Title
Genetic and Biochemical Origins of Diversity in Cobamides: Nature's Most Beautiful Cofactors

Permalink
https://escholarship.org/uc/item/8vc6k0vs

Author
Crofts, Terence Spencer

Publication Date
2013

Peer reviewed|Thesis/dissertation

Genetic and Biochemical Origins of Diversity in Cobamides: Nature's Most Beautiful Cofactors

By

Terence Spencer Crofts

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Microbiology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Michiko E. Taga, Chair
Professor Michelle Chang
Professor Krishna Niyogi
Professor Patricia C. Zambryski

Fall 2013

Abstract

Genetic and Biochemical Origins of Diversity in Cobamides: Nature's Most Beautiful Cofactors

by

Terence Spencer Crofts

Doctor of Philosophy in Microbiology

University of California, Berkeley

Professor Michiko E. Taga, Chair

The cobamide class of small molecules includes the essential micronutrient cobalamin (Vitamin B_{12}). Cobamides catalyze a variety of reactions involving one-carbon transfers and radical mediated molecular rearrangements. These reactions are vital parts of biochemical pathways found in both humans (one-carbon metabolism, fatty acid metabolism) and prokaryotes (methanogenesis, acetogenesis, fermentation and bioremediation).

These chemically difficult reactions are made possible by the complex structure of cobamides. Cobamides are tetrapyrroles related to heme, chlorophyll and Factor F_{430}, but are unique in containing a cobalt atom that binds an upper and lower ligand. The upper ligand of cobamide cofactors is made up of a cobalt-carbon covalent bond and varies depending on the metabolic role of the cofactor. The lower ligand can vary widely and is largely responsible for the diversity of cobamides found in nature. Lower ligands can take the form of benzimidazoles (5,6-dimethylbenzimidazole in the case of cobalamin), purines (such as adenine), or phenolics (such as *p*-cresol).

The first chapter of my thesis will provide more background into the functions and structures of cobalamin and other cobamides. I will focus on the known diversity of cobamides in nature and how that diversity is the result of the metabolisms of cobamide producing organisms.

The next three chapters detail a series of experiments that I have performed to expand our knowledge of how cobamide diversity is maintained. The first set of experiments (Chapter 2) describes how the CobT enzyme, the first enzyme in the lower ligand attachment pathway, is responsible for determining which lower ligands are incorporated into cobamides. I examined this at the organismal and genetic levels and collaborated on a companion project examining these mechanisms at the enzymatic level. I also demonstrated that altering this specificity can have deleterious effects on the host organism. In Chapter 3, building on the previous chapter, I and a collaborator in the laboratory (Amrita Hazra) conducted a series of experiments to show that some CobT enzymes also specify the orientation of the cobamide lower ligand. Organismal, genetic, and biochemical analyses showed that CobT enzymes and cobamide producing bacteria can synthesize pairs of products as structural isomers. These findings add a new potential source

of structural variability in cobamides and reveal that CobT specificity extends to the orientation as well as the identity of the lower ligand. Finally, Chapter 4 describes a bioassay I developed to detect benzimidazole lower ligands in the environment and the results of its application to a variety of samples. I found that benzimidazoles are ubiquitous in the environment, a finding that has important implications for cobamide diversity in microbial communities.

Across these experiments I have explored the means by which cobamide diversity is maintained in nature through the actions of CobT and the availability of lower ligand bases. The final chapter summarizes my findings and places them in the context of the field. This section also contains novel hypotheses generated by my results about the significance of cobamide diversity in nature as well as future experiments to test these hypotheses.

Table of Contents

Acknowledgements

During the past several weeks of writing this dissertation, I have been reminded of all the help and effort others have given me. Of these I need to first thank Michi Taga for being a great mentor and advisor. When I joined the lab, Michi gave me a chance to figure out what I wanted to do and she has been very supportive ever since. She taught me how to be a careful experimenter and a better writer. I am also very grateful for the freedom she gave me by allowing me to choose what questions I wanted to answer while also always providing guidance. I also want to thank the other members of my qualifying exam and dissertation committees: Bob Buchanan, David Zusman, Pat Zambryski, and Michelle Chang. I received a lot of valuable feedback from meetings with this group of professors and their suggestions (e.g. Gibson cloning) improved the pace and quality of my research. I also want to acknowledge my undergraduate advisor, David Kranz, for first letting me join his lab and showing me how fun research can be.

