant activity

5.1 Introduction

As with iron in heme, nickel in F_{430} or magnesium in chlorophyll, cobalt in cobalamin is the central atom in a modified tetrapyrrole that is essential for its function. The insertion of these key metal components often represent the branch point of the different biosynthetic pathways such as the heme b pathway and the chlorophyll pathway which diverge at protoporphyrin IX after insertion of magnesium or iron, respectively (Bryant et al., 2020). In the case of cobalamin, along the extended pathway leading from precursor uroporphyrinogen III to vitamin B_{12}, cobalt insertion is done via one of two ways, depending on which route is followed. In the oxygen-independent pathway, cobalt insertion happens early on, before the macrocycle is contracted. This step is catalysed by a class II single-subunit ATP-independent cobaltochelatase. In the oxygen-dependent pathway, however, it is one of the last steps before the two pathways merge again. In this pathway, the insertion poses a bigger challenge, as the contracted ring structure leaves less space for the cobalt ion (Figure 5.1). This technically challenging step is mediated by a class I heterotrimeric cobaltochelatase in a reaction that requires significant quantities of ATP and is analogous to the reactions mediated with the insertion of magnesium in chlorophyll synthesis.



Hydrogenobyrinic acid *a,c*-diamide (HBAD)

Cobyrinic acid *a,c*-diamide

Figure 5.1. Cobalt insertion step in the aerobic biosynthetic pathway of vitamin B_{12}.

This class I cobaltochelatase is a large protein complex that is composed of three subunits, which are termed CobN, CobS and CobT (Debussche et al., 1992). CobN, the largest subunit at 140 kDa, is believed to be responsible for binding the tetrapyrrole-derived substrate, hydrogenobyrinic acid *a,c*-diamide (HBAD), as well as the Co(II) ion (Leeper et al., 2012). CobS and CobT are believed to form a complex with a combined mass of 450 kDa. This complex consists of two stacked rings, each made up as a trimer of dimers (Lundqvist et al., 2009). Sequence analysis has identified characteristic ATP binding motifs, indicating this complex is involved in ATP hydrolysis which is thought to power the cobalt chelation in the CobN subunit.

As mentioned above, the class I cobaltochelatase shares similarity with the magnesium chelatase system in (bacterio)chlorophyll biosynthesis. In contrast to the magnesium chelatase, the cobaltochelatase has not been studied intensively. The magnesium chelatase been studied in multiple organisms including *Rhodobacter capsulatus* (Willows & Hansson, 2003), *Chlorobium vibrioforme*, *Synechocystis PC6803* (Reid & Hunter, 2002) and *Rhodobacter shpaeroides* and therefore provides a good model enzyme. The magnesium chelatase is also composed of 3 peptides: ChlH/BchH, which at 145 kDa shares similarity with CobN, and ChlI/BchI and D (40 kDa and 65 kDa) which form the AAA+ like chaperone complex responsible for ATPase activity much like the CobST complex is for the cobaltochelatase.

Previous research from the Warren group has resulted in co-crystallisation of a CobN fragment with its substrate, HBAD. This structure reveals the binding pocket and residues which might be involved in the peptide-molecule interaction. Based on this information, the group of Neil Hunter designed the corresponding fragment of ChlH from *T. elongatus* for their research into magnesium chelatase. Their publication of the active site of magnesium chelatase found an essential glutamate (E660), which was proposed to be the key catalytic residue for magnesium insertion (Adams et al., 2020). They also found the mutation of D1177 to an alanine to improve the catalytic efficiency. Additionally, the structure of a substrate-free *Mycobacterium tuberculosis* cobaltochelatase CobN has recently been reported. Here the authors have proposed a triad of residues involved in cobalt binding, formed from H522, E561 and D1037, residues that are found within the interior of CobN (J.-H. Zhang et al., 2021). The E561 residue in *Mt* CobN corresponds to E660 found in the ChlH subunit of the magnesium chelatase.

5.1.1 Project aim

The work presented in this chapter describes the characterisation of the cobaltochelatase from *Brucella melitensis*. Based on the research into ChlH and *Mt*CobN (described above), *Bmei*CobN was mutated at specific residues to investigate the impact of these amino acids on the activity of the enzyme with the ultimate aim of improving the turnover rate to improve vitamin B_{12} production.

5.2 Production and purification of HBAD

To be able to perform activity assays with the cobaltochelatase enzyme, first its substrate HBAD needs to be made available. As HBAD is not commercially available, it needs to be prepared by extraction from a bacterial strain that overproduces this intermediate. In this section, the production of HBAD and its subsequent purification is described.

E. coli strain ED661, which contains all the genes that encode for the enzymes required for HBAD production, was kindly provided by Dr. Evelyne Deery and grown as described in Section 2.2.11. In actual fact, ED661 houses the genes that encode the enzymes for hydrogenobyric acid, which is the fully amidated form of hydrogenobyrinic acid. For reasons that are not fully understood, this strain secretes hydrogenobyric acid but retains HBAD within the cell. Thus by collecting the cells and then lysing them, the extract contains a good proportion of HBAD, which can then be subsequently purified. The purification process was optimised to include the use of a recently-reported B_{12}-binding protein called BtuG2. BtuG2 is a surface-exposed lipoprotein that has been shown to be essential for B_{12} transport in the gut commensal *Bacteroides thetaiotaomicron* (Wexler et al., 2018). It was anticipated that BtuG2 may be able to bind a variety of cobamide intermediates including HBAD.

5.2.1 Column purification of HBAD

Recombinant *B. thetaiotamicron* BtuG2 was purified from a recombinant *E. coli* strain (see Section 2.2.7) after the gene had been cloned into a pET14b plasmid to allow the protein to be produced without its N-terminal signal sequence but with a N-terminal His-tag. The resultant ^{6xHis}BtuG2 was immobilised on a Ni^{2+}-Sepharose column after which residual proteins and from the cell lysate were washed away with a low amount of imidazole. Boiled ED661 cell lysate, containing HBAD (and other amidated B_{12}-intermediates), was centrifuged to obtain the soluble fraction. This was applied to the BtuG2-loaded Sepharose column and purified as described in Section 2.4.5. Figure 5.2 shows the colour retention of the Ni^{2+}-Sepharose column after washing away excess protein, indicating that the HBAD had successfully bound to BtuG2. The elution of the HBAD was easily monitored by following the movement of colour.

Figure 5.2 Immobilised BtuG on Ni^{2+} - sepharose columns, trapping HBAD-like B_{12} intermediates. *ED661 cell lysate was loaded onto BtuG-columns (left), and bound HBAD (right).*

The eluted mix at this point contained hydrogenobyrinic acid in different amidated states, all bound to BtuG2. This includes HBAD, but also the mono-amide – all the way up to the hexa-amide (hydrogenobyric acid). This mixture was boiled to release HBAD from the BtuG2 protein, which resulted in its precipitation. The precipitate was spun down and the soluble fraction, containing the intermediates, was loaded onto a diethylaminoethyl sepharose (DEAE) anion-exchange column. As the HBAD-like molecules are negatively charged at neutral pH, these are bound by the positively charged protonated amine within the exchange resin. The different numbers of carboxylic acid groups within the amidated forms of hydrogenobyrinic acid results in differences in charge between the different intermediates. This allows for their separation on this column through application of increasing concentrations of salt solutions (Figure 5.3). Fraction elution was based on colour movement through the column which resulted in 5 separate fractions which were termed: flow-through, 50 mM NaCl eluate, 100 mM, and two bands at 150 mM.

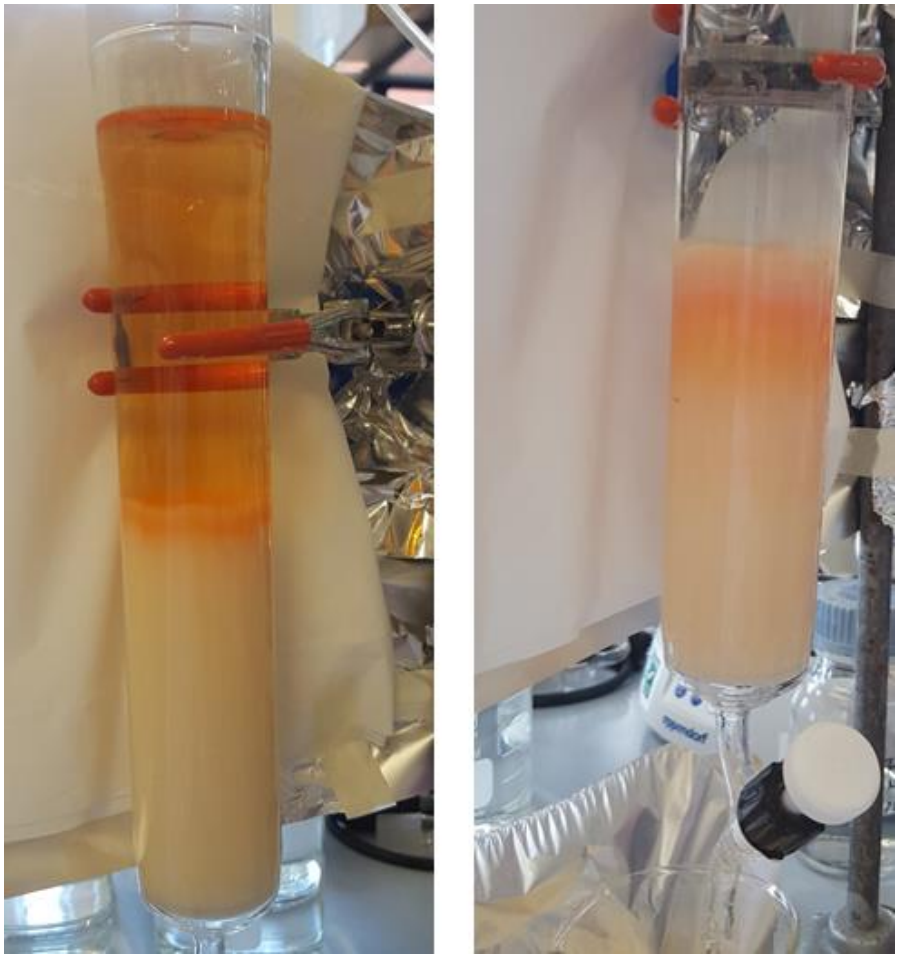


Figure 5.3. Separation of multi-amidated vitamin B_{12} intermediates via DEAE anion exchange column. *Loading of HBAD onto DEAE column (left) and separation of bands after applying salt (right).*

The salt was removed from the sample by applying the samples to a reverse-phase column. Each eluted fraction was adjusted to pH 4 before being bound to an RP18 column. Elution with 100% MeOH resulted in concentrated samples which were dried down by vacuum centrifugation. To obtain an indication of which fraction contained the relevant intermediate, as well as the purity of these fractions, samples were run on a TAE gel. Agarose electrophoresis allows the amidated forms of hydrogenobyrinic acid to be separated on the basis of their charge. Previously authenticated samples of HBA and HBAD were kindly provided by Dr Evelyne Deery to include as control samples. The gel is seen in Figure 5.4.

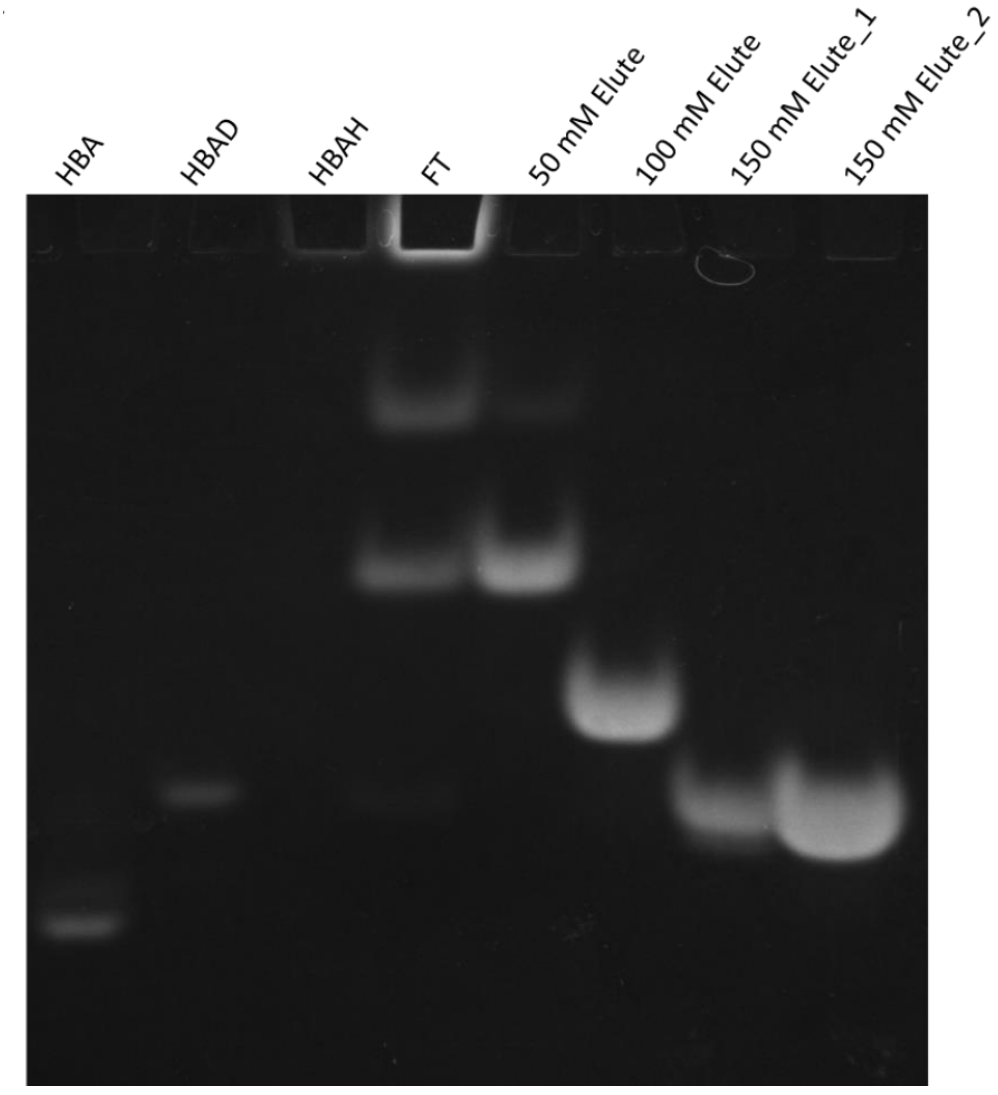


Figure 5.4 TAE gel of purified B_{12} intermediates. *HBAD-like intermediates were purified via BtuG and DEAE columns. The elution salt concentrations for each fraction is stated above each lane. Samples were concentrated via RP18 column of which 20 µL sample – along with 5 µL glycerol – was loaded to be separated based on charge. Pure HBA and HBAD samples were included as a control, which are found in lanes 1 and 2.*

As the gel indicated the 150 mM salt elution fractions contained pure HBAD, these were pooled and dried down. Although these were collected as separate fractions, the bands, whilst moving downwards, broaden, increasing the difficulty of seeing where one band ends and the next starts. In this case, it appears both belonged to the same sample.

5.2.2 HBAD sample verification via mass spectrometry

The HBAD sample shown in Figure 5.4 appears to run as a single species on the gel. For a more accurate indication of the purity of the sample, a small amount was analysed via HPLC-MS. Samples were separated on an Agilent 1100 series HPLC coupled to a micrOTOF-Q II (Bruker) mass spectrophotometer using an Ace 5 AQ column (2.1 x 150 mm; Advanced Chromatography Technologies) maintained at 30˚C and with a flow rate of 0.2 mL/min. The mobile phase consisted of 0.1% TFA in water (solvent A) and acetonitrile (solvent B). The initial conditions were set at 100% of solvent A. The concentration of solvent B was increased with a linear gradient to 100% over 50 min and back to the initial conditions at 60 min. The mass spectrometer was operated in the positive electrospray mode.

The HPLC chromatogram at 200-800 nm, seen at the top of Figure 5.5, shows a single peak coming off the column between 26.6 and 29.1 minutes, indicating a high level of purity of the analysed sample. When looking at the chromatogram at a wavelength of 329 nm, one peak is shown again at the same retention time. As this is the specific wavelength for HBAD, it is a further indication that the peak corresponds to HBAD. To confirm this, the sample was further analysed via mass-spectrometry, the results of which are seen in the bottom two graphs of Figure 5.5. Here, the full spectrum of the peak at r_t 30 min is given, which shows a spectrum identical to HBAD; a peak at 329 nm and the double peak around 495 and 522 nm. Below this spectrum, the MS data is shown, which, combined with the spectrum, allows for a positive identification of HBAD which has a mass of 879.45 as is seen in this graph.

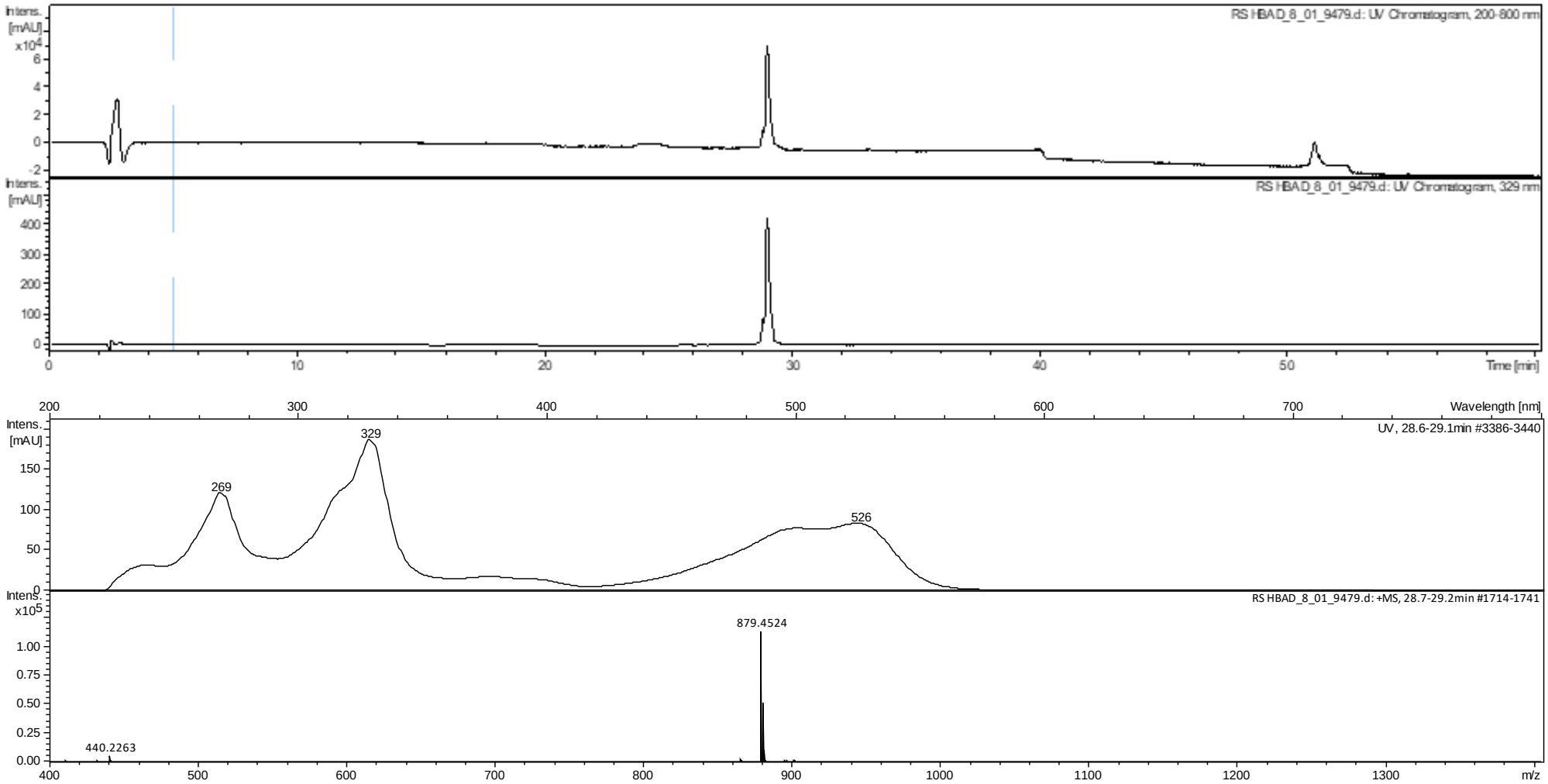


Figure 5.5 LC-MS analysis of HBAD sample. Top: HPLC chromatogram at 200-800 nm, showing a single peak coming off the column between 26.6 and 29.1 min indicating a pure sample. Below, the chromatogram at 329 nm which is the specific wavelength for HBAD. The same peak is shown, further indicating this relates to HBAD. UV analysis of this peak (third graph) shows distinctive spectrum of HBAD; a single peak at 329 nm, and the double peak at 495 and 522 nm. Bottom: MS analysis. The molecule can be confirmed HBAD with a mass of 879.

5.3 Cloning of Bmei ^{H}cobN - mutants

As mentioned in the Section 5.1, the combined results achieved in the Warren lab, the structure of a CobN from *M. tuberculosis* and the detailed analysis of the Mg-chelatase system prompted a need to investigate whether the CobNST chelatase system could be improved in terms of its catalytic turnover. The sequences of CobN from *M. tuberculosis* and ChlH from *Synechocystis sp.* PCC 6803 were aligned with the *B. melitensis* CobN to identify key residues within the latter sequence that may be essential for controlling either substrate/product binding or catalytic turnover (Figure 5.6). From these alignments, the cobalt triad initially identified within the *M. tuberculosis* sequence corresponds to H554, E593 and D1101 in the *B. melitensis* CobN. Furthermore, based on the structural data obtained in the Warren lab, R912, H1094, D1098 and Q1104 were identified as residues of interest as they appear to mediate interaction between CobN and its substrate, HBAD (Figure 5.7)

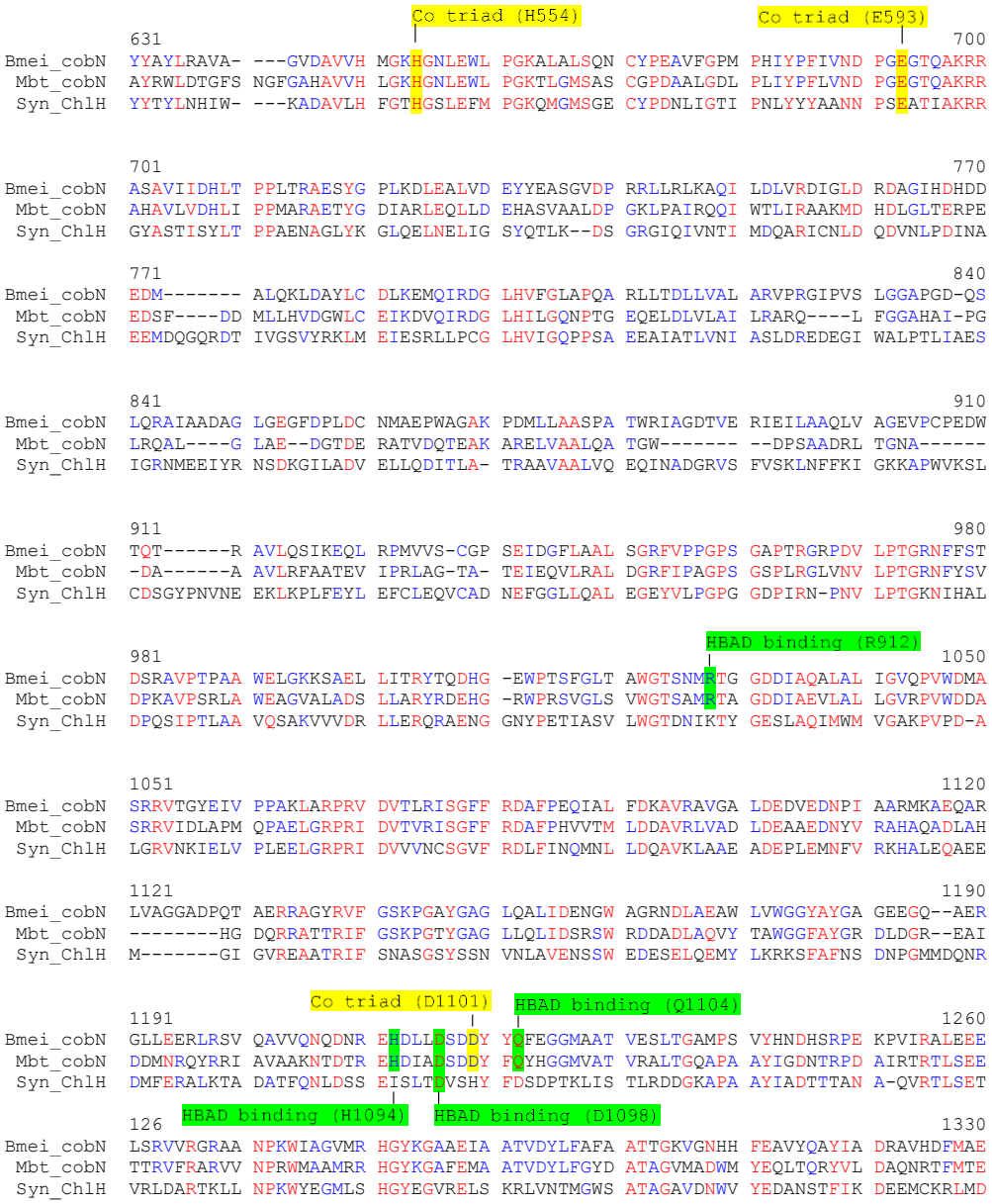


Figure 5.6 Alignment of *Brucella melitensis* CobN, *Mycobaterium tuberculosis* CobN and *Synechocystis* sp PCC 6803 ChlH. *Highlighted in yellow are the residues identified as the cobalt triad by Zhang et al. This includes corresponding E660 in ChlH. The corresponding residues responsible for cobalt binding in Brucella melitensis were identified as H554, E593 and D1101. Highlighted in green are residues within B. melitensis cobN which are likely involved in binding of its substrate; hydrogenobyrinic acid a,c- diamide, based on structural data of a partial co-crystallisation of B. melitensis CobN with HBAD.*

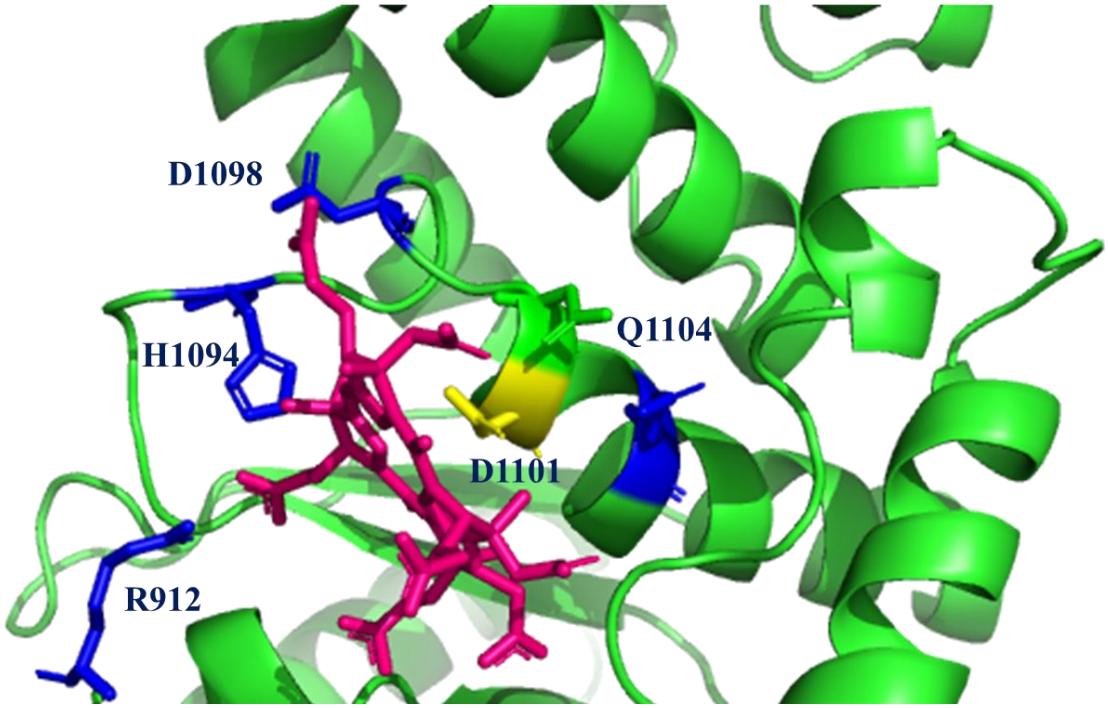


Figure 5.7 Partial structure of the CobN binding pocket of with HBAD bound. *Highlighted in yellow is D1101, which is likely involved in cobalt binding. The other two residues which make up the cobalt triad are not part of this partial structure and are therefore not shown. Amino acids coloured blue were targeted for mutation design as they seem to be involved in the HBAD-protein interaction. Based on data obtained by Dr Evelyne Deery and Dimitrios Ladikis.*

To investigate the role of these residues of interest they were indivually changed to an alanine. This amino acid is neutral in charge, which would change the function both in case of H554 and H1094 (involved in cobalt and HBAD binding, respectively), which could be dependent on their positive charge for function, and E593, D1098 and D1101, which is negatively charged. However, it is hydrophobic and has a small side chain which minimises the chance of a structural change in the peptide's folding as a whole and it is well-known that alanine residues can be accommodated into the various secondary structure elements that are found within protein structures.

The wild-type cobaltochelatase from *Brucella melitensis* was kindly provided by Dr Evelyne Deery on 2 plasmids; pET14b - *Bmei* ^{6xHis}*cobN*, and pET3a- *Bmei cobS*-^{6xHis}*cobT*. Mutations in the *cobN* gene were introduced using overlap extension PCR (Sambrook & Russell, 2001). Forward and reverse mutagenic primers were designed and matched to *cobN*-specific primers to allow for introduction of the mutation (Figure 5.8). In the first round of PCR, the forward mutagenic primer and the reverse *cobN*-specific primer were applied in one reaction, whilst the reverse mutagenic primer and the forward *cobN*-specific primer were used. The resulting DNA strains were added 1:1 for a second round of PCR,

where the *cobN*-specific primers were added to create the complete strand including the mutation.

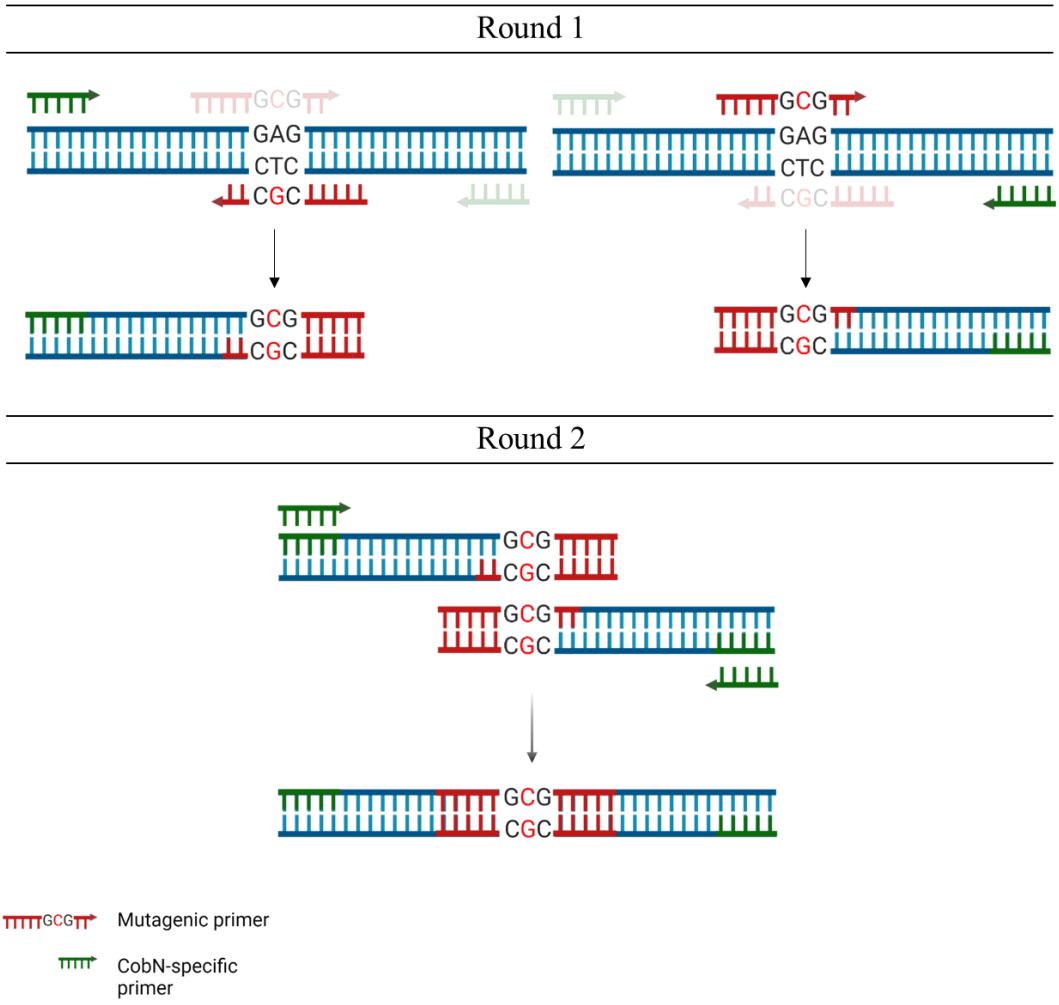


Figure 5.8 Overlap extension PCR procedure. *In the first round of PCR, the forward mutagenic primer was paired with the reverse* cobN*-specific primer. Simultaneously, in a second reaction, the reverse mutagenic primer was paired with the forward* cobN*-specific primer. The resulting strands were added in a 1:1 ratio and the* cobN*-specific primers were used to create the final strand of DNA containing the desired mutation.*

These were used to perform PCR using pET14b - *Bmei* ^{6xHis}*cobN* as the template. These two PCR products were then used as the template for a subsequent overlap extension PCR. This final fragment was digested with restriction enzymes and inserted into pET14b - *Bmei* ^{6xHis}*cobN*. The primers used for each reaction can be found in Table 5.1 along with the restriction enzymes used to clone the mutation into the vector.

Table 5.1. Cloning design for each *cobN* mutant.

Mutation	1st round PCR primers		2nd round PCR primers	Restriction enzymes for cloning
	Fw reaction	Rev reaction		
E593A	BmeiCobN_E593A_F BmeiN_AmpTAA_Rev	BmeiCobN_E593A_R Bmei N3	Bmei_AmpTAA_Rev BmeiN3	NdeI + MluI
H1094A	BmeiCobN_H1094A_F T7T	BmeiCobN_H1094A_R BmeiN-mut-Nde	T7T Bmei-mut-Nde	StuI + AscI
D1098A	BmeiCobN_D1098A_F T7T	BmeiCobN_D1098A_R BmeiN-mut-Nde	T7T Bmei-mut-Nde	StuI + AscI
Q1104A	BmeiCobN_Q1104A_F T7T	BmeiCobN_Q1104A_R BmeiN-mut-Nde	T7T Bmei-mut-Nde	StuI + AscI

The process of cloning pET14b - *Bmei* $^{6xHis}cobN^{E593A}$ is illustrated in Figure 5.9. This was repeated using primers and enzymes listed in Table 5.1 to produce mutant plasmids pET14b - *Bmei* $^{6xHis}cobN^{D1098A}$, pET14b - *Bmei* $^{6xHis}cobN^{H1094A}$, and pET14b - *Bmei* $^{6xHis}cobN^{Q1104}$. Test digests were performed using the relevant restriction enzymes and subsequent sequencing confirmed the correct constructs had been generated.

Despite best efforts, *Bmei* $^{6xHis}cobN^{H554A}$ could not be constructed. This cloning was designed as a one-round PCR, as the mutation site was next to a restriction site. PCR products were created at the expected band size, however, cloning of the restricted fragment into the vector remained unsuccessful. COVID related time-restraints prevented the re-design of the cloning approach.

The vectors containing mutations R912A and D1101A were kindly donated by Dr Evelyne Deery.

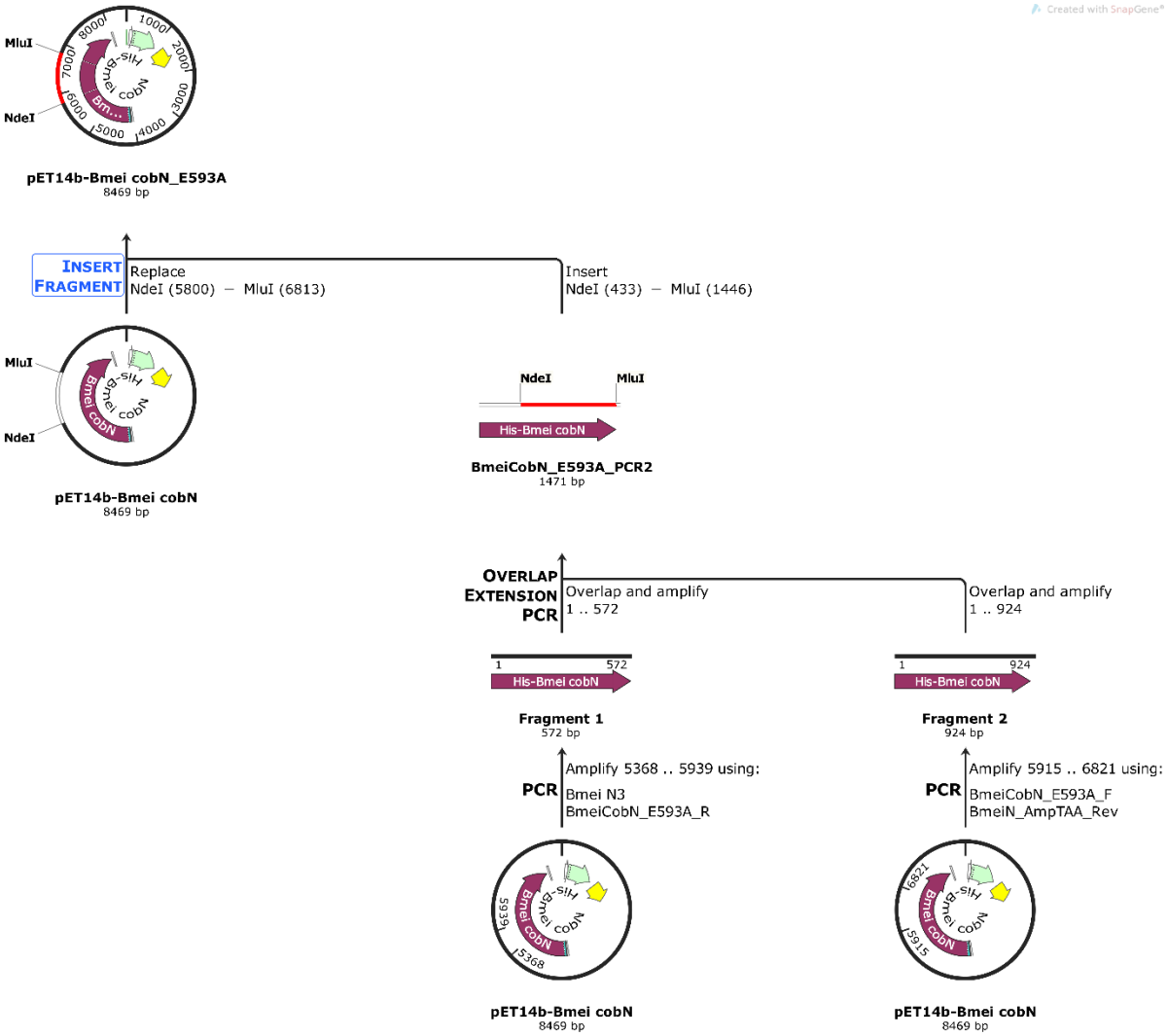


Figure 5.9. Cloning history pET14b - *Bmei* ^{6xHis}*cobN*^{E593A}. *Mutations were introduced by a forward and reverse PCR using primers containing the E593A mutation, and pET14b - Bmei* ^{6xHis}*cobN as template. The PCR products were subsequently used for an extension overlap PCR. This final product was digested with restriction enzymes NdeI and MluI and inserted into the original vector.*

5.4 Expression and purification of the cobaltochelatase CobN subunit from *Brucella Melitensis*.

For the purpose of comparing wild-type to mutant versions of the *B. melitensis* cobaltochelatase, the CobN subunit was produced and purified separately from CobST. Competent BL21 Star (DE3) pLysS cells were used to express the wild-type and mutant proteins as described in Section 2.2.7. Purification of the hexa-histidine tagged proteins was achieved via nickel affinity chromatography. The fractions produced were analysed via SDS PAGE as shown below. Fractions containing the most concentrated and purest CobN were identified as indicated with the red box in Figure 5.10. These fractions were pooled and concentrated before being applied to a PD10 column to exchange the buffer to 20 mM HEPES, pH 7.5 containing 100 mM NaCl, which is the buffer used for the assay work. The concentrations of the purified protein fractions were determined by measuring the absorbance at 280 nm and treating this with the Lambert-Beer law (Section 2.4.3) where ε_{cobN} at 280 nm is 167900 M^{-1} cm^{-1}. The purified protein was stored at -20°C in 5% glycerol.

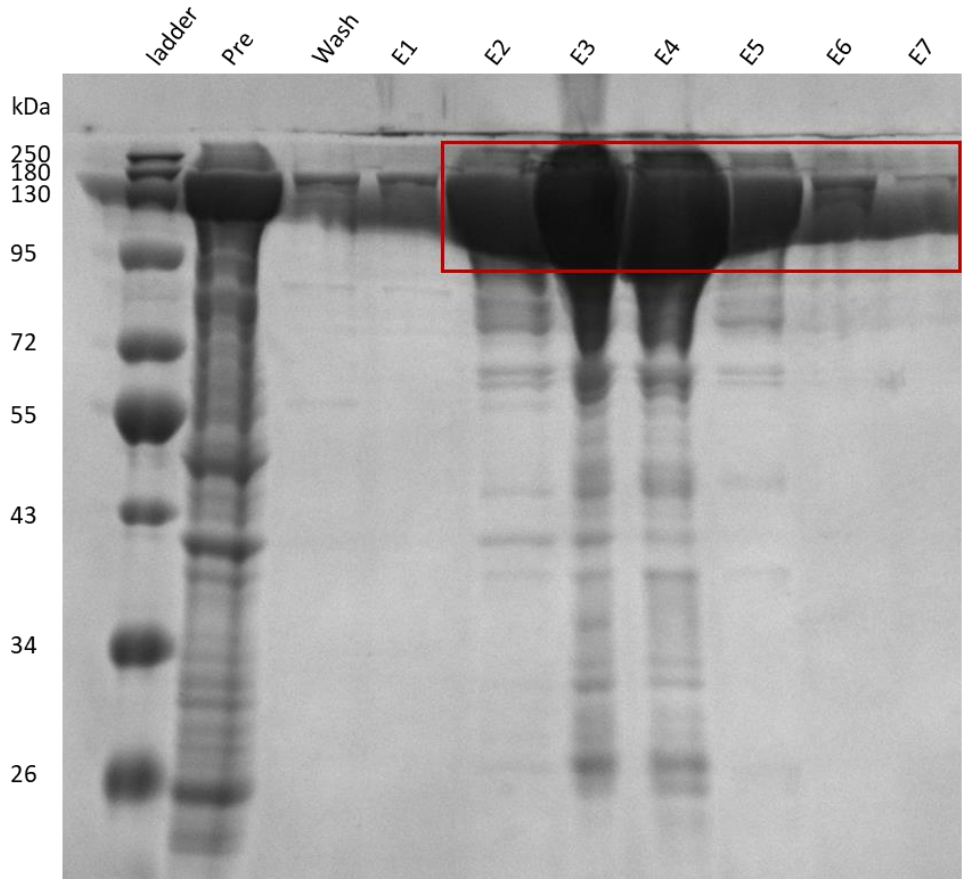


Figure 5.10 SDS PAGE analysis of *B. melitensis* CobN fractions as procured by nickel affinity chromatography. *Pre indicates the pre-column sample, directly taken from the soluble fraction of the cell lysate. Wash corresponds to washing the column with 20 mM HEPES pH 7.5 containing 500 mM NaCl and 60 mM imidazole. E1 to E7 correspond to eluted protein fractions via application of 400 mM imidazole. Sample and 2X SDS sample buffer were mixed 1:1 and 5µl was loaded on the gel. With a predicted mass of 140 kDa, CobN was identified in all fractions. The red box outlines the fractions which were pooled for further work.*

5.5 Production and purification of the cobaltochelatase CobST subunits from *Brucella melitensis*.

CobS and CobT were expressed in *E. coli* BL21star DE3 pLysS-Rc DNAJ as a complex where CobT was His-tagged to enable purification as described in Section 2.2.7. Purifying CobST as a complex ensured the ideal ratio of CobS to CobT, which is crucial for the enzyme's activity. The fractions obtained from the nickel affinity chromatography were analysed via denaturing polyacrylamide gel electrophoresis in order to identify those that contained the protein complex. The SDS-PAGE of the fractions is shown in Figure 5.11.