My fellow Taga lab researchers have also been great companions and sources of valuable advice, both technical and philosophical. I have been very lucky to have had access to the experience and imagination of Kenny Mok, Amrita Hazra, Erica Dirks-Seth, and Andrew Han. Amrita and Erica have been great to collaborate with and Kenny has always been ready to help me with whatever experimental question I had. I also want to thank my collaborators in the Lisa Alvarez-Cohen lab, especially Shan Yi and Yujie Men for all their help with LCMS. Finally I have also had the pleasure of working with several dedicated undergraduate students. Andrea Oneto, Jasmine Aimua and especially Shelley Shi and Vadim Osadchiy have all provided hours of valuable assistance in the lab.

This research also would not have been possible without the help of other members of the PMB department. I'd especially like to thank Dana Jantz and Rocio Sanchez for the incredible job they have done to keep things running smoothly and shielding me from hours of dealing with the Berkeley bureaucracy. I also want to thank my fellow PMB'ers, especially those in the Lindow lab, including in particular Russell Scott, who were always happy to lend a reagent or give advice and Lisa Kronstad for her experiment-saving cloning tricks.

Finally, my family has continually been a source of support these past years. My parents were the first to get me interested in science and have always supported me in everything I do: even now I am eating food from a dissertation writing care package sent by them. I have also appreciated the kind words and support from my two sisters and grandmother, as well as my sister-, mother-, and father-in law. Lastly, my wonderful wife Caterina Gratton has been the most important factor in getting me to this point. In addition to her emotional support, throughout graduate school she has always led the way for me. Her company and belief in me made pursuing my PhD an even more enjoyable experience.

Dedicated, of course, to my wife Caterina C. Gratton.
She has always encouraged and supported me.
Without her love, and help with plating bacteria, I would not be here.

Chapter 1

Overview of cobamide history, function, and diversity

1.1   Introduction

Cobalamin, or vitamin B_{12} (Figure 1.1A), has at various points been called a pigment of life (Battersby, 2000), nature's most beautiful cofactor (Stubbe, 1994), and the largest natural product (a title it has since lost) (Roth et al., 1993). I will provide a short summary of the properties of cobalamin, but will focus primarily on cobalamin analogs. Cobalamin and its related cofactors are together termed cobamides, which are themselves part of the corrinoid family of tetrapyrroles (Figure 1.2A). Research on this group of compounds has been largely overshadowed by cobalamin. Here, I will review research on cobamide diversity and its biological significance.

Cobalamin and other related cobamides belong to the tetrapyrrole family of molecules which include heme, chlorophyll and coenzyme F_{430} among others (Figure 1.2B). Cobamides differ in several ways from these molecules, most notably in containing a contracted ring (Figure 1.2B), coordinated upper and lower axial ligands (Figure 1.2A), and the presence of a nucleotide loop to covalently anchor the lower axial ligand (Figure 1.1A) (Battersby, 2000). In addition, cobamides are produced solely by prokaryotes, making cobalamin the only vitamin that humans cannot obtain from plants (Roth et al., 1996). Among naturally occurring cobamides, structural variation has been observed in three areas. First, the β-axial ligand, or upper ligand, varies depending on the reaction being catalyzed. The cofactor forms of cobalamin contain either a methyl or an adenosyl group as the upper ligand. When methylcobalamin is utilized in an enzymatic reaction, the methyl group is donated and needs to be replaced. Conversely, in adenosylcobalamin the adenosyl group is regenerated each catalytic cycle and as such, this form is often referred to as the coenzyme form. The reactions that utilize these two forms of the cofactor are discussed later in this chapter. When in solution it is possible for a water molecule to take the place of the upper ligand, forming hydroxocobalamin. Finally, during purification of cobalamin it is common for the compound to pick up a cyanide ligand to become cyanocobalamin which is the vitamin form of cobalamin (Figure 1.2A). The other two areas in which variation are observed are the α-axial, or lower, ligand and the presence or absence of methyl group 177 (Figure 1.1A) which are covered in more detail later.

1.2   Brief history of cobalamin

The history of cobalamin encompasses almost 100 years of scientific progress, beginning in 1926 with the discovery by Minot and Murphy that an extract from liver could treat the disease known as pernicious anemia (Minot and Murphy, 1926), now known to be an autoimmune disease that destroys Intrinsic Factor, the cobalamin receptor located in the stomach (Stabler, 1999). This contribution towards the study and treatment of anemia earned the Nobel Prize in medicine and physiology in 1934 for Minot, Murphy, and Whipple. This finding began the race to purify the underlying small molecule. With the development of the still used *Lactobacillus lactis* bioassay to detect cobalamin, Rickes *et al* from Merck were able to isolate the pure compound from liver

in 1948 followed soon after by Smith and Parker at Glaxo laboratories (Rickes et al., 1948a, 1948b; Smith and Parker, 1948).
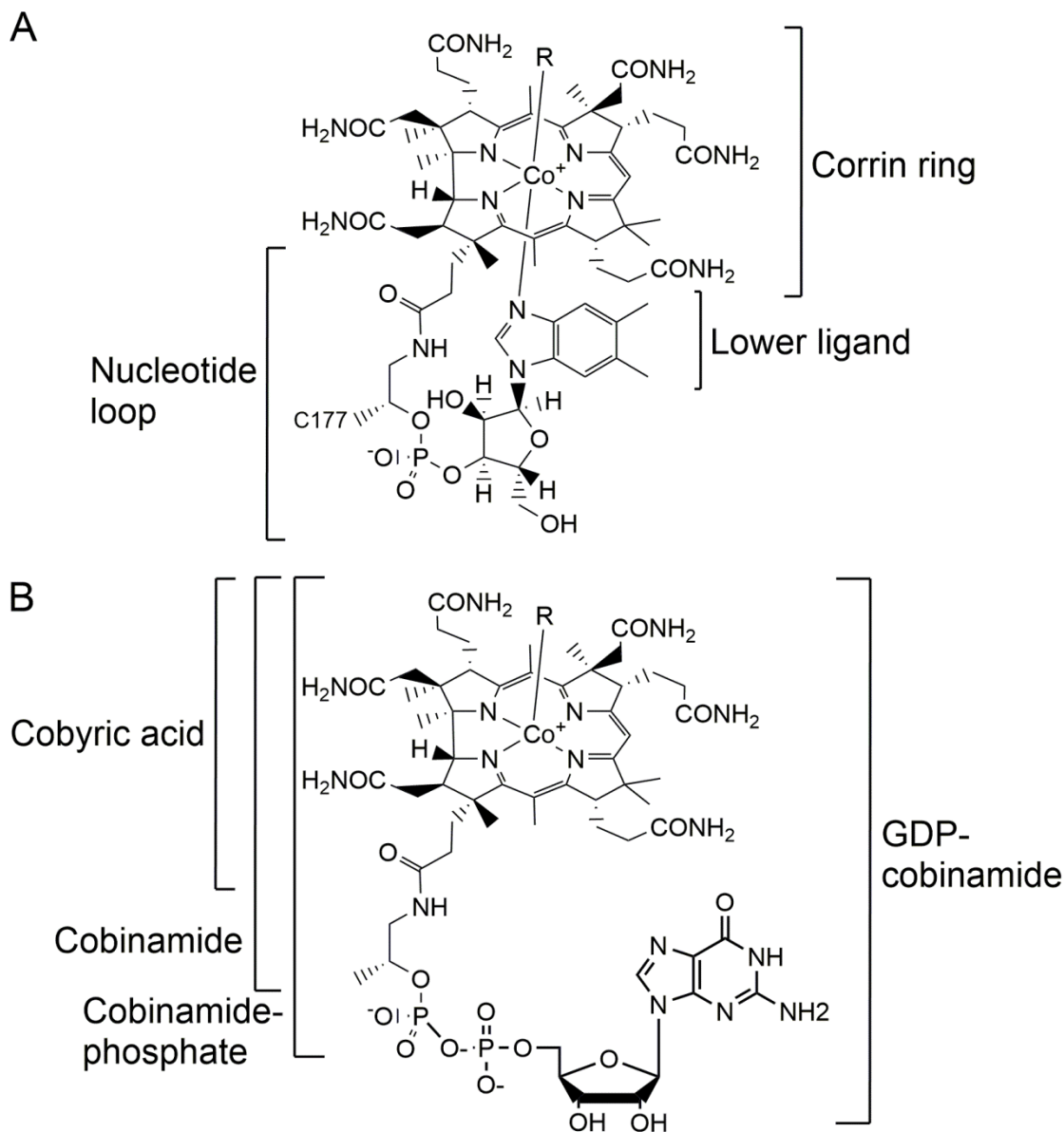