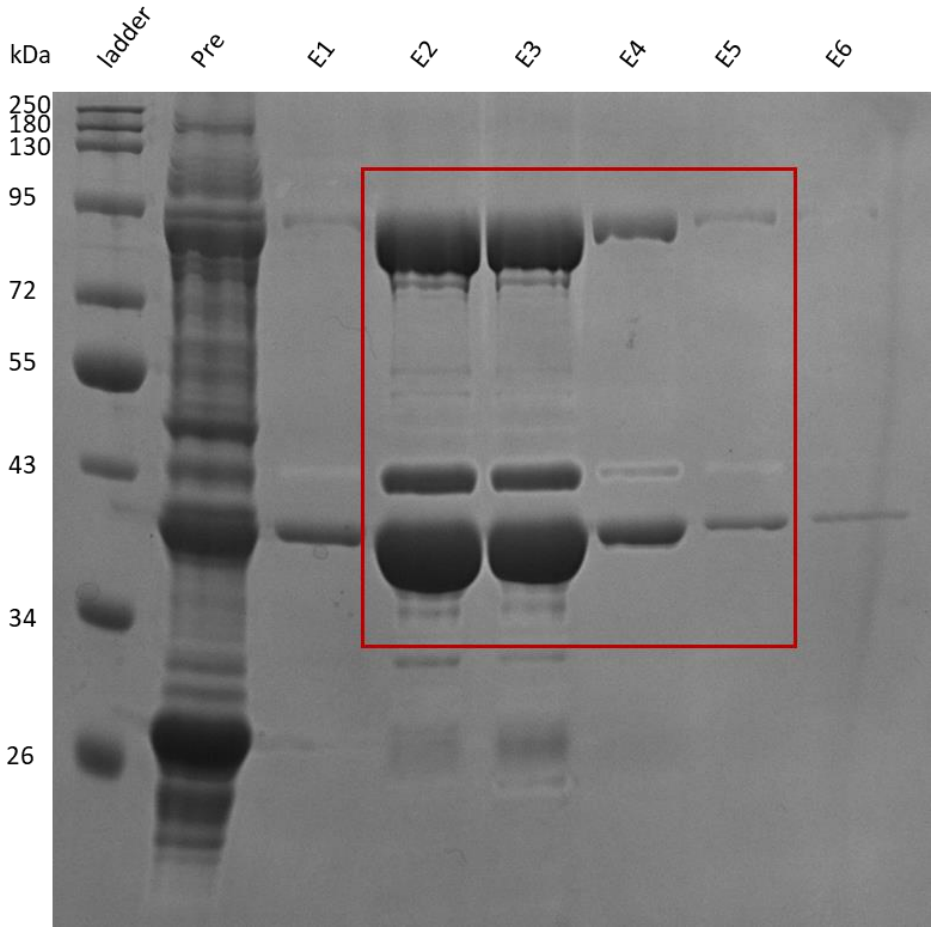


Figure 5.11 SDS PAGE analysis of *B. melitensis* CobST fractions as procured by nickel affinity chromatography. *Pre indicates the pre-column sample, directly taken from the soluble fraction of the cell lysate. Wash corresponds to washing the column with 20 mM HEPES pH 7.5 containing 500 mM NaCl and 60 mM imidazole. E1 to E6 correspond to eluted protein fractions via application of 400 mM imidazole. Sample and 2X SDS sample buffer were mixed 1:1 and 15 µL was loaded on the gel. Of protein standard ladder, 5 µL was loaded. Predicted mass of His-tagged B. melitensis CobT is 72.8 kDa and 36.9 kDa for CobS. The red box highlights the fractions which were pooled for further work.*

The gel shows the bands of CobS and His-tagged CobT at the predicted masses of 36.9 kDa and 72.8 kDa, respectively, in fractions E2-E5. There is another visible band around 41 kDa which could not successfully be removed by altering the wash steps and, therefore, possibly represents a breakdown product of CobT. The pooled fractions, as indicated by the red outline in Figure 5.11, were concentrated and buffer-exchanged via PD10 column into 20 mM HEPES, pH 7.5, containing 100 mM NaCl. Protein concentration was again determined by measuring the absorbance at 280 nm and using the extinction coefficient ε_{cobST} = 88997.5 M^{-1} cm^{-1} to convert the values with the Lambert-Beer law.

5.6 Functional characterisation of cobaltochelatase (-mutants) of Brucella *melitensis*

To compare the activity of the cobaltochelatase wild-type to those which have CobN-mutations, a multi-step characterisation was done. Before the enzyme activity could be measured, a K_D value needed to be obtained for each CobN – CobST as these proteins were expressed and purified separately. This was done via assays where [CobN] was fixed, and [CobST] was varied, creating a Michaelis – Menten (MM) curve where CobST acted as the substrate for CobN. Using this K_D value, the total amount of functional enzyme could be calculated. Subsequently, MM curves were plotted where the substrate cobalt was varied to produce kinetic parameters K_M, k_{cat} and k_{cat}/K_M. These values give an initial insight into the functional character of an enzyme, allowing a comparison between the various mutants and wild-type CobN to be made. Equation 2 represents a simplified reaction mechanism which Michaelis and Menten based their first definitions upon (K. A. Johnson & Goody, 2011).

Equation 2

$$\mathbf{E + S} \underset{\boldsymbol{k_2}}{\overset{\boldsymbol{k_1}}{\rightleftharpoons}} \mathbf{ES} \xrightarrow{\boldsymbol{k_3}} \mathbf{E + P}$$

The Michaelis constant K_M ($= \frac{[E][S]}{[ES]} = \frac{k2+k3}{k1}$) represents the concentration at which half the maximal velocity is observed, and is obtained by fitting a MM curve. k_{cat} ($= \frac{Vmax}{[E]} = k3$) is a catalytic constant and represents the steady state turnover number. The specificity constant k_{cat}/K_M ($= \frac{k1k3}{k2+k3}$) gives insight into the specificity of an enzyme for the substrate.

5.6.1 CobN-ST binding curves

The activity of the wild-type as well as the mutant variants of *B. melitensis* CobNST was investigated via steady state kinetics. As CobN, including its variant forms, and CobST were expressed separately, the functional amount of enzyme needed to be determined for each batch. CobN-ST binding curves were initially generated, where CobN was fixed at 0.04 μM whilst CobST was added in a range of 0 – 4 μM. This resulted in MM curves where CobST acted as the substrate for CobN. Furthermore, the reaction mix contained 20 mM HEPES, 100 mM NaCl,10 mM $MgCl_2$ and 5 mM ATP. Under these conditions the substrates cobalt and HBAD were added in excess at 0.02 mM and 0.5 μM, respectively. CobN and CobST were pipetted into a 96-well plate and each was adjusted to the same volume with buffer (20 mM HEPES, pH 7.5, 100 mM NaCl). The plate was then pre-incubated at 35 °C to approach the enzyme's optimal *in vivo* conditions. The reaction was initiated by injecting a mastermix, containing $MgCl_2$, ATP, cobalt and HBAD in buffer. The reaction was followed for 50 minutes by monitoring a decrease in fluorescent emission signal at 620 nm when excited at 510 nm. This measured the conversion of HBAD to cobyric acid *a,c*- diamide, where the insertion of the metal ion into HBAD quenches the fluorescence. All reactions were measured in a BMG labtech Omega plate reader. The slopes of these reactions are seen as raw data below. Each set of assays contained the WT cobaltochelatase to ensure identical conditions for comparisons.

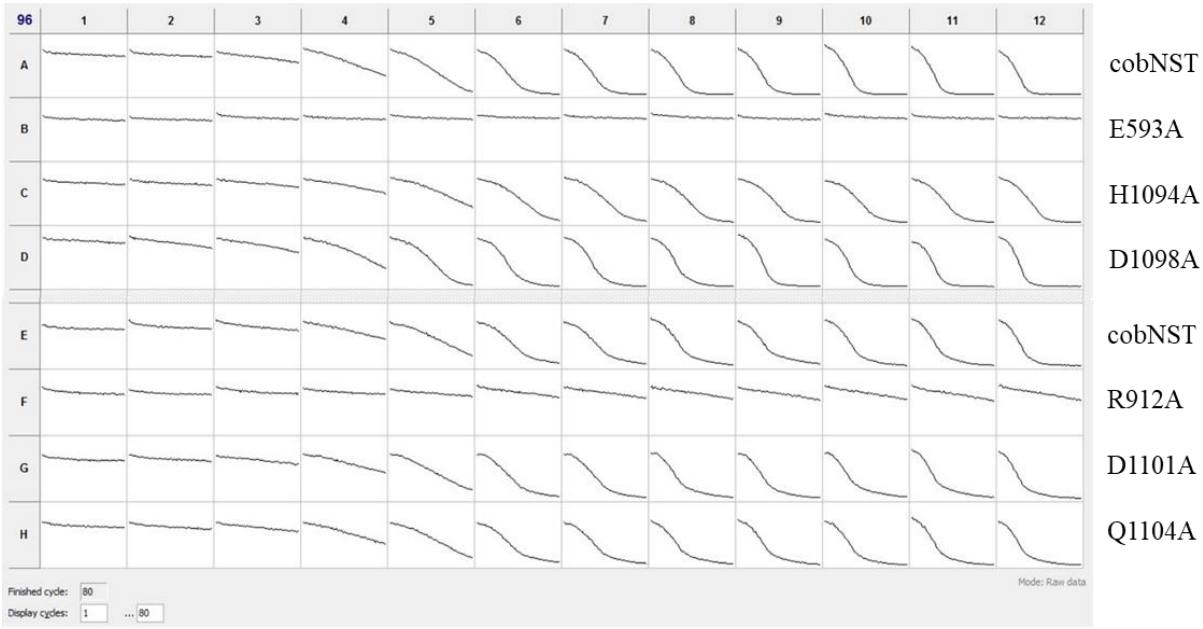


Figure 5.12 Enzyme activity Bmei cobaltochelatase WT and mutants. *Image of raw data from the plate reader. Reduction of fluorescence is seen as the result of HBAD cobalt chelation. Lack of activity is seen for mutants E539A and R912A. Numbers 1-12 represent the range of CobST of 0 to 4 μM.*

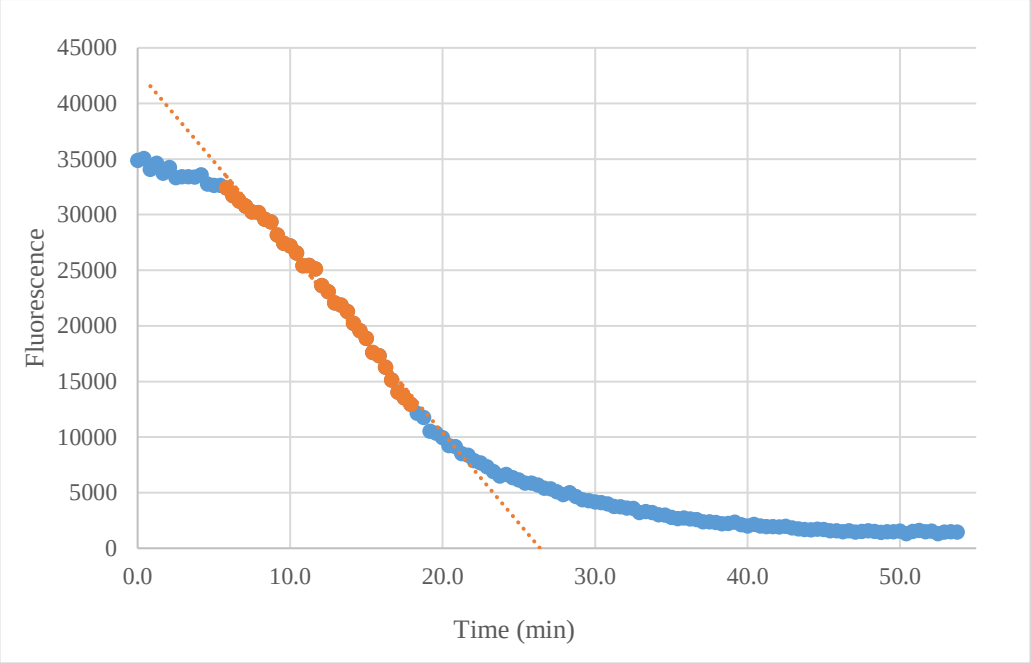


Figure 5.13 Reaction curve showing decreased fluorescent signal as a result of cobalt insertion into HBAD by wild-type *B. melitensis* cobaltochelatase. *Reaction shows a lag phase of approximately 6 minutes after which the reaction speeds up. The slope of the trend line indicates the reaction speed within the time frame of 6-18 minutes, within which each reaction velocity was measured.*

Figure 5.13 shows a single reaction progression where there is a lag phase of approximately 6 minutes, after which the reaction speeds up. This could be associated with a conformational change to activate CobN into a "ready" state. As there are many factors at play in this reaction, however – the combining of CobN-CobST to form a fully functional cobaltochelatase complex, the ATP hydrolysis which powers the chelation of HBAD, and then the reaction of cobalt insertion into HBAD itself (which is what is measured) – one cannot say which factor is shown by this trend. For this reason, the change in signal over time was taken between time points 6 – 18 minutes, where the slope was steepest, and where all reactions still showed a linear trend. Based on this raw data (Figure 5.12), it can be concluded that mutant CobN^{E593A} is rendered inactive, whereas the cobaltochelatase containing the mutation R912A is highly impaired.

To convert the ΔFluorescence/min to ΔHBAD/min, a standard curve of HBAD was generated where CobN was fixed at 0.04 µM to which a range of 0 – 0.5 µM HBAD was added (Figure 5.14). Two separate conditions were then tested. In one set of conditions, the volume was adjusted by adding just buffer (orange curve), whilst to the other the assay conditions were mimicked by adding 0.02 mM $CoCl_2$, 10 mM $MgCl_2$, and 5 mM ATP (blue curve). The mixes were left to incubate for 2 minutes before being measured in triplicate. Though minor, there is a slight difference between the two curves, showing that the presence of ATP and cobalt results in a slight increase in fluorescence. The curve more closely mimicking the assay conditions was therefore used to convert the data.

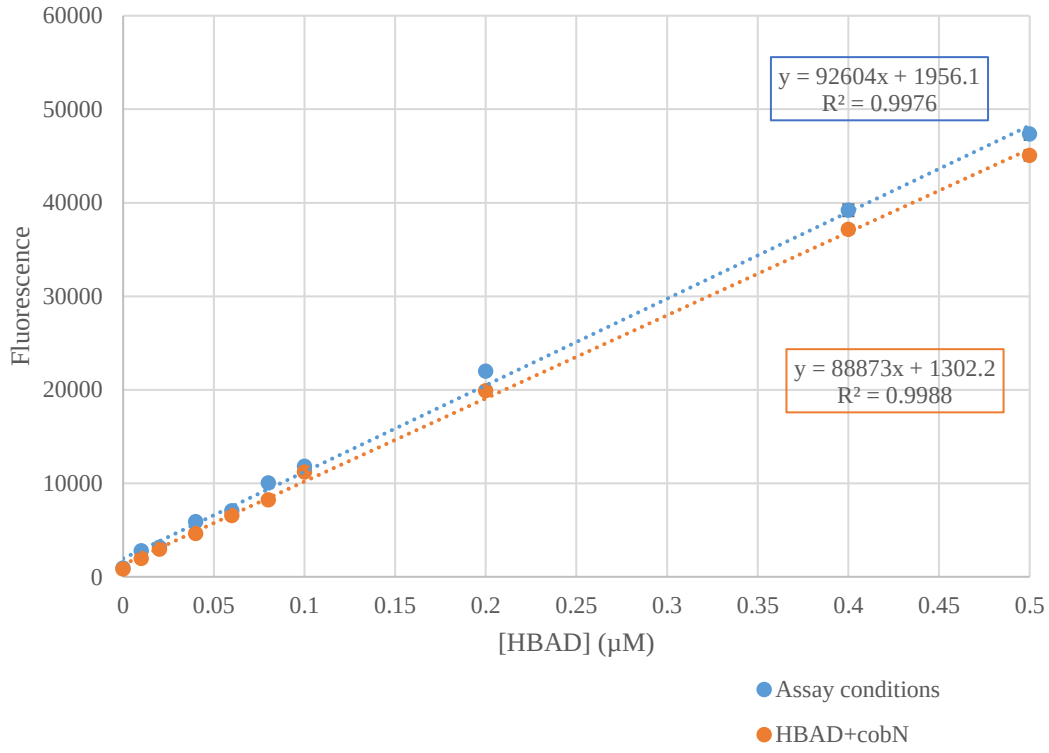


Figure 5.14. HBAD standard curve. *The fluorescent signal of HBAD was monitored under two sets of conditions; 0.04 µM CobN added to a range of HBAD in buffer (orange) and the same curve, now including cobalt and ATP (blue), more closely mimicking activity assay conditions.*

The conversion via a standard curve is important, as the binding of HBAD to CobN has been shown to increase the fluorescent signal of the former (Figure 5.15). This unexpected effect was measured as the fluorescent signal of 0.5 µM HBAD in buffer, with CobN varied from 0 – 1.5 µM. The increase of fluorescence as a result of HBAD binding to CobN could be the result of increased delocalisation of the double bonds in HBAD upon binding to the protein.

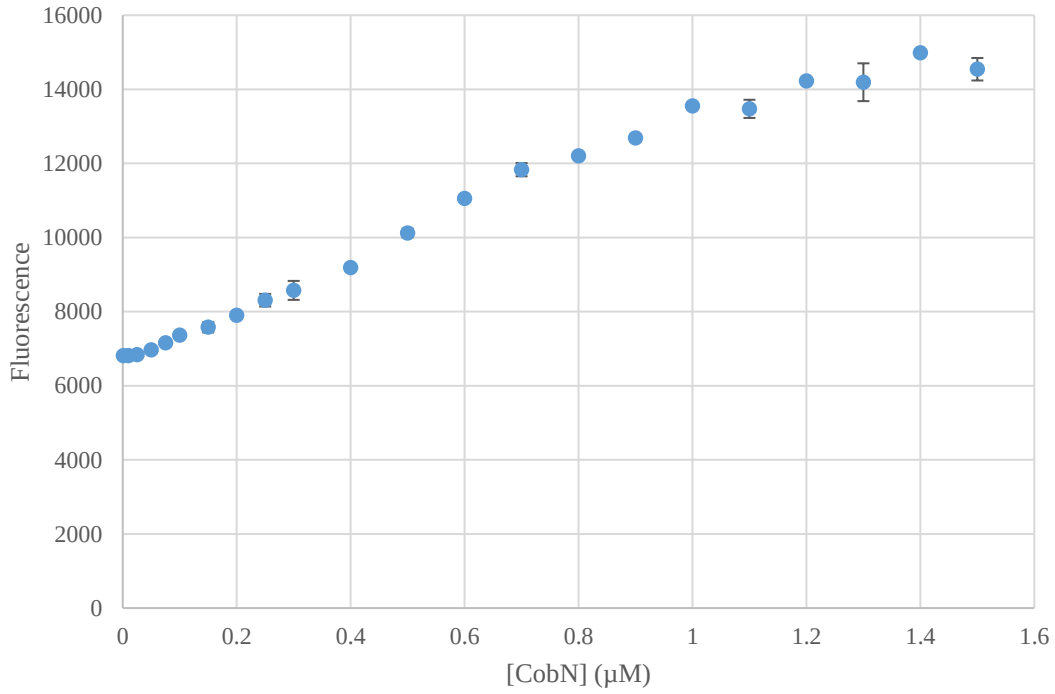


Figure 5.15 CobN-HBAD binding curve. *Increase in fluorescent signal measured as a function of 0.5 µM HBAD with CobN added in a range of 0 – 1.5 µM. This curve was executed in duplicate; the error bars indicate the SD of the two values.*

The calculated reaction velocities in nM min^{-1} were plotted in OriginPro and fitted to Equation 4, derived from the Michaelis-Menten equation (Equation 3), resulting in the curves shown in Figure 5.16. Here, the K_D value represents the concentration of CobST in μM, where 50% of fully functional enzyme is formed under these conditions.

Equation 3

$$v = \frac{Vmax * [S]}{K_M + [S]}$$

Equation 4

$$[CobNST] = \frac{[CobN] * [CobST]}{K_D + [CobST]}$$

It is evident that mutants E593A was non-functional from the lack of activity measured. It does seem that mutant R912A is somewhat active, however, a MM curve could not be produced from the data. A further investigation into what causes this reduced activity is described in Section 5.6.3.

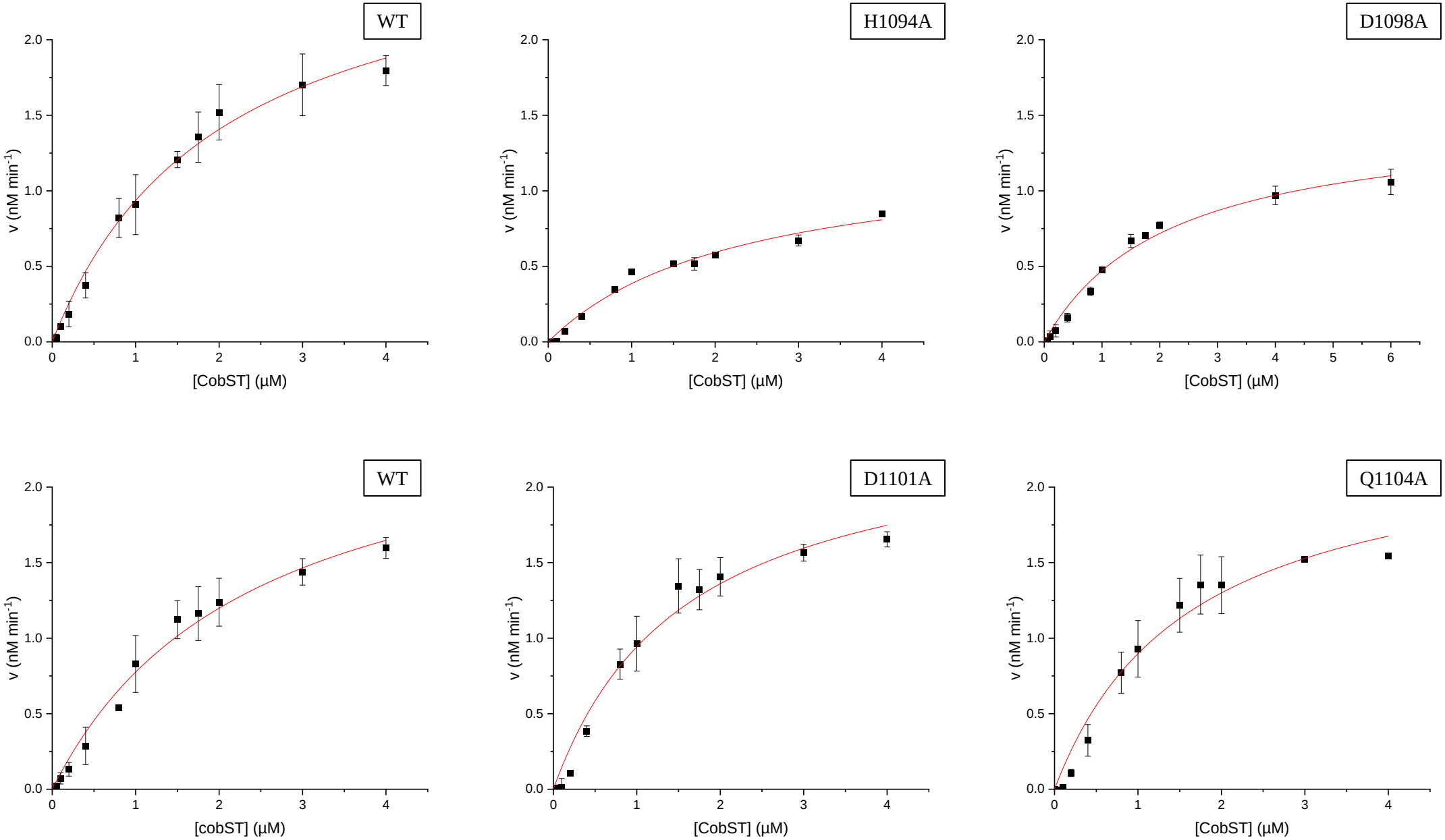


Figure 5.16 Binding curves of cobN(-mutants) to cobST. *A range of cobST was added to 0.04 µM of CobN and the reduction of fluorescent signal was measured in duplicate, indicating the chelation of HBAD. The wild-type cobN is compared to the mutants H1094A, D1098A, D1101A and Q1104A. Top and bottom row were enzymes expressed in two separate batches.*

The aforementioned K_D values produced by the fit of the plots are listed in Table 5.2. Here, we see that the binding of CobST to the wild-type vs the mutant CobNs are similar. It can be concluded that none of the mutations in CobN have a major effect on its conformation where the subunit interacts with CobST. A difference in V_{max}, however, is seen for mutants H1094 and D1098, the activity of which is ~ 30 % lower than for the wild-type CobN under these conditions. Based on the results obtained in this experiment, the concetration of CobST was determined which, in combination with 0.04 μM CobN(-variant), would form a fully functional cobaltochelatase. This enabled the determination of an apparent K_M curve for cobalt, which is described in Section 5.6.2.

Table 5.2

cobN	**K_D (μM)**	**Vmax (nM min^{-1})**
WT	2.0 ± 0.3	2.8 ± 0.2
H1094A	2.3 ± 0.5	1.3 ± 0.1
D1098A	2.2 ± 0.3	1.5 ± 0.1
WT	2.4 ± 0.4	2.6 ± 0.3
D1101	1.6 ± 0.3	2.4 ± 0.2
Q1104A	1.6 ± 0.4	2.4 ± 0.3

5.6.2 Michaelis-Menten curves for cobalt

To compare the activities of the wild-type cobaltochelatase to the various mutants, the catalytic rate was recorded as a function of the substrate concentration of cobalt. CobN was again fixed at 0.04 μM. CobST was added at 4 μM, roughly twice the K_D value for the various CobNs. Cobalt was ranged from 0 – 128 μM and the assays were performed in duplicate. The resulting change in fluorescence was converted to Δ[HBAD] min^{-1} as described previously. This was then divided by the total enzyme concentration used in each reaction as calculated using Equation 3. As such the value of V_{max} was divided by [E], and is therefore converted to the catalytic turnover number, k_{cat}. These plots were most accurately fitted to a substrate inhibition model (Equation 5), the graphs of which are seen in Figure 5.17. The kinetic parameters K_M, k_{cat}, k_{cat}/K_M as well as the inhibition constant K_i produced from these fits are listed in Table 5.3.

Equation 5

$$v = \frac{Vmax * [S]}{K_M * [S]\left(1 + \frac{[S]}{K_i}\right)}$$

Table 5.3 Kinetic parameters resulting from substrate inhibition curve. *Cobaltochelatase catalytic function of each CobN mutant was recorded as a function of the cobalt substrate concentration. The resulting plots fitted most accurately to a substrate inhibition model.*

cobN	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M (min^{-1} $μM^{-1}$)	K_i (μM)
WT	4.4 ± 0.5	0.08 ± 0.004	0.018 ± 0.006	87 ± 10
H1094A	3.3 ± 0.4	0.04 ± 0.002	0.012 ± 0.006	89 ± 12
D1098A	5.37 ± 0.95	0.06 ± 0.005	0.011 ± 0.015	80 ± 16
WT	6 ± 0.3	0.09 ± 0.003	0.015 ± 0.002	44 ± 3
D1101	6 ± 0.5	0.08 ± 0.004	0.013 ± 0.004	49 ± 5
Q1104A	5 ± 0.4	0.07 ± 0.003	0.014 ± 0.003	54 ± 4

The discrepancy between the values for the wild-type cobaltochelatase enzymes from the different batches is likely due to a difference in protein purity as well as how long the protein had been stored since its purification. As Covid-related work shifts prevented the ability to purify these proteins further via a gel-filtration column, slight differences in purity are expected. Additionally, depending on equipment availability, not every batch was tested for activity the same number of days after their purification. As this enzyme is not particularly stable, the freshness of the protein could have had an effect on the catalytic values measured. Taking this into regard, the mutants are compared to the wild-type enzymes in their own batch.

The catalytic turnover values indicate the mutation from histidine to alanine at position 1094 results in a 50% drop of activity of the overall enzyme, where mutant D1098 is 25% less active than the wild-type enzyme in this batch. These trends are not necessarily due to altered affinity to the substrate, cobalt, as we see from H1094A. This mutant has a lower K_M value even than the wild-type enzyme, resulting in a similar affinity (k_{cat}/K_M) for cobalt to the wild-type enzyme. With a higher K_M value, the affinity of D1098A for cobalt is significantly reduced, as can be seen from its lower k_{cat}/K_M value. It could be theorized that the lower catalytic function for the H1094A might therefore be due to another factor, such as a lower affinity for the substrate HBAD. Interestingly, the difference in affinity for cobalt did not result in a significant difference in inhibition by the substrate, as can be seen from the K_i values in this batch.

In the second set, the mutation of glutamine at 1104 to alanine is seen to result in a ~20% drop in catalytic activity, whereas the D1101 mutation – a residue that is part of the cobalt triad – seemed to result in a ~ 10% decrease. For Q1104, a lower K_M value is observed than seen with the other enzymes in this batch, however, the efficiency with which it turns over cobalt is reduced compared to the wild-type enzyme. Again, the lower catalytic function for this mutant could be due an increased binding affinity for the substrate where the intrinsic binding energy is manifested as tight binding rather than in increased catalytic turnover (Jencks, 1975).

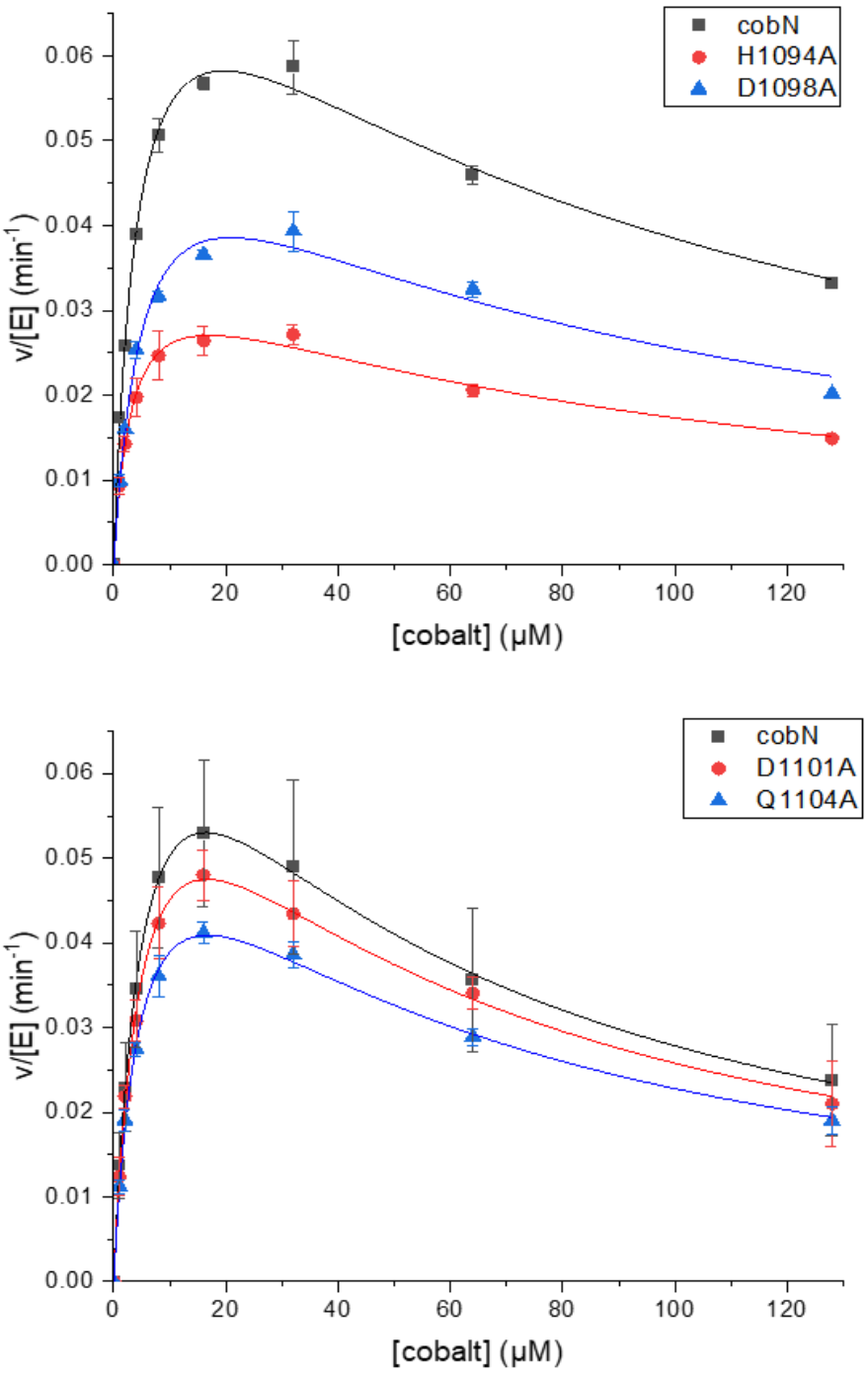


Figure 5.17 Cobaltochelatase . *The cobaltochelatase activity was measured in duplicate as function of substrate concentration of cobalt, which was ranged 0-64 μM. Velocity rates obtained by converting ΔFluorescence min -1 to Δ[HBAD] min-1 via a HBAD standard curve. Rates were divided by total enzyme concentration for each mutant as well as the wild-type for comparisons sake. Error bars indicate the SD of the values from the replicates.*

5.6.3 Binding assays HBAD-CobN (-mutants)

From the results obtained in the previous sections, it can be concluded that mutations of CobN at E593A and R912A render the enzyme inactive. E593 is part of the cobalt triad which could account for the enzyme's inactivity when mutated. As this is not the case for R912, an alternative possibility is this residue is involved with the binding of HBAD to the enzyme. A lower activity is also clear for the $CobN^{H1094A}$ and $CobN^{D1098A}$ variants. To test the affinity of these CobN-mutants for HBAD, different concentrations of CobN (0 - 1.6 µM) were mixed added to 0.5 µM HBAD. Volumes were made up to be equal with assay-buffer. The fluorescent signal was measured in triplicate.

Figure 5.18 shows the binding curves of the wild-type CobN to HBAD as well as the mutant variants. The binding of HBAD to CobN results in an increase of the fluorescent signal, which indicates that the binding of HBAD to the protein results in a greater delocalised state presumably through modulation of the double bond conformation. For all curves, a linear trend is seen up to ~ 0.5 µM, after which saturation is evident from the curve flattening. This point of saturation is shown by the intersection between the trend lines of the linear section and the horizontal section. Whilst there is some variation, the ratio of protein to HBAD looks to be 1:1. Under these circumstances, the change in activity seen previously, cannot be explained by a decreased ability to bind HBAD.

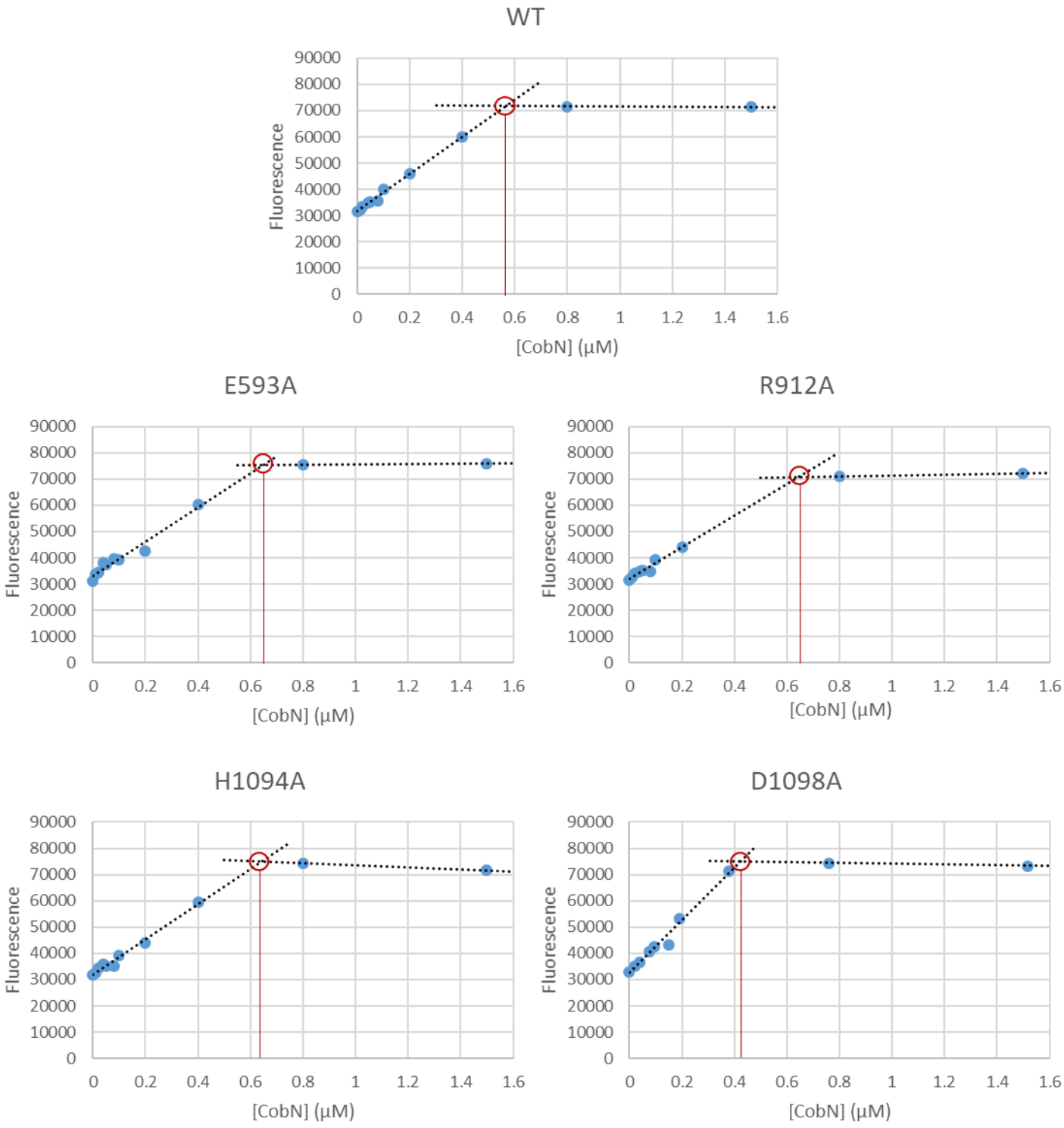


Figure 5.18 HBAD – CobN binding curves. *0.5 µM HBAD was added to a range of CobN (-mutants), and the effect of substrate-protein binding is seen by a rise in fluorescence signal.*

5.7 *In vivo* characterisation of CobN-mutants

Following the results of the mutations in the CobN subunit of the cobaltochelatase complex, the effect of these altered activities were studied *in vivo*. As mentioned previously, cobalt chelation of HBAD is one of the most challenging steps in the biosynthetic pathway of cobalamin. It was therefore theorised that implementing a cobaltochelatase-mutant with a higher catalytic activity into the best *E. coli* B_{12} producing strain, would result in an improved yield. Although a mutant with higher catalytic activity was not found, if indeed this step is a rate-limiting one in the pathway, a lower enzymatic function should then also be reflected in the B_{12} production. To change the gene in strain ED656-3B, the mutation would have to be made via CRISPR. Strain ED722A however, contained the entire pathway for cobalamin, excluding CobWNST. The deletion of the cobaltochelatase on the genome allowed for an easier insertion of the mutants via a plasmid.

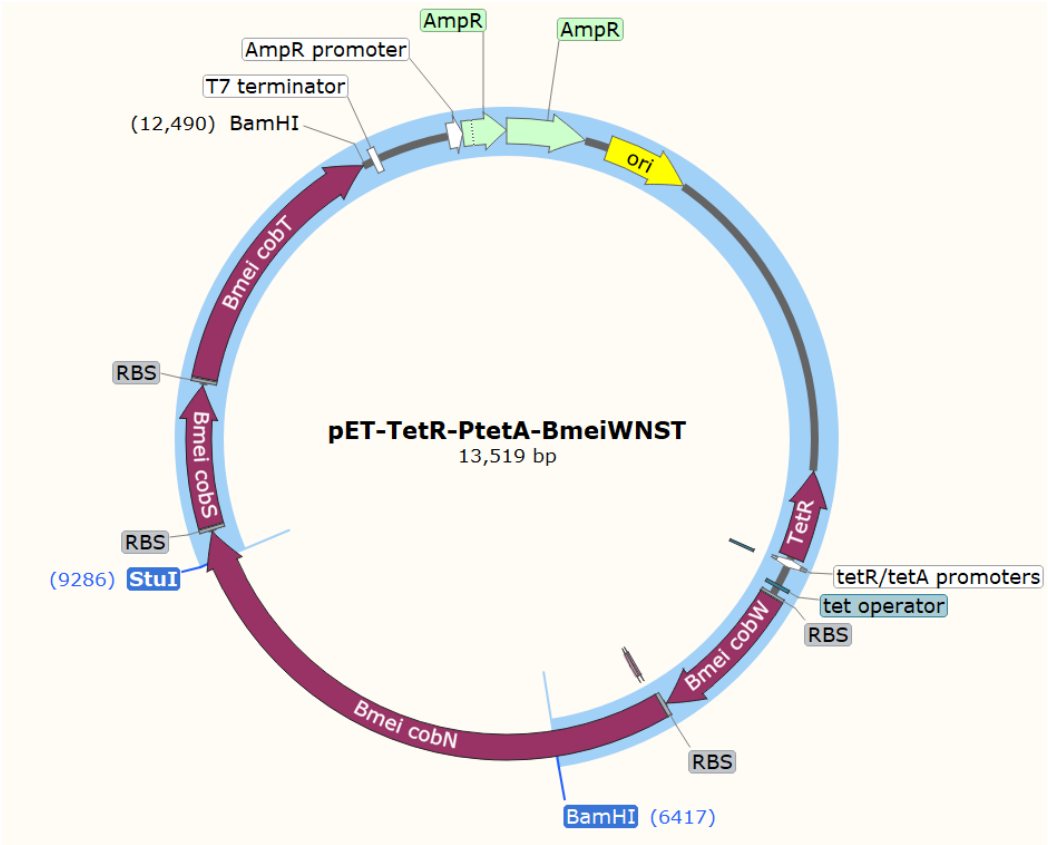


Figure 5.19 Plasmid map pf pET-TetR-PtetA-BmeiWNST. *Containing the cobaltochelatase mechanism from Brucella melitensis; cobN, cobS and cobT, as well as cobW, cobalt transporter protein. To obtain the desired vector (highlighted), a partial digest was made for BamHI, as well as a full digest for StuI.*

5.7.1 Cloning of pET-TetR-PtetA

pET-TetR-PtetA-BmeiWNST provided the vector for the mutated *cobN* variants. All *cobN* mutations made in Section 5.2.1 were located in the section between StuI and BamHI as shown in Figure 5.19. A straightforward restriction digest was made impossible by the lack of single-cutter restriction recognition sites on both sides – BamHI recognises another restriction site downstream from the *cobT* gene. To obtain the desired vector, therefore, a partial digest was done for BamHI, with a full StuI digest. To find the optimal digest, 5 reactions were incubated where StuI was added to 0.5 U/µl, and BamHI ranged at 0.06, 0.05, 0.025, 0.0125, and 0.005 U/ µL. The resulting band patterns are seen in Figure 5.20. A band of 10 kb is seen in the lanes which were digested with 0.06 and 0.05 U/µl BamHI, which corresponds with the size of the correct vector. This band was gel extracted and subsequently used as the vector for further cloning.