Figure 1.1 Structure of cobalamin and its precursors. A) The structure of cobalamin is shown, with the central corrin ring, nucleotide loop, and lower ligand indicated. The α-axial ligand, or lower ligand, of cobalamin is 5,6-dimethylbenzimidazole (DMB). The β-axial ligand, or upper ligand, is variable and shown here as R. In the vitamin form of cobalamin R is a cyano group, in the cofactor forms of cobalamin R is a methyl or an adenosyl group. Carbon 177, missing in the norcobamides, is labeled. B) The structure of guanosine diphosphate-cobinamide (GDP-Cbi), an incomplete corrinoid that is the immediate precursor to cobalamin and other cobamides. Also indicated are the biosynthetic precursors to GDP-Cbi: Cobyric acid, cobinamide (Cbi), and Cbi-phosphate.

The availability of a pure compound opened the door to the analytical phase of cobalamin research. This phase of cobalamin history also contains some of the great achievements of chemistry, notably in the fields of organic synthesis and structure determination. First, the

elucidation of the three dimensional crystal structure of cyanocobalamin by Dorothy Hodgkin was recognized by the 1964 Nobel Prize in chemistry (Hodgkin et al., 1956). Knowledge of the structure of cobalamin aided in the *de novo* synthesis of cobalamin by Woodward and Eschenmoser, a feat that to this day stands as a crowning achievement of organic synthesis and contributed towards Woodward's 1965 Nobel Prize in chemistry (Woodward, 1973). The development of new techniques for this synthesis also led, in part, to the creation of the influential Woodward-Hoffman rules (Battersby, 2000) which were recognized by the 1981 Nobel Prize in chemistry.
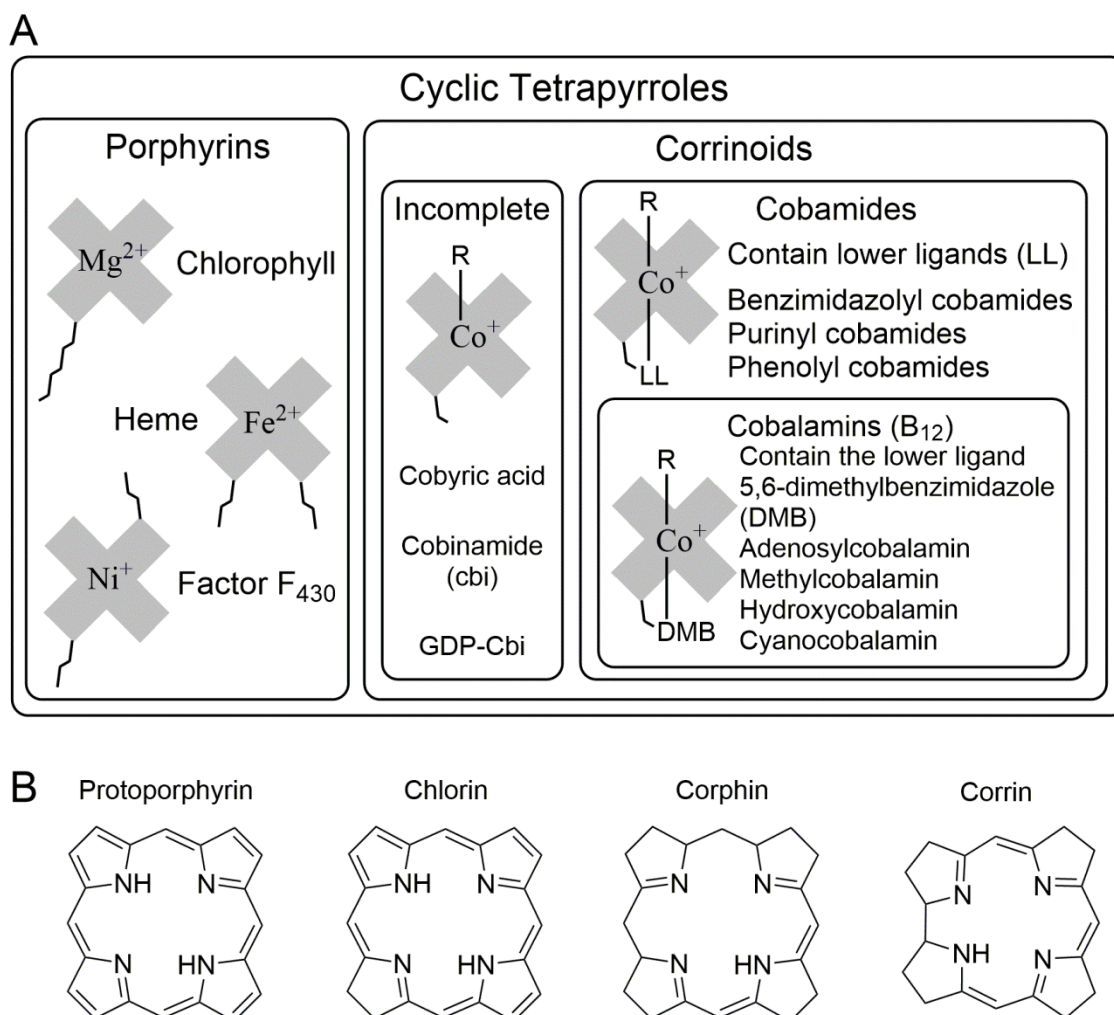
Figure 1.2 Corrinoids and other cyclic tetrapyrroles. A) A Venn diagram illustrating the relationships between cobalamin, cobamides, corrinoids, and other tetrapyrroles. Corrinoids containing both upper and lower ligands are termed cobamides, while cobalamin refers specifically to cobamides which have 5,6-dimethylbenzimidazole (DMB) as the lower ligand. Other cobamides contain benzimidazole bases (e.g. 5-hydroxybenzimidazole), purine bases (e.g. adenine), or phenolic bases (e.g. *p*-cresol). Cobamides are named based on their lower ligand base (e.g. [lower ligand]Cba). B) Examples of ring structures of cyclic tetrapyrroles. From left to right: Protoporphyrin ring found in heme, chlorin ring found in chlorophyll a, corphin ring found in factor F_{430}, corrin ring found in cobalamin.