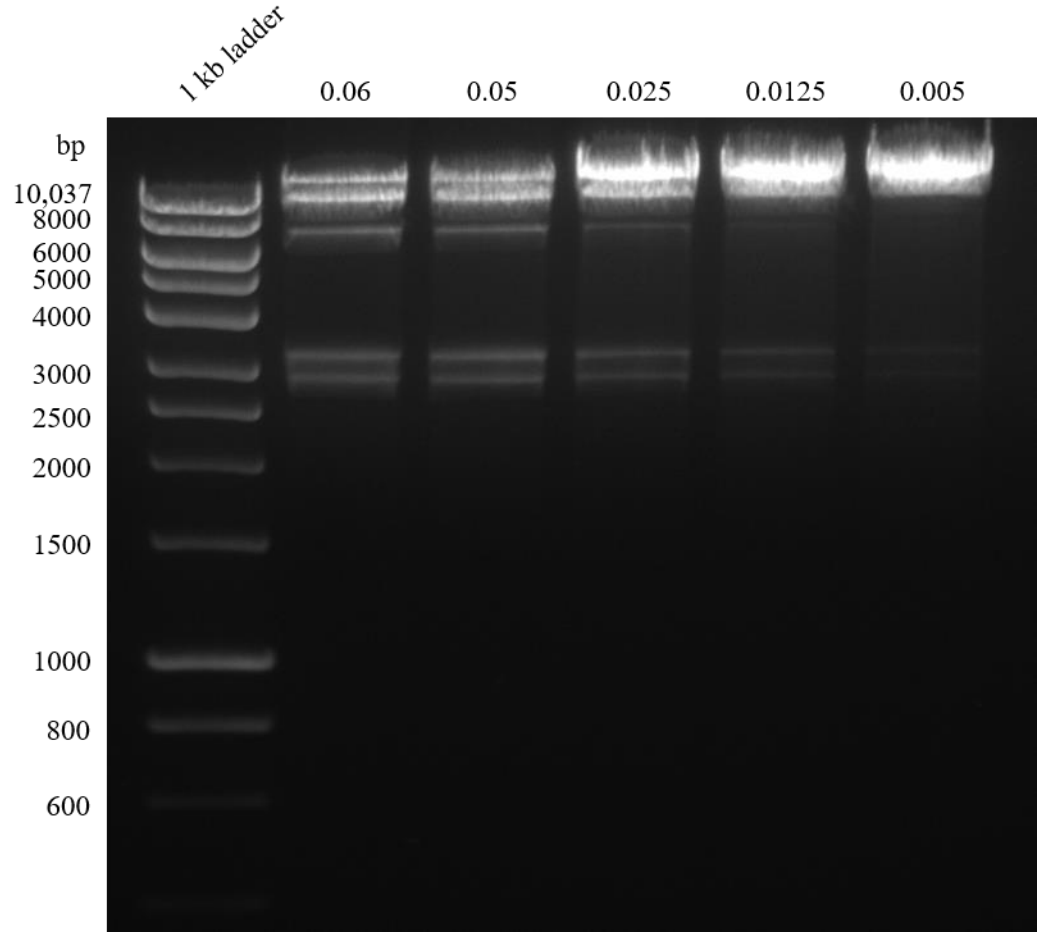


Figure 5.20. Band pattern of partial digest pET-TetR-pTetA-BmeiWNST on 1 % agarose gel. *Plasmid was digested fully with StuI (0.5 U/µl), and partially with BamHI (enzyme concentrations per digest given at the top of each lane). Electrophoresis of the gel was executed over 3 hours at 40V. A band corresponding with the size of the correct vector (~10 kb) is seen for 0.06 and 0.05 U/µl BamHI.*

The *cobN*-mutant inserts were digested from the plasmids pET14b - *Bmei* $^{6xHis}cobN^{E593A}$, pET14b - *Bmei* $^{6xHis}cobN^{D1098A}$, pET14b - *Bmei* $^{6xHis}cobN^{H1094A}$, pET14b - *Bmei* $^{6xHis}cobN^{Q1104}$, pET14b - *Bmei* $^{6xHis}cobN^{R912A}$, and pET14b - *Bmei* $^{6xHis}cobN^{D1101}$ (described in Section 5.3) with restriction enzymes StuI and BamHI. These inserts were then ligated into the digested pET-TetR-pTetA-BmeiW-ST vector isolated from the gel shown in Figure 5.20. The mutation insertion of $cobN^{R912A}$ was confirmed via sequencing. Cloning of the other mutants remained unsuccessful. It was realised that the band at ~10 kb could contain two possible partial digest; the desired vector as shown in Figure 5.19 as well as the highlighted section shown in Figure 5.21. To remove this second optional band in an attempt to successfully clone the rest of the mutant *cobN* genes into pET-TetR-pTetA-BmeiWNST, MfeI was added to the mix at 0.5 U/µl for a full digest. This enzyme cuts in the (unwanted) Bmei cobN gene on vector plasmid pET-TetR-pTetA-BmeiWNST, thereby eliminating the second band at 10 kb. Additionally, the gel electrophoresis time was extended to 24 hours, at 10V, to maximise the band separation. The resulting band pattern is shown in Figure 5.22.

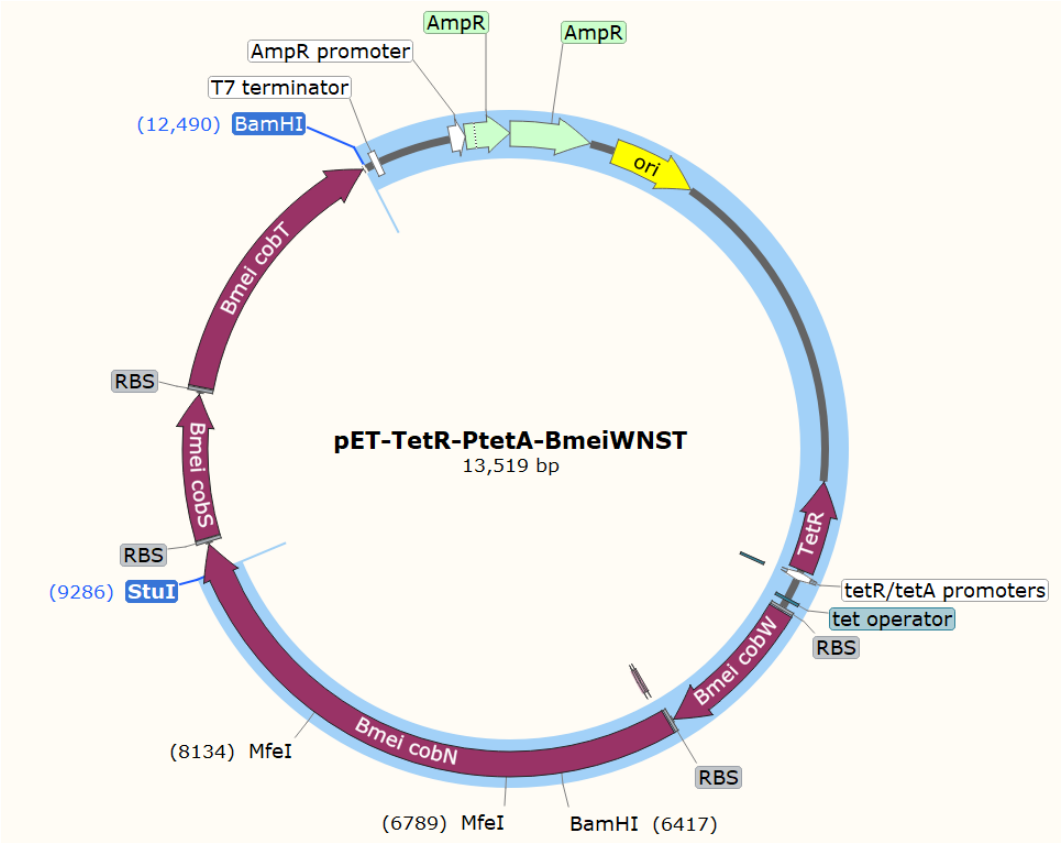


Figure 5.21 Optional digest of pET-TetR-PtetA-BmeiWNST when partially digested with BamHI and fully digested with StuI. *To remove this unwanted band, MfeI was added as an additional digest.*

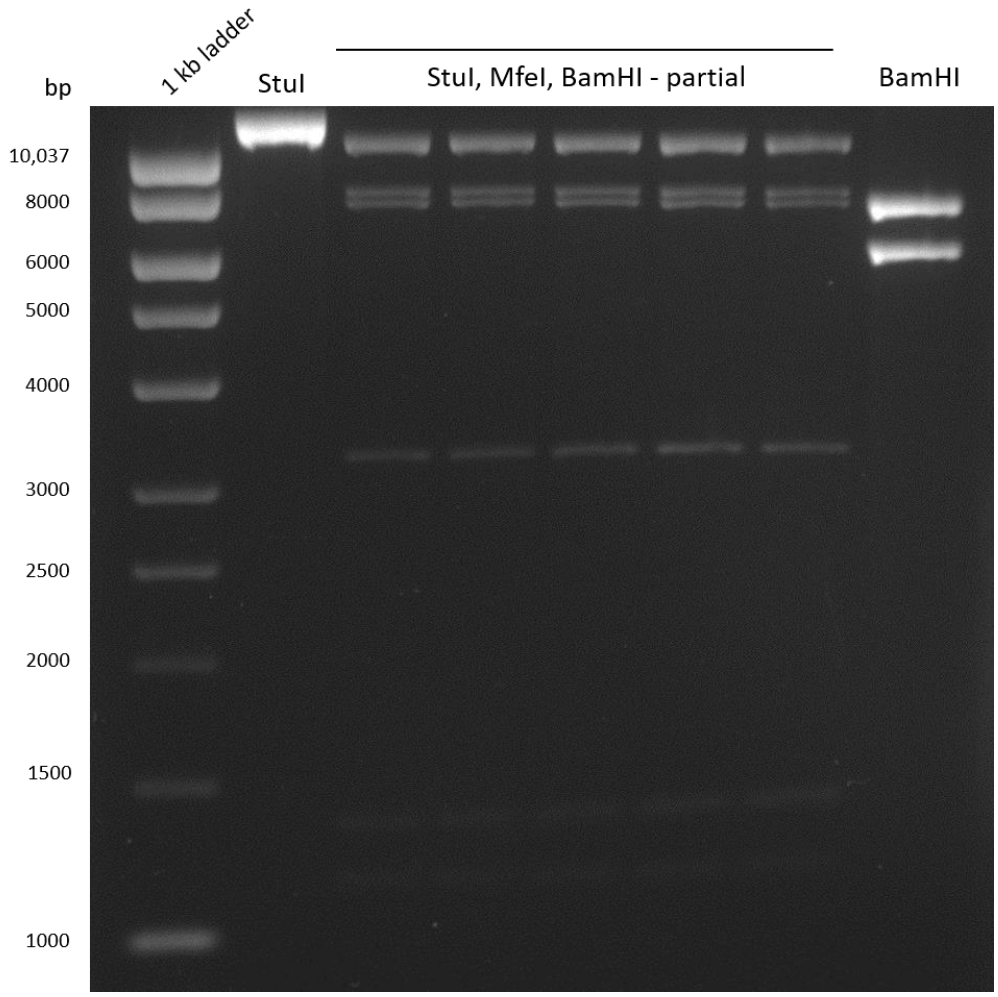


Figure 5.22 Band pattern of partial digest pET-TetR-pTetA-BmeiWNST on 1 % agarose gel. *Lane 2 contains the plasmid fully digested with StuI, lane 8 the fully digested plasmid with BamHI. The full digest with StuI and MfeI (0.5 U/µl), and partial digest with BamHI (0.1 U/µl) resulted in the band pattern seen in lanes 3-7. Electrophoresis of the gel was extended to 24 hours at 10V. No band is shown at the size corresponding to the vector (10 kb).*

Figure 5.22 shows the full digest of pET-tetR-pTetA-BmeiWNST with restriction enzymes StuI and MfeI in combination with a partial digest by restriction enzyme BamHI. Two samples of full digests by either StuI or BamHI were included as controls. These controls show full activity of both StuI and BamHI as pET-TetR-pTetA-BmeiWNST is cut in the expected size bands. The combination of full digests by StuI and MfeI with the partial digest by BamHI resulted in a band pattern which lacks the vector at ~10 kb. This unexpected result is difficult to explain as the enzymes StuI and BamHI appear to be fully functional from the controls. From the difference between the band patterns displayed in Figure 5.22, it is evident MfeI is also functional. Still, the desired fragment of 10 kDa is not produced. At this point, time-constraints restricted further progress of this project.

5.8 Discussion

An optimised method of purifying HBAD is described, which uses BtuG2 to trap HBAD and hydrogenobyrinic acid that has undergone varying degrees of amidation. A further combination of column purification steps generated pure HBAD as confirmed via HPLC-MS.

The cobaltochelatase subunit CobN from *Brucella melitensis* was successfully mutated to produce alanine substitutions at 4 different locations; E593, H1094, D1098 and Q1104. The remaining residue in the cobalt triad, H554, was not successfully mutated. The difficulty in mutating H554 could be due to the different approach in cloning, and would have been pursued further, had time allowed. The mutation design involves a two-step overlap extension PCR approach that is described in Section 5.2.1. The CobN mutants were expressed and purified separately from the CobST complex, and thus, the formation of fully functional enzyme was investigated via MM-style binding studies where CobN was fixed and CobST varied, allowing the latter complex to act as the 'substrate'. The resulting activity curves generated apparent K_D values from which the total enzyme concentration was calculated for subsequent cobalt catalytic turnover assays. These K_D curves highlighted that cobalt demonstrated substrate inhibition at higher concentrations. This is an important point as to make large amounts of vitamin B_{12}, the cell requires higher levels of cobalt to be added to the cell. Thus, cobalt may actually limit the ability of the cell to produce the nutrient not through its scarcity but through its over-abundance. Hence, there is good reason to study and modulate the availability of this central metal core of B_{12}.

From the binding curves, a difference in activity between some of the mutants became evident. The raw data (Figure 5.12), shows that E593 is vital for chelation of HBAD in *B. melitensis* CobN as the mutation of the amino acid to alanine renders the enzyme as a whole inactive. This preserved residue was identified to be one of the 'cobalt triad' in the CobN subunit of *Mycobacterium tuberculosis* (J.-H. Zhang et al., 2021). The second mutant of the cobalt triad which was tested was D1101. As seen from the kinetic parameters obtained from the cobalt curve, this mutation caused a 10% drop in catalytic activity, when compared to the wild-type enzyme. The improvement of catalytic efficiency which is described for the magnesium chelatase (Adams et al., 2020) caused by this mutation is not seen in *Bmei* CobN. In their research, the increased catalytic efficiency was likely due to a lower affinity for the substrate and product. As product-release is thought to be the rate-limiting factor of this reaction, a lower affinity would be preferable. This effect is not seen for any of the mutations tested here. The mutations implemented in *B. melitensis* CobN which were made have not succeeded in achieving the higher turnover they were sought.

Following the *in vitro* characterisation of the cobaltochelatase mutants, the difference in activity was set out to be studied *in vivo*. One mutant was successfully cloned, resulting in pET-TetR-PtetA-*Bmei*WNR^{912A}ST, however, inserting the cobaltochelatase mutant into strain ED722A to test the difference in cobalamin production was not achieved due to time limitations.

Chapter 6

Discussion

The majority of vitamin B_{12} production has been restricted to China for the past couple decades, which has caused the price to fluctuate greatly as manufacture has varied due to factory closures over safety concerns. With a strong global demand for the nutrient from combined applications in human health, livestock farming, animal welfare and biotechnology, the need for access to sensibly priced vitamin B_{12} has been increasing. Entering a competitive market requires a novel production method at distinctly lower costs. While classic production methods have relied on optimising organisms that contain the cobalamin biosynthetic pathway, recently attention has steered towards production via alternative organisms as a way of overcoming the limits presented in native strains. Research performed by other groups have led to cobalamin-producing *E. coli* strains with yields of 307.00 μg g^{-1} cdw and 0.65 μg g^{-1} cdw (Fang et al., 2018b; Ko et al., 2014). In the Warren lab, a cobalamin-producing *E. coli* strain was engineered that produces a respectable 40 mg/L which is ~2 mg g^{-1} cdw (using conversion values of 0.39 g L^{-1} cdw for *E. coli* per one unit OD_{600} (Glazyrina et al., 2010)). To reach market potential, however, production levels of 1 g/L would provide a strain that far outstrips all current production strains. This thesis describes three ways in which strain optimisation was approached. The first approach involves relieving metabolic stress by switching the promoters controlling operons holding the B_{12} biosynthetic pathways from a strong T7 promoter to weaker, constitutive promoters. The second describes methods to improve metabolic flux within the biosynthetic pathway as well as points feeding into the pathway by improving essential secondary metabolic processes linked to B_{12} biosynthesis. Finally, the current B_{12}-producing strain exhibits a metabolic pinch-point in the pathway with the accumulation of HBAD, the last intermediate before cobalt insertion, despite the strain encoding two sets of cobaltochelatases on the genome. Therefore, the most active of the two cobaltochelatases - responsible for the crucial step of cobalt insertion - was investigated in order to try and improve its catalytic turnover.

The first approach attempted to alleviate the high metabolic stress of producing cobalamin via its 20+ step biosynthetic pathway under the control of a strong T7 promoter. Induction of the T7 promoter initiates an approximate 9-hour metabolic race for survival, since the T7 promoter system effectively prevents normal cellular transcription that would result in cell death. The cell therefore selects against the T7 system and the resulting cells will therefore not produce B_{12}. The benefit of switching to weaker, constitutive promoters was thought to be two-fold; as the metabolic strain lessens, the organism could allow for a longer production phase; secondly, a lack of inductive agent is preferable with an eye on up-scaling the process for industrial production as this lowers the production costs. A library of promoters of varying strengths to obtain optimal growth and product formation

phenotypes has been previously described and is available (Alper et al., 2005; *Promoters/Catalog/Anderson - Parts.Igem.Org*, n.d.). From this panel a strong (A2), medium (A3) and weak (M2) promoter were identified. The B_{12}-producing strain ED656-3B was modified by replacement of the T7 promoter with one of these constitutive promoters in the A- site, B- site, or both, in various combinations, via CRISPR-Cas9. The resulting strains displayed a distinct decline in B_{12} yield when the A-site T7 promoter was changed to any constitutive promoter. As this operon holds all the genes of the B_{12} biosynthetic pathway, this effect could be due to lower levels of the encoded enzymes from the pathway being produced. This is supported by the lower levels of β-galactosidase activity, enabled by a *lacZ* gene at the end of this operon. Lower yields could also be caused by the constant burden of such an energy-intensive pathway operating in as early as the lag phase even if the overall production is lower. In contrast, under the control of the T7 promoter, the cells can reach exponential phase before burdening the metabolic system. However, despite the observation that strains that contain a constitutive promoter in the A-site produce lower B_{12} levels than those whose A-site operon is under the control of the T7 promoter, it is also observed that those strains with an A-site constitutive promoter show an increase in β-galactosidase activity in the time period of 24 to 48 hours, whereas the A-site T7 strains show a decrease. This indicates that expression of the genes of the B_{12} pathway does continue longer when under the control of a constitutive promoter when these strains are cultured in flasks. However, it does not translate into elongation of the production phase, as cobalamin levels did not continue to climb after the 24-hour time point. Testing mRNA levels as well as the stability of these transcripts could give further insight into how these results are linked. This could be pursued by qPCR. As the A-site operon holds over 20 genes, the mRNA transcript is a substantial product. Depending on the stability of this mRNA, the enzymes that are encoded by the genes at the end of the operon could be expressed to a much lesser extent in comparison to those encoded for by those genes at the start of the operon. Balancing the expression levels of all the enzymes in the B_{12} pathway in relation to each other remains unexplored by this approach, and is definitely worth exploring further. This could be attempted by changing the order of the genes within the A-site operon, or changing some of the ribosome binding sites that precede the individual genes. Changes of the T7 promoter to a constitutive promoter in the B-site has a much lower effect that the change in the A-site, as cobalamin production levels remain similar to the original strain. Increasing the copy number of various genes may also prove beneficial, and a tactic that has been used for instance in the production of antibiotics (Hopwood & Chater, 1980).

The second approach towards production improvements of strain ED656-3B was to improve key steps in the biosynthetic pathway of B_{12} as well as secondary metabolites linked to the pathway. As mentioned above, the previous method targeted the operons as a whole, without looking into the effect of individual genes within the pathway. Here, genes of interest were selected, based on their position within the B_{12} pathway, and added to the B_{12}-producing *E. coli* strain ED656-3B through their inclusion on vectors. The genes selected included *Rc hemA*^{C4}, *Mbar cobA*, *Pd cobA*^{K187A}, *Bmei cobI*, *ySAMS*, *hSAMS*, and *rat SAHH*. These genes encode enzymes that are either involved directly in the early B_{12} biosynthetic pathway, or feed into the pathway through the provision or removal of certain metabolites. HemA^{C4} catalyses the synthesis of 5-ALA, the first precursor in the tetrapyrrole pathway. 5-ALA synthesis is highly regulated and is shown to be one of the rate limiting steps in tetrapyrrole biogenesis (Lascelles, 1978). As *E. coli* inherently produces 5-ALA via the C5 pathway, overexpression of *Rc* HemA^{C4} enables strain ED656-3B to produce 5-ALA via the C4 pathway as well as the C5 pathway. This has proven beneficial to 5-ALA yields in *E. coli* as overexpression of *R. sphaeroides* HemA resulted in a 5-fold increase (Werf & Zeikus, 1996). CobA catalyses the reaction resulting in precorrin-2 (Blanche et al., 1989), the last common precursor for several tetrapyrroles (Figure 1.3) before committing to the B_{12} biosynthetic pathway in the next step, catalysed by CobI (Thibaut et al., 1990). A change of a lysine at position 187 to an alanine in CobA from *P. denitrificans* cured this enzyme of substrate inhibition. Overexpressing these steps could direct flow into the cobalamin pathway and thereby increase cobalamin yield. These first 3 out of 8 methylation steps, all involve the transfer of methyl groups from S-adenosyl methionine (SAM) to the central corrin ring (Figure 1.3). Increasing SAM synthesis would therefore provide more substrate to be fed into the B_{12} pathway. This was attempted by increasing SAM synthesis as well as S-adenosylhomocysteine hydrolase. The latter reaction loops back into the methionine recovery pathway (Matthews et al., 2008) further providing SAM substrate (Figure 4.1).

Expression of the aforementioned genes in ED656-3B under the control of a T7 promoter confirmed that the metabolic system is burdened to the point of negative phenotypical results. Growth rate is often used as a measure of fitness costs associated with overexpressing heterologous genes (Pope et al., 2010). Here, the decreased fitness can be seen as decreased cell growth as is seen when *P. denitrificans* CobA^{K187A} is overexpressed, or even cell death, as is seen when ED656-3B is transformed with *M. barkeri cobA,* despite the T7 expression being supressed by addition of glucose to the plate media. More significantly, it is measured by diminished B_{12} levels produced by the strain when overexpressing any of the genes. The effect of expression of any of these genes when

incorporated into the B_{12} pathway is poorly reflected by T7 expression levels as these are much higher than would result from single genomic transcription. The genes were therefore placed under the control of the TetR-PtetA promoter system, which is more tightly regulated by the levels of inducer. In this system, ED656-3B was transformed with *M. barkeri cobA*, and expression of the enzyme resulted in a two-fold increase in cobalamin production, indicating this step in the biosynthetic pathway is rate limiting. Additionally, expression of HemA from *R. capsulatus* and rat *SAH hydrolase* increased cobalamin yields by 30%. Strain improvement could therefore be achieved by integrating these genes into the genome of ED656-3B. Disappointingly, attempts at CRISPR cloning for the integration of *M. barkeri cobA* into the B_{12} biosynthetic pathway of ED656-3B proved unsuccessful. Recent method developments have resulted in a CRISPR cloning approach with a higher success rate which could be employed. However, integration of the *M. barkeri cobA* was attempted by switching out the existing *R. capsulatus cobA* in the biosynthetic pathway, which is located directly at the 3' end of the T7 promoter. Although this is genomic, it might still result in the same lethal effect as was oberved when the gene was added on a plasmid without induction. For future work, the *M. barkeri cobA* could be incorporated into the genome using the new CRISPR cloning method, as well as using ribosome binding sites of different strength to adjust the optimal expression of CobA for B_{12} production. The next step would be the integration of *R. capsulatus hemA* and rat *SAH hydrolase* for further strain improvement.

Finally, the cobalt insertion step of the biosynthetic pathway was investigated as a means to increase B_{12} yields. The enzyme that catalyses the step is slow and intermediate build-up shows it is a crucial and rate-limiting part of the pathway that needs improving. The cobaltochelatase CobNST of *B. melitensis*, which is located in the B-site operon of ED656-3B, was characterised as well as some mutant variants of the CobN subunit.

Research into the cobaltochelatase from *M. tuberculosis* highlighted three residues involved in cobalt binding in the active site of the CobN subunit, thus dubbed the 'cobalt triad' (J.-H. Zhang et al., 2021). The equivalent amino acids were located in the *B. melitensis* CobN to be H554, E593 and D1101. Cloning of *B. melitensis* $cobN^{H554A}$ was not achieved, and should be pursued further. The other two variants were successfully cloned and characterised. Interestingly, the mutation of glutamic acid to alanine at position 593 rendered an enzyme variant that has no activity and is unable to chelate the substrate HBAD with cobalt. Whilst such a dramatic effect was not observed for $CobN^{D1101A}$, it did show a decrease in catalytic activity by 10%. These results suggest both these residues are vital for catalytic turnover in the *B. melitensis* CobN.

Another mutant was inspired by research into the related magnesium chelatase enzyme. More research has been done into this enzyme which is structurally similar to the cobaltochelatase, and therefore represents a useful model. When changing D1177 in the ChlH subunit – the magnesium chelatase equivalent to CobN in the cobaltochelatase –an increased catalytic turnover was reported (Adams et al., 2020). The corresponding residue in the *B. melitensis* CobN was found to be a glutamine at position 1104 and a $cobN^{Q1104A}$ mutant was cloned. Unfortunately, characterisation of this mutant did not see the increase in catalytic number, but rather a decrease by ~20%. This lower catalytic function could be due to an increased binding affinity for the substrate as a lower K_M value is measured. This indicates the intrinsic binding energy manifests itself as tight binding rather than in increased catalytic turnover (Jencks, 1975).

Further mutations were designed based on a partial crystal structure of the *B. melitensis* CobN binding pocket with its substrate HBAD produced in the Warren lab, which highlighted residues involved in HBAD binding. As product release is thought to be the rate limiting factor, these residues, R912, H1094 and D1098, were also investigated. When characterised, the mutation of the arginine at 912 to an alanine is shown to result in a significant drop in activity – where generating a Michaelis-Menten curve was impossible. $CobN^{D1098A}$ was found to be 25% less active than the wild-type enzyme and a 50% drop in activity is seen after mutation of H1094 to an alanine. D1098A shows a decreased affinity for cobalt, which could explain its lower catalytic turnover, however, this is not the case for H1094A. A reduction in binding affinity of the other substrate, HBAD, was investigated to explain the reduction in activity for H1094A and R912A. HBAD binding curves indicate, however, that the binding of HBAD is not diminished. This leaves the explanation that the binding is altered to result in non-productive binding, where the conformational change does not support catalytic turnover. It would be very interesting to obtain a Michaelis-Menten curves for HBAD. As this substrate gives the signal by which the reaction is followed, however, this is not possible in the current method.

For each reaction, wild-type or mutant, the reaction trajectory showed a lag phase of approximately 6 minutes before speeding up, which could indicate a conformational change taking place to activate CobN. However, there are many factors at play in this reaction, such as the combining of the subunits to form a fully functional enzyme as well as the ATP hydrolysis taking place in the CobST-subunits. Elucidating the relationship between ATP hydrolysis and the cobalt chelation reactions should be pursued further.

The Michaelis-Menten curves for cobalt revealed a substrate inhibition at higher concentrations of cobalt for the wild-type CobN as well as all mutants. This is interesting as high levels of cobalt have been added to the media to boost B_{12} production. However, these results indicate that this over-abundance of cobalt may limit the cell's ability to produce cobalamin, rather than enhance it. This, as well as the negative impact of cobalt on the iron-sulfur cluster-containing CobG (Ranquet et al., 2007), highlights the need to study and modulate the availability of this cobalt ion that forms the central metal core of B_{12}.

Overall, future work should focus on characterising the entire vitamin B_{12} pathway within the ED656-3B. The approaches taken here to improve the cobalamin yield were designed rationally, but also subject to available materials in the lab. Further investigation of the pathway through quantifying mRNA levels as well as pathway-specific enzymes' half-life could identify more specific approaches or targets for improved cobalamin yield. These mRNA levels could be quantified via qPCR and would give insight into the transcription levels when under the control of the different types of promoters. Additionally, designing different primer sets at the beginning, middle and end of the pathway would show the stability of the transcripts. When combined with more in-depth proteomic – in the form of 2D SDS PAGE combined with MALDI-TOF - would enable the identification as well as quantification of the corresponding expressed enzymes of the pathway. Such information would give insight into how the different expression systems truly affect B_{12} synthesis as well as potentially identify additional bottlenecks. Furthermore, the work described in chapter 4 identifies the addition of M. barkeri cobA to the strain via a vector to be a very promising route towards increasing B_{12} yields of the *E. coli* strain. Replacing or even just adding this gene to the genome of ED656-3B via the beforementioned new approach of CRISPR-Cas9 (page 151) is strongly advised. Finally, the catalytic mechanism of CobNST should be characterised more fully, as the rate limiting step of product release is not experimentally proven. Elucidation of the reaction rates could provide insight into how these are ordered, and which is rate-limiting. Additionally, it would be valuable to gain insight into the order and relation between the ATP hydrolysis reaction taking place in CobST and the cobalt chelation of HBAD taking place in CobN.

6.1 Conclusions

The approaches towards increasing B_{12} production in engineered *E. coli* strain ED656-3B have highlighted the importance of regulation of gene expression in the interest of increased cell fitness. T7 controlled expression of the genes of the B_{12} pathway is effective, but strains the metabolic system, which manifests as a 9-hour clock starting from the moment of induction, at the end of which expression is stopped. When this entire pathway is placed under the control of constitutive promoters of varying strengths, this expression time frame does seem to widen, however, this does not translate in higher B_{12} yields. This straining effect of T7-controlled expression on the desired phenotype is further manifested in the second approach where transformation with *M. barkeri cobA* leads to cell toxicity. When placed under a more tightly controlled TetR-PtetA promoter system however, expression of this gene increased the B_{12} yield of ED656-3B by two-fold. Balancing between driving the cell to maximum performance and keeping the metabolic system running in a sustainable manner is shown to be key.

Furthermore, a key step in the biosynthesis of B_{12} is catalysed by cobaltochelatase. Strain ED656-3B holds two copies of this enzyme. The more active of the two, originating from *B. melitensis*, shows substrate inhibition for cobalt indicating intracellular cobalt levels should be monitored for optimised production. Mutations in cobalt and HBAD binding site have negative impact on catalytic turnover, and highlight how much more can be learned about the manner in which this enzyme operates.

In a day and age where we are confronted with climate change we are more and more faced with the finding alternatives to methods of production which have served us well thus far. This has given the rise to an exciting branch of science in the form of 'synthetic biology'; a field in which biological systems are used and adapted to produce valuable chemicals in a more sustainable manner. Key words such as 'metabolic engineering' and 'cell-factories' are applied to create idyllic pictures of a green future. And whilst there are certainly a number of tremendous advances being made, the simplified presentation of setting up such mechanisms often involves a lot of trial and error as the rules associated with the engineering of complex systems remain poorly defined and understood. Our understanding of metabolic systems is vast, however, it is not all-encompassing. This often shows in rational design not approaching the predicted results, and serendipitous results being ground-breaking. This does not lead to the conclusion that these approaches are therefore not worth-while, as there are many success-stories around. There is a lot still left to learn about the biological systems and how they can be manipulated for useful purposes. This is

exactly what is seen in the body of work presented in this thesis. Most approaches of rational design did not lead to the ultimate aim of a higher productivity in our B_{12}-producing strain, however, a lot of knowledge has been gained, leading to further, more informed choices of how to proceed towards this aim.

References

Adams, N. B. P., Bisson, C., Brindley, A. A., Farmer, D. A., Davison, P. A., Reid, J. D., & Hunter, C. N. (2020). The active site of magnesium chelatase. *Nature Plants*, *6*(12), 1491–1502. https://doi.org/10.1038/s41477-020-00806-9

Allen, L. H., Miller, J. W., de Groot, L., Rosenberg, I. H., Smith, A. D., Refsum, H., & Raiten, D. J. (2018). Biomarkers of Nutrition for Development (BOND): Vitamin B-12 Review. *The Journal of Nutrition*, *148*(suppl_4), 1995S-2027S. https://doi.org/10.1093/jn/nxy201

Allen, R. H., Seetharam, B., Podell, E., & Alpers, D. H. (1978). Effect of Proteolytic Enzymes on the Binding of Cobalamin to R Protein and Intrinsic Factor. *Journal of Clinical Investigation*, *61*(1), 47–54. https://doi.org/10.1172/JCI108924

Alper, H., Fischer, C., Nevoigt, E., & Stephanopoulos, G. (2005). Tuning genetic control through promoter engineering. *Proceedings of the National Academy of Sciences*, *102*(36), 12678–12683. https://doi.org/10.1073/pnas.0504604102

Astner, I., Schulze, J., Heuvel, J., van der, Jahn, D., Schubert, W.-D., & Heinz, D. W. (2005). Crystal structure of 5-aminolevulinate synthase, the first enzyme of heme biosynthesis, and its link to XLSA in humans. *The EMBO Journal*, *24*(18), 3166–3177. https://doi.org/10.1038/sj.emboj.7600792

Banerjee, R. (1997). The Yin-Yang of cobalamin biochemistry. *Chemistry & Biology*, *4*(3), 175–186. https://doi.org/10.1016/S1074-5521(97)90286-6

Banerjee, R. (1999). *Chemistry and Biochemistry of B12*. John Wiley & Sons.

Banerjee, R., & Ragsdale, S. W. (2003). The Many Faces of Vitamin B $_{12}$: Catalysis by Cobalamin-Dependent Enzymes. *Annual Review of Biochemistry*, *72*(1), 209–247. https://doi.org/10.1146/annurev.biochem.72.121801.161828

Barker, H. A., Smyth, R. D., Weissbach, H., Munch-Petersen, A., Toohey, J. I., Ladd, J. N., Volcani, B. E., & Wilson, R. M. (1960). Assay, purification, and properties of the adenylocobamide coenzyme. *The Journal of Biological Chemistry*, *235*(1), 181–190. https://doi.org/10.1016/S0021-9258(18)69607-3

Barker, H. A., Weissbach, H., & Smyth, R. D. (1958). A coenzyme containing pseudovitamin B12. *Proceedings of the National Academy of Sciences of the United States of America*, *44*(11), 1093–1097. https://doi.org/10.1073/pnas.44.11.1093

Battersby, A. R. (1994). How nature builds the pigments of life: The conquest of vitamin B12. *Science*, *264*(5165), 1551–1557. https://doi.org/10.1126/science.8202709

Battersby, A. R. (2007). B $_{12}$ -Biosynthesis in an Aerobic Organism: How the Pathway was Elucidated. In B. Kräutler, D. Arigoni, & B. T. Golding (Eds.), *Vitamin B 12 and B 12 -Proteins* (pp. 47–61). Wiley-VCH Verlag GmbH. https://doi.org/10.1002/9783527612192.ch02

Battersby, A. R., Fookes, C. J. R., Gustafson-Potter, K. E., McDonald, E., & Matcham, G. W. J. (1982). Biosynthesis of porphyries and related macrocycles. Part 18. Proof by spectroscopy and synthesis that unrearranged hydroxymethylbilane is the product from deaminase and the substrate for cosynthetase in the biosynthesis of uroporphyrinogen-III. *Journal of the Chemical Society, Perkin Transactions 1*, *0*, 2427–2444. https://doi.org/10.1039/P19820002427

Beale, S. I. (1990). Biosynthesis of the Tetrapyrrole Pigment Precursor, δ-Aminolevulinic Acid, from Glutamate 1. *Plant Physiology*, *93*(4), 1273–1279. https://doi.org/10.1104/pp.93.4.1273

Bentley, W. E., Mirjalili, N., Andersen, D. C., Davis, R. H., & Kompala, D. S. (1990). Plasmid-encoded protein: The principal factor in the "metabolic burden" associated with recombinant bacteria. *Biotechnology and Bioengineering*, *35*(7), 668–681. https://doi.org/10.1002/bit.260350704

Blanche, F., Cameron, B., Crouzet, J., Debussche, L., Levy-Schil, S., & Thibaut, D. (1998). *Polypeptides involved in the biosynthesis of cobalamines and/or cobamides, dna sequences coding for these polypeptides, and their preparation and use* (European Union Patent No. EP0516647B1). https://patents.google.com/patent/EP0516647B1/en

Blanche, F., Cameron, B., Crouzet, J., Debussche, L., Thibaut, D., Vuilhorgne, M., Leeper, F. J., & Battersby, A. R. (1995). Vitamin B12: How the Problem of Its Biosynthesis Was Solved. *Angewandte Chemie International Edition in English*, *34*(4), 383–411. https://doi.org/10.1002/anie.199503831

Blanche, F., Couder, M., Debussche, L., Thibaut, D., Cameron, B., & Crouzet, J. (1991). Biosynthesis of vitamin B12: Stepwise amidation of carboxyl groups b, d, e, and g of cobyrinic acid a,c-diamide is catalyzed by one enzyme in Pseudomonas denitrificans.

Journal of Bacteriology, *173*(19), 6046–6051. https://doi.org/10.1128/jb.173.19.6046-6051.1991

Blanche, F., Debussche, L., Famechon, A., Thibaut, D., Cameron, B., & Crouzet, J. (1991). A bifunctional protein from Pseudomonas denitrificans carries cobinamide kinase and cobinamide phosphate guanylyltransferase activities. *Journal of Bacteriology*, *173*(19), 6052–6057. https://doi.org/10.1128/jb.173.19.6052-6057.1991

Blanche, F., Debussche, L., Thibaut, D., Crouzet, J., & Cameron, B. (1989). Purification and characterization of S-adenosyl-L-methionine: Uroporphyrinogen III methyltransferase from Pseudomonas denitrificans. *Journal of Bacteriology*, *171*(8), 4222–4231. https://doi.org/10.1128/jb.171.8.4222-4231.1989

Blanche, F., Famechon, A., Thibaut, D., Debussche, L., Cameron, B., & Crouzet, J. (1992). Biosynthesis of vitamin B12 in Pseudomonas denitrificans: The biosynthetic sequence from precorrin-6y to precorrin-8x is catalyzed by the cobL gene product. *Journal of Bacteriology*, *174*(3), 1050–1052. https://doi.org/10.1128/jb.174.3.1050-1052.1992

Blanche, F., Maton, L., Debussche, L., & Thibaut, D. (1992). Purification and characterization of Cob(II)yrinic acid a,c-diamide reductase from Pseudomonas denitrificans. *Journal of Bacteriology*, *174*(22), 7452–7454. https://doi.org/10.1128/jb.174.22.7452-7454.1992

Blanche, F., Thibaut, D., Couder, M., & Muller, J.-C. (1990). Identification and quantitation of corrinoid precursors of cobalamin from Pseudomonas denitrificans by high-performance liquid chromatography. *Analytical Biochemistry*, *189*(1), 24–29. https://doi.org/10.1016/0003-2697(90)90038-B

Blanche, F., Thibaut, D., Famechon, A., Debussche, L., Cameron, B., & Crouzet, J. (1992). Precorrin-6x reductase from Pseudomonas denitrificans: Purification and characterization of the enzyme and identification of the structural gene. *Journal of Bacteriology*, *174*(3), 1036–1042. https://doi.org/10.1128/jb.174.3.1036-1042.1992

Bommer, M., Kunze, C., Fesseler, J., Schubert, T., Diekert, G., & Dobbek, H. (2014). Structural basis for organohalide respiration. *Science*, *346*(6208), 455–458. https://doi.org/10.1126/science.1258118

Bridwell-Rabb, J., & Drennan, C. L. (2017). Vitamin B 12 in the spotlight again. *Current Opinion in Chemical Biology*, *37*, 63–70. https://doi.org/10.1016/j.cbpa.2017.01.013

Bridwell-Rabb, J., Zhong, A., Sun, H. G., Drennan, C. L., & Liu, H. (2017). A B 12 -dependent radical SAM enzyme involved in oxetanocin A biosynthesis. *Nature*, *544*(7650), 322–326. https://doi.org/10.1038/nature21689

Bryant, D. A., Hunter, C. N., & Warren, M. J. (2020). Biosynthesis of the modified tetrapyrroles—The pigments of life. *Journal of Biological Chemistry*, *295*(20), 6888–6925. https://doi.org/10.1074/jbc.REV120.006194

Bunbury, F., Deery, E., Sayer, A., Bhardwaj, V., Harrison, E., Warren, M. J., & Smith, A. G. (2022). *Exploring the onset of B12-based mutualisms using a recently evolved Chlamydomonas auxotroph and B12-producing bacteria* (p. 2022.01.04.474942). bioRxiv. https://doi.org/10.1101/2022.01.04.474942

Cameron, B., Blanche, F., Rouyez, M. C., Bisch, D., Famechon, A., Couder, M., Cauchois, L., Thibaut, D., Debussche, L., & Crouzet, J. (1991). Genetic analysis, nucleotide sequence, and products of two Pseudomonas denitrificans cob genes encoding nicotinate-nucleotide: Dimethylbenzimidazole phosphoribosyltransferase and cobalamin (5'-phosphate) synthase. *Journal of Bacteriology*, *173*(19), 6066–6073. https://doi.org/10.1128/jb.173.19.6066-6073.1991

Carmel, R. (2000). Current Concepts in Cobalamin Deficiency. *Annual Review of Medicine*, *51*(1), 357–375. https://doi.org/10.1146/annurev.med.51.1.357

Cattle/cow population worldwide 2012-2020. (n.d.). Statista. Retrieved 1 June 2020, from https://www.statista.com/statistics/263979/global-cattle-population-since-1990/

Chopra, I., & Roberts, M. (2001). Tetracycline Antibiotics: Mode of Action, Applications, Molecular Biology, and Epidemiology of Bacterial Resistance. *Microbiology and Molecular Biology Reviews*, *65*(2), 232–260. https://doi.org/10.1128/MMBR.65.2.232-260.2001

Crouzet, J., Cauchois, L., Blanche, F., Debussche, L., Thibaut, D., Rouyez, M. C., Rigault, S., Mayaux, J. F., & Cameron, B. (1990). Nucleotide sequence of a Pseudomonas denitrificans 5.4-kilobase DNA fragment containing five cob genes and identification of structural genes encoding S-adenosyl-L-methionine: Uroporphyrinogen III methyltransferase and cobyrinic acid a,c-diamide synthase. *Journal of Bacteriology*, *172*(10), 5968–5979. https://doi.org/10.1128/jb.172.10.5968-5979.1990

Crouzet, J., Levy-Schil, S., Cameron, B., Cauchois, L., Rigault, S., Rouyez, M. C., Blanche, F., Debussche, L., & Thibaut, D. (1991). Nucleotide sequence and genetic analysis of a 13.1-kilobase-pair Pseudomonas denitrificans DNA fragment containing five cob genes and identification of structural genes encoding Cob(I)alamin adenosyltransferase, cobyric acid synthase, and bifunctional cobinamide kinase-cobinamide phosphate guanylyltransferase. *Journal of Bacteriology*, *173*(19), 6074–6087. https://doi.org/10.1128/jb.173.19.6074-6087.1991

Data,https://www.reportsanddata.com, R. and. (n.d.). *Cyanocobalamin Market Trends, Demand & Forecast, 2019-2027*. Retrieved 2 June 2020, from https://www.reportsanddata.com/report-detail/cyanocobalamin-market/amp

Debussche, L., Couder, M., Thibaut, D., Cameron, B., Crouzet, J., & Blanche, F. (1991). Purification and partial characterization of Cob(I)alamin adenosyltransferase from Pseudomonas denitrificans. *Journal of Bacteriology*, *173*(19), 6300–6302. https://doi.org/10.1128/jb.173.19.6300-6302.1991

Debussche, L., Couder, M., Thibaut, D., Cameron, B., Crouzet, J., & Blanche, F. (1992). Assay, purification, and characterization of cobaltochelatase, a unique complex enzyme catalyzing cobalt insertion in hydrogenobyrinic acid a,c-diamide during coenzyme B12 biosynthesis in Pseudomonas denitrificans. *Journal of Bacteriology*, *174*(22), 7445–7451. https://doi.org/10.1128/jb.174.22.7445-7451.1992

Debussche, L., Thibaut, D., Cameron, B., Crouzet, J., & Blanche, F. (1990). Purification and characterization of cobyrinic acid a,c-diamide synthase from Pseudomonas denitrificans. *Journal of Bacteriology*, *172*(11), 6239–6244. https://doi.org/10.1128/jb.172.11.6239-6244.1990

Debussche, L., Thibaut, D., Cameron, B., Crouzet, J., & Blanche, F. (1993). Biosynthesis of the corrin macrocycle of coenzyme B12 in Pseudomonas denitrificans. *Journal of Bacteriology*, *175*(22), 7430–7440. https://doi.org/10.1128/jb.175.22.7430-7440.1993

Deery, E., Schroeder, S., Lawrence, A. D., Taylor, S. L., Seyedarabi, A., Waterman, J., Wilson, K. S., Brown, D., Geeves, M. A., Howard, M. J., Pickersgill, R. W., & Warren, M. J. (2012). An enzyme-trap approach allows isolation of intermediates in cobalamin biosynthesis. *Nature Chemical Biology*, *8*(11), 933–940. https://doi.org/10.1038/nchembio.1086

Delobel, J. (2008). [Master thesis]. Université des sciences et technologie de Lille.

Doniach, D., Roitt, I. M., & Taylor, K. B. (1963). Autoimmune Phenomena in Pernicious Anaemia. *British Medical Journal*, *1*(5342), 1374–1379. https://doi.org/10.1136/bmj.1.5342.1374

Drennan, C. L., Matthews, R. G., & Ludwig, M. L. (1994). Cobalamin-dependent methionine synthase: The structure of a methylcobalamin-binding fragment and implications for other B12-dependent enzymes. *Current Opinion in Structural Biology*, *4*(6), 919–929. https://doi.org/10.1016/0959-440X(94)90275-5

Drummond, J. T., Huang, S., Blumenthal, R. M., & Matthews, R. G. (1993). Assignment of enzymic function to specific protein regions of cobalamin-dependent methionine synthase from Escherichia coli. *Biochemistry*, *32*(36), 9290–9295. https://doi.org/10.1021/bi00087a005

Dunn-Coleman, N. S., Gatenby, A. A., & Valle, F. (2006). *Method for the production of 1,3-propanediol by recombinant organisms comprising genes for coenzyme B12 synthesis* (United States Patent No. US7074608B1). https://patents.google.com/patent/US7074608B1/en

Duplessis, M., Lapierre, H., Pellerin, D., Laforest, J.-P., & Girard, C. L. (2017). Effects of intramuscular injections of folic acid, vitamin B12, or both, on lactational performance and energy status of multiparous dairy cows. *Journal of Dairy Science*, *100*(5), 4051–4064. https://doi.org/10.3168/jds.2016-12381

Elliott, T., Avissar, Y. J., Rhie, G. E., & Beale, S. I. (1990). Cloning and sequence of the Salmonella typhimurium hemL gene and identification of the missing enzyme in hemL mutants as glutamate-1-semialdehyde aminotransferase. *Journal of Bacteriology*, *172*(12), 7071–7084. https://doi.org/10.1128/jb.172.12.7071-7084.1990

Escalante-Semerena, J. C. (2007). Conversion of Cobinamide into Adenosylcobamide in Bacteria and Archaea. *Journal of Bacteriology*, *189*(13), 4555–4560. https://doi.org/10.1128/JB.00503-07

Eschenmoser, A. (1974). Organische Naturstoffsynthese heute Vitamin B12 als Beispiel. *Naturwissenschaften*, *61*(12), 513–525. https://doi.org/10.1007/BF00606511

Eschenmoser, A. (2015). Corrin Syntheses. Part I: Introduction and Overview. *Helvetica Chimica Acta*, *98*(11–12), 1483–1600. https://doi.org/10.1002/hlca.201400277

Evans, J. C., & Mizrahi, V. (2015). The application of tetracyclineregulated gene expression systems in the validation of novel drug targets in Mycobacterium tuberculosis. *Frontiers in Microbiology*, *6*, 812. https://doi.org/10.3389/fmicb.2015.00812

Fang, H., Li, D., Kang, J., Jiang, P., Sun, J., & Zhang, D. (2018a). Metabolic engineering of Escherichia coli for de novo biosynthesis of vitamin B12. *Nature Communications*, *9*(1), 4917. https://doi.org/10.1038/s41467-018-07412-6

Fang, H., Li, D., Kang, J., Jiang, P., Sun, J., & Zhang, D. (2018b). Metabolic engineering of Escherichia coli for de novo biosynthesis of vitamin B12. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-07412-6

Fanica-Gaignier, M., & Clement-Metral, J. (1973). 5-Aminolevulinic-Acid Synthetase of Rhodopseudomonas spheroides Y. *European Journal of Biochemistry*, *40*(1), 19–24. https://doi.org/10.1111/j.1432-1033.1973.tb03164.x

Fatma, Z., Hartman, H., Poolman, M. G., Fell, D. A., Srivastava, S., Shakeel, T., & Yazdani, S. S. (2018). Model-assisted metabolic engineering of Escherichia coli for long chain alkane and alcohol production. *Metabolic Engineering*, *46*, 1–12. https://doi.org/10.1016/j.ymben.2018.01.002

Fayet, O., Ziegelhoffer, T., & Georgopoulos, C. (1989). The groES and groEL heat shock gene products of Escherichia coli are essential for bacterial growth at all temperatures. *Journal of Bacteriology*, *171*(3), 1379–1385. https://doi.org/10.1128/jb.171.3.1379-1385.1989

Fedosov, S. N., Brito, A., Miller, J. W., Green, R., & Allen, L. H. (2015). Combined indicator of vitamin B12 status: Modification for missing biomarkers and folate status and recommendations for revised cut-points. *Clinical Chemistry and Laboratory Medicine (CCLM)*, *53*(8), 1215–1225. https://doi.org/10.1515/cclm-2014-0818

Ferreira, G. C., & Gong, J. (1995). 5-Aminolevulinate synthase and the first step of heme biosynthesis. *Journal of Bioenergetics and Biomembranes*, *27*(2), 151–159. https://doi.org/10.1007/BF02110030

Fincker, M., & Spormann, A. M. (2017). Biochemistry of Catabolic Reductive Dehalogenation. *Annual Review of Biochemistry*, *86*(1), 357–386. https://doi.org/10.1146/annurev-biochem-061516-044829

Frank, S., Deery, E., Brindley, A. A., Leech, H. K., Lawrence, A., Heathcote, P., Schubert, H. L., Brocklehurst, K., Rigby, S. E. J., Warren, M. J., & Pickersgill, R. W. (2007). Elucidation of Substrate Specificity in the Cobalamin (Vitamin B12) Biosynthetic Methyltransferases structure and function of the C20 methyltransferase (CbiL) from Methanothermobacter thermoautotrophicus. *Journal of Biological Chemistry*, *282*(33), 23957–23969. https://doi.org/10.1074/jbc.M703827200

Fresquet, V., Williams, L., & Raushel, F. M. (2004). Mechanism of Cobyrinic Acid a,c-Diamide Synthetase from Salmonella typhimurium LT2. *Biochemistry*, *43*(33), 10619–10627. https://doi.org/10.1021/bi048972x

Frey, P. A., Hegeman, A. D., & Ruzicka, F. J. (2008). The Radical SAM Superfamily. *Critical Reviews in Biochemistry and Molecular Biology*, *43*(1), 63–88. https://doi.org/10.1080/10409230701829169

Friedmann, H. C., & Cagen, L. M. (1970). Microbial Biosynthesis of B12-Like Compounds. *Annual Review of Microbiology*, *24*(1), 159–208. https://doi.org/10.1146/annurev.mi.24.100170.001111

Fujii, K., & Huennekens, F. M. (1974). Activation of Methionine Synthetase by a Reduced Triphosphopyridine Nucleotide-dependent Flavoprotein System. *Journal of Biological Chemistry*, *249*(21), 6745–6753. https://doi.org/10.1016/S0021-9258(19)42122-4

Gamini Kannangara, C., Gough, S. P., Bruyant, P., Kenneth Hoober, J., Kahn, A., & von Wettstein, D. (1988). tRNAGlu as a cofactor in δ-aminolevulinate biosynthesis: Steps that regulate chlorophyll synthesis. *Trends in Biochemical Sciences*, *13*(4), 139–143. https://doi.org/10.1016/0968-0004(88)90071-0

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In J. M. Walker (Ed.), *The Proteomics Protocols Handbook* (pp. 571–607). Humana Press. https://doi.org/10.1385/1-59259-890-0:571

Gibson, K. D., Neuberger, A., & Scott, J. J. (1955). The purification and properties of δ-aminolaevulic acid dehydrase. *Biochemical Journal*, *61*(4), 618–629. https://doi.org/10.1042/bj0610618

Glazyrina, J., Materne, E.-M., Dreher, T., Storm, D., Junne, S., Adams, T., Greller, G., & Neubauer, P. (2010). High cell density cultivation and recombinant protein production with Escherichia coli in a rocking-motion-type bioreactor. *Microbial Cell Factories*, *9*, 42. https://doi.org/10.1186/1475-2859-9-42

Gray, M. J., & Escalante-Semerena, J. C. (2007). Single-enzyme conversion of FMNH2 to 5,6-dimethylbenzimidazole, the lower ligand of B12. *Proceedings of the National Academy of Sciences*, *104*(8), 2921–2926. https://doi.org/10.1073/pnas.0609270104

Green, R., Allen, L. H., Bjørke-Monsen, A.-L., Brito, A., Guéant, J.-L., Miller, J. W., Molloy, A. M., Nexo, E., Stabler, S., Toh, B.-H., Ueland, P. M., & Yajnik, C. (2017). Vitamin B12 deficiency. *Nature Reviews Disease Primers*, *3*, 17040. https://doi.org/10.1038/nrdp.2017.40

Grimm, B. (1990). Primary structure of a key enzyme in plant tetrapyrrole synthesis: Glutamate 1-semialdehyde aminotransferase. *Proceedings of the National Academy of Sciences*, *87*(11), 4169–4173. https://doi.org/10.1073/pnas.87.11.4169

Guest, J. R., Friedman, S., Woods, D. D., & Lester Smith, E. (1962). A Methyl Analogue of Cobamide Coenzyme in Relation to Methionine Synthesis by Bacteria. *Nature*, *195*(4839), 340–342. https://doi.org/10.1038/195340a0

Hansson, M., Rutberg, L., Schröder, I., & Hederstedt, L. (1991). The Bacillus subtilis hemAXCDBL gene cluster, which encodes enzymes of the biosynthetic pathway from glutamate to uroporphyrinogen III. *Journal of Bacteriology*, *173*(8), 2590–2599. https://doi.org/10.1128/jb.173.8.2590-2599.1991

Hargrove, R. E., & Abraham, L. (1955). *Process for the manufacture of vitamin b12* (United States Patent No. US2715602A). https://patents.google.com/patent/US2715602A/en#citedBy

Hazra, A. B., Han, A. W., Mehta, A. P., Mok, K. C., Osadchiy, V., Begley, T. P., & Taga, M. E. (2015). Anaerobic biosynthesis of the lower ligand of vitamin B12. *Proceedings of the National Academy of Sciences*, *112*(34), 10792–10797. https://doi.org/10.1073/pnas.1509132112

Hodgkin, D. C. (n.d.-a). *Web of Stories—Dorothy Hodgkin. Meeting Lester Smith* [Interview]. https://www.webofstories.com/play/dorothy.hodgkin/31

Hodgkin, D. C. (n.d.-b). *Web of Stories—Dorothy Hodgkin. Working on Vitamin B12* [Interview]. https://www.webofstories.com/play/dorothy.hodgkin/34

Hodgkin, D. C., Kamper, J., Mackay, M., Pickworth, J., Trueblood, K. N., & White, J. G. (1956). Structure of Vitamin B 12. *Nature*, *178*(4524), 64–66. https://doi.org/10.1038/178064a0

Hodgkin, D. C., Pickworth, J., Robertson, J. H., Trueblood, K. N., Prosen, R. J., & White, J. G. (1955). Structure of Vitamin B 12: The Crystal Structure of the Hexacarboxylic Acid derived from B 12 and the Molecular Structure of the Vitamin. *Nature*, *176*(4477), 325–328. https://doi.org/10.1038/176325a0

Hogan, K. G., Lorentz, P. P., & Gibb, F. M. (1973). The diagnosis and treatment of vitamin B12 deficiency in young lambs. *New Zealand Veterinary Journal*, *21*(11), 234–237. https://doi.org/10.1080/00480169.1973.34115

Holliday, G. L., Akiva, E., Meng, E. C., Brown, S. D., Calhoun, S., Pieper, U., Sali, A., Booker, S. J., & Babbitt, P. C. (2018). Chapter One - Atlas of the Radical SAM Superfamily: Divergent Evolution of Function Using a "Plug and Play" Domain. In V. Bandarian (Ed.), *Methods in Enzymology* (Vol. 606, pp. 1–71). Academic Press. https://doi.org/10.1016/bs.mie.2018.06.004

Hopwood, D. A., & Chater, K. F. (1980). Fresh approaches to antibiotic production. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *290*(1040), 313–328. https://doi.org/10.1098/rstb.1980.0097

Hungerer, C., Troup, B., Römling, U., & Jahn, D. (1995). Regulation of the hemA gene during 5-aminolevulinic acid formation in Pseudomonas aeruginosa. *Journal of Bacteriology*, *177*(6), 1435–1443. https://doi.org/10.1128/jb.177.6.1435-1443.1995

Ilag, L. L., Jahn, D., Eggertsson, G., & Söll, D. (1991). The Escherichia coli hemL gene encodes glutamate 1-semialdehyde aminotransferase. *Journal of Bacteriology*, *173*(11), 3408–3413. https://doi.org/10.1128/jb.173.11.3408-3413.1991

Inoue, H., Nojima, H., & Okayama, H. (1990). High efficiency transformation of Escherichia coli with plasmids. *Gene*, *96*(1), 23–28. https://doi.org/10.1016/0378-1119(90)90336-P

Jahn, D., Michelsen, U., & Söll, D. (1991). Two glutamyl-tRNA reductase activities in Escherichia coli. *Journal of Biological Chemistry*, *266*(4), 2542–2548. https://doi.org/10.1016/S0021-9258(18)52279-1

Jahn, D., Verkamp, E., & So¨ll, D. (1992). Glutamyl-transfer RNA: A precursor of heme and chlorophyll biosynthesis. *Trends in Biochemical Sciences*, *17*(6), 215–218. https://doi.org/10.1016/0968-0004(92)90380-R

Jencks, W. (1975). Binding energy, specificity, and enzymic catalysis: The Circe Effect. *Advances in Enzymology and Related Areas of Molecular Biology*, *43*. https://doi.org/10.1002/9780470122884.ch4

Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., & Yang, S. (2015). Multigene editing in the Escherichia coli genome via the CRISPR-Cas9 system. *Applied and Environmental Microbiology*, *81*(7), 2506–2514. https://doi.org/10.1128/AEM.04023-14

Johnson, K. A., & Goody, R. S. (2011). The Original Michaelis Constant: Translation of the 1913 Michaelis–Menten Paper. *Biochemistry*, *50*(39), 8264–8269. https://doi.org/10.1021/bi201284u

Johnson, M. G., & Escalante-Semerena, J. C. (1992). Identification of 5,6-dimethylbenzimidazole as the Co alpha ligand of the cobamide synthesized by Salmonella typhimurium. Nutritional characterization of mutants defective in biosynthesis of the imidazole ring. *Journal of Biological Chemistry*, *267*(19), 13302–13305. https://doi.org/10.1016/S0021-9258(18)42210-7

Jost, M., Fernández-Zapata, J., Polanco, M. C., Ortiz-Guerrero, J. M., Chen, P. Y.-T., Kang, G., Padmanabhan, S., Elías-Arnanz, M., & Drennan, C. L. (2015). Structural basis for gene regulation by a B 12 -dependent photoreceptor. *Nature*, *526*(7574), 536–541. https://doi.org/10.1038/nature14950

Kieninger, C., Deery, E., Lawrence, A. D., Podewitz, M., Wurst, K., Nemoto-Smith, E., Widner, F. J., Baker, J. A., Jockusch, S., Kreutz, C. R., Liedl, K. R., Gruber, K., Warren, M. J., & Kräutler, B. (2019). The Hydrogenobyric Acid Structure Reveals the Corrin Ligand as an Entatic State Module Empowering B12 Cofactors for Catalysis. *Angewandte Chemie International Edition*, *58*(31), 10756–10760. https://doi.org/10.1002/anie.201904713

Kiuchi, F., Thibaut, D., Debussche, L., Leeper, F. J., Blanche, F., & Battersby, A. R. (1992). Biosynthesis of vitamin B12: Stereochemistry of transfer of a hydride equivalent from NADPH by precorrin-6x reductase. *Journal of the Chemical Society, Chemical Communications*, *4*, 306–308. https://doi.org/10.1039/C39920000306

Ko, Y., Ashok, S., Ainala, S. K., Sankaranarayanan, M., Chun, A. Y., Jung, G. Y., & Park, S. (2014). Coenzyme B12 can be produced by engineered Escherichia coli under both anaerobic and aerobic conditions. *Biotechnology Journal*, *9*(12), 1526–1535. https://doi.org/10.1002/biot.201400221

Kräutler, B. (2019). Chapter 20—Biological Organometallic Chemistry of Vitamin B12-Derivatives. In T. Hirao & T. Moriuchi (Eds.), *Advances in Bioorganometallic Chemistry* (pp. 399–430). Elsevier. https://doi.org/10.1016/B978-0-12-814197-7.00020-0

Kräutler, B., Moll, J., & Thauer, R. K. (1987). The corrinoid from Methanobacterium thermoautotrophicum (Marburg strain). *European Journal of Biochemistry*, *162*(2), 275–278. https://doi.org/10.1111/j.1432-1033.1987.tb10596.x

Lascelles, J. (1978). Regulation of pyrrole synthesis. *The Photosynthetic Bacteria*. https://ci.nii.ac.jp/naid/10016640487/

Lassiter, C. A., Ward, G. M., Huffman, C. F., Duncan, C. W., & Webster, H. D. (1953). Crystalline vitamin B12 requirement of the young dairy calf. *Journal of Dairy Science*, *36*, 997–1005. https://doi.org/10.3168/jds.S0022-0302(53)91587-2

Laver, W. G., Neuberger, A., & Udenfriend, S. (1958). Initial stages in the biosynthesis of porphyrins. 1. The formation of δ-aminolaevulic acid by particles obtained from chicken erythrocytes. *Biochemical Journal*, *70*(1), 4–14. https://doi.org/10.1042/bj0700004

Lawrence, A. D., Deery, E., McLean, K. J., Munro, A. W., Pickersgill, R. W., Rigby, S. E. J., & Warren, M. J. (2008). Identification, Characterization, and Structure/Function Analysis of a Corrin Reductase Involved in Adenosylcobalamin Biosynthesis. *Journal of Biological Chemistry*, *283*(16), 10813–10821. https://doi.org/10.1074/jbc.M710431200

Lawrence, J. G., & Roth, J. R. (1995). The cobalamin (coenzyme B12) biosynthetic genes of Escherichia coli. *Journal of Bacteriology*, *177*(22), 6371–6380. https://doi.org/10.1128/jb.177.22.6371-6380.1995

Layer, G., Jahn, D., Deery, E., Lawrence, A. D., & Warren, M. J. (2010). 7.13—Biosynthesis of Heme and Vitamin B12. In H.-W. (Ben) Liu & L. Mander (Eds.), *Comprehensive Natural Products II* (pp. 445–499). Elsevier. https://doi.org/10.1016/B978-008045382-8.00144-1

Leeper, F. J., Warren, M. J., Kelly, J. M., & Lawrence, A. D. (2012). Biosynthesis of Vitamin B_{12}. In K. M. Kadish, K. M. Smith, & R. Guilard, *Handbook of Porphyrin Science (Volume 25)* (Vol. 25, pp. 1–81). World Scientific Publishing Company. https://doi.org/10.1142/9789814397605_0017

Lenhert, P. G., & Hodgkin, D. C. (1961). Structure of the 5,6-Dimethylbenzimidazolylcobamide Coenzyme. *Nature*, *192*(4806), 937–938. https://doi.org/10.1038/192937a0

Li, K.-T., Liu, D.-H., Chu, J., Wang, Y.-H., Zhuang, Y.-P., & Zhang, S.-L. (2008). An effective and simplified pH-stat control strategy for the industrial fermentation of vitamin B12 by Pseudomonas denitrificans. *Bioprocess and Biosystems Engineering*, *31*(6), 605–610. https://doi.org/10.1007/s00449-008-0209-5

Li, K.-T., Liu, D.-H., Li, Y.-L., Chu, J., Wang, Y.-H., Zhuang, Y.-P., & Zhang, S.-L. (2008). Improved large-scale production of vitamin B12 by Pseudomonas denitrificans with betaine feeding. *Bioresource Technology*, *99*(17), 8516–8520. https://doi.org/10.1016/j.biortech.2008.03.023

Li, K.-T., Zhou, J., Cheng, X., & Wei, S. (2012). Study on the dissolved oxygen control strategy in large-scale vitamin B12 fermentation by Pseudomonas denitrificans. *Journal of Chemical Technology & Biotechnology*, *87*(12), 1648–1653. https://doi.org/10.1002/jctb.3804

Lindenbaum, J., Healton, E. B., Savage, D. G., Brust, J. C. M., Garrett, T. J., Podell, E. R., Margell, P. D., Stabler, S. P., & Allen, R. H. (1988). Neuropsychiatric Disorders Caused by Cobalamin Deficiency in the Absence of Anemia or Macrocytosis. *New England Journal of Medicine*, *318*(26), 1720–1728. https://doi.org/10.1056/NEJM198806303182604

Lindstrand, K., & Ståhlberg, K.-G. (1963). On Vitamin B12 Forms in Human Plasma. *Acta Medica Scandinavica*, *174*(6), 665–669. https://doi.org/10.1111/j.0954-6820.1963.tb16534.x

Lüer, C., Schauer, S., Möbius, K., Schulze, J., Schubert, W.-D., Heinz, D. W., Jahn, D., & Moser, J. (2005). Complex Formation between Glutamyl-tRNA Reductase and Glutamate-1-

semialdehyde 2,1-Aminomutase in Escherichia coli during the Initial Reactions of Porphyrin Biosynthesis. *Journal of Biological Chemistry*, *280*(19), 18568–18572. https://doi.org/10.1074/jbc.M500440200

Lundqvist, J., Elmlund, D., Heldt, D., Deery, E., Söderberg, C. A. G., Hansson, M., Warren, M., & Al-Karadaghi, S. (2009). The AAA(+) motor complex of subunits CobS and CobT of cobaltochelatase visualized by single particle electron microscopy. *Journal of Structural Biology*, *167*(3), 227–234. https://doi.org/10.1016/j.jsb.2009.06.013

Maggio-Hall, L. A., Claas, K. R., & Escalante-Semerena, J. C. Y. 2004. (n.d.). The last step in coenzyme B12 synthesis is localized to the cell membrane in bacteria and archaea. *Microbiology*, *150*(5), 1385–1395. https://doi.org/10.1099/mic.0.26952-0

Mandal, M., & Breaker, R. R. (2004). Gene regulation by riboswitches. *Nature Reviews Molecular Cell Biology*, *5*(6), 451–463. https://doi.org/10.1038/nrm1403

Markets, R. and. (2019, May 2). *Global Meat Sector Market Analysis & Forecast Report, 2019—A $1.14 Trillion Industry Opportunity by 2023*. GlobeNewswire News Room. http://www.globenewswire.com/news-release/2019/05/02/1815144/0/en/Global-Meat-Sector-Market-Analysis-Forecast-Report-2019-A-1-14-Trillion-Industry-Opportunity-by-2023.html

Markham, G. D., & Pajares, M. A. (2009). Structure-function relationships in methionine adenosyltransferases. *Cellular and Molecular Life Sciences : CMLS*, *66*(4), 636–648. https://doi.org/10.1007/s00018-008-8516-1

Martens, J.-H., Barg, H., Warren, M., & Jahn, D. (2002). Microbial production of vitamin B12. *Applied Microbiology and Biotechnology*, *58*(3), 275–285. https://doi.org/10.1007/s00253-001-0902-7

Martin, V. J. J., Pitera, D. J., Withers, S. T., Newman, J. D., & Keasling, J. D. (2003). Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nature Biotechnology*, *21*(7), 796–802. https://doi.org/10.1038/nbt833

Matthews, R. G., Koutmos, M., & Datta, S. (2008). Cobalamin-dependent and cobamide-dependent methyltransferases. *Current Opinion in Structural Biology*, *18*(6), 658–666. https://doi.org/10.1016/j.sbi.2008.11.005

McDowell, L. R. (2000). Reevaluation of the Metabolic Essentiality of the Vitamins—Review -. *Asian-Australasian Journal of Animal Sciences*, *13*(1), 115–125. https://doi.org/2000.13.1.115

McGoldrick, H., Deery, E., Warren, M., & Heathcote, P. (2002). Cobalamin (vitamin B12) biosynthesis in Rhodobacter capsulatus. *Biochemical Society Transactions*, *30*(4), 646–648. https://doi.org/10.1042/bst0300646

Medicine, I. of. (1998). *Dietary Reference Intakes for Thiamin, Riboflavin, Niacin, Vitamin B6, Folate, Vitamin B12, Pantothenic Acid, Biotin, and Choline*. The National Academies Press. https://doi.org/10.17226/6015

Miller, J., H. (1972). *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press,U.S.

Minot, G. R., & Murphy, W. P. (1926). Treatment of pernicious anemia by a special diet. *Journal of the American Medical Association*, *87*(7), 470–476. https://doi.org/10.1001/jama.1926.02680070016005

Moore, S. J., Lawrence, A. D., Biedendieck, R., Deery, E., Frank, S., Howard, M. J., Rigby, S. E. J., & Warren, M. J. (2013). Elucidation of the anaerobic pathway for the corrin component of cobalamin (vitamin B12). *Proceedings of the National Academy of Sciences*, *110*(37), 14906–14911. https://doi.org/10.1073/pnas.1308098110

Moore, S. J., Mayer, M. J., Biedendieck, R., Deery, E., & Warren, M. J. (2014). Towards a cell factory for vitamin B12 production in Bacillus megaterium: Bypassing of the cobalamin riboswitch control elements. *New Biotechnology*, *31*(6), 553–561. https://doi.org/10.1016/j.nbt.2014.03.003

Moore, S. J., & Warren, M. J. (2012). The anaerobic biosynthesis of vitamin B12. *Biochemical Society Transactions*, *40*(3), 581–586. https://doi.org/10.1042/BST20120066

Moser, J., Schubert, W.-D., Beier, V., Bringemeier, I., Jahn, D., & Heinz, D. W. (2001). V-shaped structure of glutamyl-tRNA reductase, the first enzyme of tRNA-dependent tetrapyrrole biosynthesis. *The EMBO Journal*, *20*(23), 6583–6590. https://doi.org/10.1093/emboj/20.23.6583

Murphy, E. (1985). Nucleotide sequence of a spectinomycin adenyltransferase AAD(9) determinant from Staphylococcus aureus and its relationship to AAD(3") (9). *Molecular & General Genetics: MGG*, *200*(1), 33–39. https://doi.org/10.1007/BF00383309

Nahvi, A., Barrick, J. E., & Breaker, R. R. (2004). Coenzyme B12 riboswitches are widespread genetic control elements in prokaryotes. *Nucleic Acids Research*, *32*(1), 143–150. https://doi.org/10.1093/nar/gkh167

Nakamura, C. E., & Whited, G. M. (2003). Metabolic engineering for the microbial production of 1,3-propanediol. *Current Opinion in Biotechnology*, *14*(5), 454–459. https://doi.org/10.1016/j.copbio.2003.08.005

Nogaj, L. A., & Beale, S. I. (2005). Physical and Kinetic Interactions between Glutamyl-tRNA Reductase and Glutamate-1-semialdehyde Aminotransferase of Chlamydomonas reinhardtii*. *Journal of Biological Chemistry*, *280*(26), 24301–24307. https://doi.org/10.1074/jbc.M502483200

O'Neill, G. P., Chen, M.-W., & Söll, D. (1989). δ-Aminolevulinic acid biosynthesis in Escherichia coli and Bacillus subtilis involves formation of glutamyl-tRNA. *FEMS Microbiology Letters*, *60*(3), 255–259. https://doi.org/10.1111/j.1574-6968.1989.tb03482.x

Orłowska, M., Steczkiewicz, K., & Muszewska, A. (2021). Utilization of Cobalamin Is Ubiquitous in Early-Branching Fungal Phyla. *Genome Biology and Evolution*, *13*(4), evab043. https://doi.org/10.1093/gbe/evab043

Orth, P., Schnappinger, D., Hillen, W., Saenger, W., & Hinrichs, W. (2000). Structural basis of gene regulation by the tetracycline inducible Tet repressor–operator system. *Nature Structural Biology*, *7*(3), 215–219. https://doi.org/10.1038/73324

O'Toole, G. A., & Escalante-Semerena, J. C. (1995). Purification and Characterization of the Bifunctional CobU Enzyme of Salmonella typhimurium LT2: Evidence for a CobU~GMP intermediate. *Journal of Biological Chemistry*, *270*(40), 23560–23569. https://doi.org/10.1074/jbc.270.40.23560

Payne, K. A. P., Quezada, C. P., Fisher, K., Dunstan, M. S., Collins, F. A., Sjuts, H., Levy, C., Hay, S., Rigby, S. E. J., & Leys, D. (2015). Reductive dehalogenase structure suggests a mechanism for B12-dependent dehalogenation. *Nature*, *517*(7535), 513–516. https://doi.org/10.1038/nature13901

Pel, H. J., & Hopper, S. (2010). *Genes and their encoded polypeptides involved in the biosynthetic pathway of vitamin b12, vectors and host cells comprising the genes, and process for producing vitamin b12* (Patent No. TWI328612B). https://patents.google.com/patent/TWI328612B/en

Pope, C. F., McHugh, T. D., & Gillespie, S. H. (2010). Methods to Determine Fitness in Bacteria. In S. H. Gillespie & T. D. McHugh (Eds.), *Antibiotic Resistance Protocols: Second Edition* (pp. 113–121). Humana Press. https://doi.org/10.1007/978-1-60327-279-7_9

Promoters/Catalog/Anderson—Parts.igem.org. (n.d.). Retrieved 1 February 2021, from http://parts.igem.org/Promoters/Catalog/Anderson

Ragsdale, S. W. (2008). Catalysis of Methyl Group Transfers Involving Tetrahydrofolate and B12. In *Vitamins & Hormones* (Vol. 79, pp. 293–324). Academic Press. https://doi.org/10.1016/S0083-6729(08)00410-X

Ramamoorthy, S., & Kushner, D. J. (1975). Binding of mercuric and other heavy metal ions by microbial growth media. *Microbial Ecology*, *2*(2), 162–176. https://doi.org/10.1007/BF02010436

Randau, L., Schauer, S., Ambrogelly, A., Salazar, J. C., Moser, J., Sekine, S., Yokoyama, S., Söll, D., & Jahn, D. (2004). TRNA Recognition by Glutamyl-tRNA Reductase. *Journal of Biological Chemistry*, *279*(33), 34931–34937. https://doi.org/10.1074/jbc.M401529200

Ranquet, C., Ollagnier-de-Choudens, S., Loiseau, L., Barras, F., & Fontecave, M. (2007). Cobalt Stress in Escherichia coli: THE EFFECT ON THE IRON-SULFUR PROTEINS*. *Journal of Biological Chemistry*, *282*(42), 30442–30451. https://doi.org/10.1074/jbc.M702519200

Raux, E., Lanois, A., Levillayer, F., Warren, M. J., Brody, E., Rambach, A., & Thermes, C. (1996). Salmonella typhimurium cobalamin (vitamin B12) biosynthetic genes: Functional studies in S. typhimurium and Escherichia coli. *Journal of Bacteriology*, *178*(3), 753–767. https://doi.org/10.1128/jb.178.3.753-767.1996

Raux, E., Lanois, A., Rambach, A., Warren, M. J., & Thermes, C. (1998). Cobalamin (vitamin B12) biosynthesis: Functional characterization of the Bacillus megaterium cbi genes required to convert uroporphyrinogen III into cobyrinic acid a,c-diamide. *Biochemical Journal*, *335*(1), 167–173. https://doi.org/10.1042/bj3350167

Raux, E., Lanois, A., Warren, M. J., Rambach, A., & Thermes, C. (1998). Cobalamin (vitamin B12) biosynthesis: Identification and characterization of a Bacillus megaterium cobI operon. *Biochemical Journal*, *335*(1), 159–166. https://doi.org/10.1042/bj3350159

Raux, E., Schubert, H. L., & Warren*, M. J. (2000). Biosynthesis of cobalamin (vitamin B12): A bacterial conundrum. *Cellular and Molecular Life Sciences CMLS*, *57*(13), 1880–1893. https://doi.org/10.1007/PL00000670

Reid, J. D., & Hunter, C. N. (2002). Current understanding of the function of magnesium chelatase. *Biochemical Society Transactions*, *30*(4), 643–645. https://doi.org/10.1042/bst0300643

Rickes, E. L., Brink, N. G., Koniuszy, F. R., Wood, T. R., & Folkers, K. (1948). Crystalline Vitamin B12. *Science*, *107*(2781), 396–397. https://doi.org/10.1126/science.107.2781.396

Rieble, S., & Beale, S. I. (1991). Purification of glutamyl-tRNA reductase from Synechocystis sp. PCC 6803. *Journal of Biological Chemistry*, *266*(15), 9740–9745. https://doi.org/10.1016/S0021-9258(18)92883-8

Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in Escherichia coli: Advances and challenges. *Frontiers in Microbiology*, *5*. https://doi.org/10.3389/fmicb.2014.00172

Roth, J. R., Lawrence, J. G., Rubenfield, M., Kieffer-Higgins, S., & Church, G. M. (1993). Characterization of the cobalamin (vitamin B12) biosynthetic genes of Salmonella typhimurium. *Journal of Bacteriology*, *175*(11), 3303–3316. https://doi.org/10.1128/jb.175.11.3303-3316.1993

Sambrook, J., & Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual* (Vol. 2). CSHL Press.

Sattler, I., Roessner, C. A., Stolowich, N. J., Hardin, S. H., Harris-Haller, L. W., Yokubaitis, N. T., Murooka, Y., Hashimoto, Y., & Scott, A. I. (1995). Cloning, sequencing, and expression of the uroporphyrinogen III methyltransferase cobA gene of Propionibacterium freudenreichii (shermanii). *Journal of Bacteriology*, *177*(6), 1564–1569. https://doi.org/10.1128/jb.177.6.1564-1569.1995

Schnellbaecher, A., Binder, D., Bellmaine, S., & Zimmer, A. (2019). Vitamins in cell culture media: Stability and stabilization strategies. *Biotechnology and Bioengineering*, *116*(6), 1537–1555. https://doi.org/10.1002/bit.26942

Schroeder, S., Lawrence, A. D., Biedendieck, R., Rose, R.-S., Deery, E., Graham, R. M., McLean, K. J., Munro, A. W., Rigby, S. E. J., & Warren, M. J. (2009). Demonstration That CobG, the Monooxygenase Associated with the Ring Contraction Process of the Aerobic Cobalamin (Vitamin B12) Biosynthetic Pathway, Contains an Fe-S Center and a Mononuclear Non-heme Iron Center. *Journal of Biological Chemistry*, *284*(8), 4796–4805. https://doi.org/10.1074/jbc.M807184200

Schubert, H. L., Blumenthal, R. M., & Cheng, X. (2003). Many paths to methyltransfer: A chronicle of convergence. *Trends in Biochemical Sciences*, *28*(6), 329–335. https://doi.org/10.1016/S0968-0004(03)00090-2

Shemin, D., Corcoran, J. W., Rosenblum, C., & Miller, I. M. (1956). On the biosynthesis of the porphyrinlike moiety of vitamin B12. *Science (New York, N.Y.)*, *124*(3215), 272. https://doi.org/10.1126/science.124.3215.272

Shemin, D., & Kikuchi, G. (1958). Enzymatic synthesis of δ-aminolevulinic acid. *Annals of the New York Academy of Sciences*, *75*(1), 122–128. https://doi.org/10.1111/j.1749-6632.1958.tb36857.x

Shemin, D., & Rittenberg, D. (1945). The utilization of glycine for the synthesis of a porphyrin. *Journal of Biological Chemistry*, *159*(2), 567–568. https://doi.org/10.1016/S0021-9258(19)52819-8

Smith, A. D., Warren, M. J., & Refsum, H. (2018). Vitamin B12. In *Advances in Food and Nutrition Research* (Vol. 83, p. 363). Elsevier Inc.

Smith, E. L. (1948). Purification of Anti-pernicious Anæmia Factors from Liver. *Nature*, *161*(4095), 638–639. https://doi.org/10.1038/161638a0

Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F., & Miller, N. E. (2001). Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: Functional characterization using new analysis and information visualization methods. *Nucleic Acids Research*, *29*(5), 1097–1106. https://doi.org/10.1093/nar/29.5.1097

Stabler, S. P., & Allen, R. H. (2004). Vitamin B12 Deficiency as a Worldwide Problem. *Annual Review of Nutrition*, *24*(1), 299–326. https://doi.org/10.1146/annurev.nutr.24.012003.132440

Stich, T. A., Buan, N. R., Escalante-Semerena, J. C., & Brunold, T. C. (2005). Spectroscopic and Computational Studies of the ATP:Corrinoid Adenosyltransferase (CobA) from Salmonella enterica: Insights into the Mechanism of Adenosylcobalamin Biosynthesis. *Journal of the American Chemical Society*, *127*(24), 8710–8719. https://doi.org/10.1021/ja042142p

Stupperich, E., Eisinger, H. J., & Kräutler, B. (1988). Diversity of corrinoids in acetogenic bacteria. *European Journal of Biochemistry*, *172*(2), 459–464. https://doi.org/10.1111/j.1432-1033.1988.tb13910.x

Stupperich, E., Eisinger, H.-J., & Schurr, S. (1990). Corrinoids in anaerobic bacteria. *FEMS Microbiology Letters*, *87*(3), 355–359. https://doi.org/10.1016/0378-1097(90)90478-9

Stupperich, E., & Kräutler, B. (1988). Pseudo vitamin B12 or 5-hydroxybenzimidazolyl-cobamide are the corrinoids found in methanogenic bacteria. *Archives of Microbiology*, *149*(3), 268–271. https://doi.org/10.1007/BF00422016

Suh, S.-J., & Escalante-Semerena, J. C. (1993). Cloning, sequencing and overexpression of cob A which encodes ATP:corrinoid adenosyltranferase in Salmonella typhimurium. *Gene*, *129*(1), 93–97. https://doi.org/10.1016/0378-1119(93)90701-4

Tabor, S. (1990). Expression Using the T7 RNA Polymerase/Promoter System. *Current Protocols in Molecular Biology*, *11*(1), 16.2.1-16.2.11. https://doi.org/10.1002/0471142727.mb1602s11

Taylor, K. B., Roitt, I. M., Doniach, D., Couchman, K. G., & Shapland, C. (1962). Autoimmune Phenomena in Pernicious Anaemia: Gastric Antibodies. *British Medical Journal*, *2*(5316), 1347–1352. https://doi.org/10.1136/bmj.2.5316.1347

Thibaut, D., Couder, M., Crouzet, J., Debussche, L., Cameron, B., & Blanche, F. (1990). Assay and purification of S-adenosyl-L-methionine:precorrin-2 methyltransferase from Pseudomonas denitrificans. *Journal of Bacteriology*, *172*(11), 6245–6251. https://doi.org/10.1128/jb.172.11.6245-6251.1990

Thibaut, D., Couder, M., Famechon, A., Debussche, L., Cameron, B., Crouzet, J., & Blanche, F. (1992). The final step in the biosynthesis of hydrogenobyrinic acid is catalyzed by the cobH gene product with precorrin-8x as the substrate. *Journal of Bacteriology*, *174*(3), 1043–1049.

Toh, B.-H., van Driel, I. R., & Gleeson, P. A. (1997). Pernicious Anemia. *New England Journal of Medicine*, *337*(20), 1441–1448. https://doi.org/10.1056/NEJM199711133372007

Toohey, J. I., & Barker, H. A. (1961). Isolation of Coenzyme B12 from Liver. *Journal of Biological Chemistry*, *236*(2), 560–563. https://doi.org/10.1016/S0021-9258(18)64403-5

Trzebiatowski, J. R., O'Toole, G. A., & 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*. https://doi.org/10.1128/jb.176.12.3568-3575.1994

Uzar, H. C., Battersby, A. R., Carpenter, T. A., & Leeper, F. J. (1987). Biosynthesis of porphyrins and related, macrocycles. Part 28. Development of a pulse labelling method to determine the C-methylation sequence for vitamin B12. *Journal of the Chemical Society, Perkin Transactions 1*, *0*, 1689–1696. https://doi.org/10.1039/P19870001689

Vandamme, E. J., & Revuelta, J. L. (2016). *Industrial Biotechnology of Vitamins, Biopigments, and Antioxidants*. John Wiley & Sons.

Vévodová, J., Graham, R. M., Raux, E., Schubert, H. L., Roper, D. I., Brindley, A. A., Ian Scott, A., Roessner, C. A., Stamford, N. P. J., Elizabeth Stroupe, M., Getzoff, E. D., Warren, M. J., & Wilson, K. S. (2004). Structure/Function Studies on a S-Adenosyl-l-methionine-dependent Uroporphyrinogen III C Methyltransferase (SUMT), a Key Regulatory Enzyme of Tetrapyrrole Biosynthesis. *Journal of Molecular Biology*, *344*(2), 419–433. https://doi.org/10.1016/j.jmb.2004.09.020

Vitamin B12 market size worldwide 2021 forecast. (n.d.). Statista. Retrieved 15 May 2020, from https://www.statista.com/statistics/967092/total-cobalamin-market-size-globally/

Warren, M. J., Raux, E., Schubert, H. L., & Escalante-Semerena, J. C. (2002). The biosynthesis of adenosylcobalamin (vitamin B12). *Natural Product Reports*, *19*(4), 390–412. https://doi.org/10.1039/b108967f

Warren, M. J., & Scott, A. I. (1990). Tetrapyrrole assembly and modification into the ligands of biologically functional cofactors. *Trends in Biochemical Sciences*, *15*(12), 486–491. https://doi.org/10.1016/0968-0004(90)90304-T

Watanabe, F., Katsura, H., Takenaka, S., Fujita, T., Abe, K., Tamura, Y., Nakatsuka, T., & Nakano, Y. (1999). Pseudovitamin B12 Is the Predominant Cobamide of an Algal Health Food, Spirulina Tablets. *Journal of Agricultural and Food Chemistry*, *47*(11), 4736–4741. https://doi.org/10.1021/jf990541b

Werf, M. J. van der, & Zeikus, J. G. (1996). 5-Aminolevulinate production by Escherichia coli containing the Rhodobacter sphaeroides hemA gene. *Applied and Environmental Microbiology*. https://doi.org/10.1128/aem.62.10.3560-3566.1996

Wexler, A. G., Schofield, W. B., Degnan, P. H., Folta-Stogniew, E., Barry, N. A., & Goodman, A. L. (2018). Human gut Bacteroides capture vitamin B12 via cell surface-exposed lipoproteins. *ELife*, *7*, e37138. https://doi.org/10.7554/eLife.37138

Whipple, G. H., & Robscheit-Robbins, F. S. (1925). Blood regeneration in severe anemia. II. Favorable influence of liver, heart and skeletal muscle in diet. *Am. J. Physiol.*, *72*(3), 408–418. Scopus. https://doi.org/10.1152/ajplegacy.1925.72.3.408

Willows, R. D., & Hansson, M. (2003). 77—Mechanism, Structure, and Regulation of Magnesium Chelatase. In K. M. Kadish, K. M. Smith, & R. Guilard (Eds.), *The Porphyrin Handbook* (pp. 1–47). Academic Press. https://doi.org/10.1016/B978-0-08-092387-1.50007-2

Winter, C. A., & Mushett, C. W. (1950). Absence of toxic effects from single injections of crystalline vitamin B12. *Journal of the American Pharmaceutical Association*, *39*, 360–361.

Xia, W., Chen, W., Peng, W., & Li, K. (2015). Industrial vitamin B12 production by Pseudomonas denitrificans using maltose syrup and corn steep liquor as the cost-effective fermentation substrates. *Bioprocess and Biosystems Engineering*, *38*(6), 1065–1073. https://doi.org/10.1007/s00449-014-1348-5

Young, T. R., Martini, M. A., Foster, A. W., Glasfeld, A., Osman, D., Morton, R. J., Deery, E., Warren, M. J., & Robinson, N. J. (2021). Calculating metalation in cells reveals CobW acquires CoII for vitamin B12 biosynthesis while related proteins prefer ZnII. *Nature Communications*, *12*(1), 1195. https://doi.org/10.1038/s41467-021-21479-8

Zayas, C. L., & Escalante-Semerena, J. C. (2007). Reassessment of the Late Steps of Coenzyme B12 Synthesis in Salmonella enterica: Evidence that Dephosphorylation of Adenosylcobalamin-5′-Phosphate by the CobC Phosphatase Is the Last Step of the Pathway. *Journal of Bacteriology*, *189*(6), 2210–2218. https://doi.org/10.1128/JB.01665-06

Zhang, C., & Hong, K. (2020). Production of Terpenoids by Synthetic Biology Approaches. *Frontiers in Bioengineering and Biotechnology*, *8*. https://doi.org/10.3389/fbioe.2020.00347

Zhang, J.-H., Yuan, H., Wang, X., Dai, H.-E., Zhang, M., & Liu, L. (2021). Crystal structure of the large subunit of cobaltochelatase from Mycobacterium tuberculosis. *Proteins: Structure, Function, and Bioinformatics*, *89*(4), 462–467. https://doi.org/10.1002/prot.26023

Zhang, K., Sawaya, M. R., Eisenberg, D. S., & Liao, J. C. (2008). Expanding metabolism for biosynthesis of nonnatural alcohols. *Proceedings of the National Academy of Sciences*, *105*(52), 20653–20658. https://doi.org/10.1073/pnas.0807157106

Zhang, X., & Bremer, H. (1995). Control of the Escherichia coli rrnB P1 Promoter Strength by ppGpp*. *Journal of Biological Chemistry*, *270*(19), 11181–11189. https://doi.org/10.1074/jbc.270.19.11181

Zhang, Y. (2009). *New round of price slashing in Vitamin B12 Sector—Zhang yemei China Chemical Report No.3 Jan 26th 2009*. http://blog.tianya.cn/blogger/post_show.asp?BlogID=2968076&PostID=24434221