1.3   Cobalamin-dependent reactions

While many could argue that the sheer size and complexity of cobalamin has justified its study – it has been called the Mt. Everest of biosynthesis (Blanche et al., 1995) – it is also an essential cofactor for humans, other animals, and many prokaryotes. Cobalamin-catalyzed reactions can be divided into three broad classes: the isomerases, methyl transferases, and dehalogenases (Figure 1.3).

The isomerase reactions are generally catalyzed by adenosylcobalamin, the cofactor form in which the upper ligand is an adenosyl group (Figure 1.2A). During catalysis, the cobalt-carbon bond of adenosylcob(III)alamin is broken homolytically, generating an adenosyl radical and cob(II)alamin (note, the roman numeral notation signifies the oxidation state of the cobalt atom). This radical can abstract an electron from a substrate carbon or hydrogen atom (Gruber et al., 2011). The resulting substrate radical can then undergo radical mediated skeletal rearrangement. At the end of these rearrangements, the radical is generally regenerated on the adenosyl group and from there adenosylcob(III)alamin is formed once again (Figure 1.3A). An important component of these reactions is control of the radical. In the absence of substrate, the adenosyl radical will generally recombine with cob(II)alamin but even after the formation of a substrate radical the tendency is to recombine. Therefore, control of the reaction depends on so-called conformational toggling of the radical substrate to move the radical away from the adenosyl moiety (Banerjee and Ragsdale, 2003).

Isomerase enzymes can be divided into five classes that differ in their substrates, the conformation in which they bind adenosylcobalamin, and how the radical is propagated. First are the skeleton mutases, an example of which is methylmalonyl-CoA mutase, shown in Figure 1.3A. These enzymes bind cobalamin in a conformation in which the coordinate bond of the lower ligand is broken (the "base-off" conformation) and is replaced by the imidazole group of a histidine residue on the enzyme (the "His-on" conformation) (Figure 1.3D). Mutases are involved in a variety of metabolisms, including the fermentation of lactic acid by the bacterium *Veillonella parvula* (Ng and Hamilton, 1971) and glutamate degradation by *Clostridium tetanomorphum* (Barker et al., 1960a). In contrast, isomerases belonging to the diol dehydratase, and ethanolamine lyase family bind cobamides in the "base-on" conformation in which the lower ligand remains coordinated to the cobalt ion (Figure 1.3D). These two classes of enzymes catalyze molecular rearrangement reactions resulting in the liberation of water and ammonia respectively (Gruber et al., 2011). *Salmonella enterica*, as an example, utilizes a 1,2-propanediol dehydratase and ethanolamine ammonia lyase when growing on 1,2-propanediol and ethanolamine, respectively (Roth et al., 1996). The fourth type of isomerase, aminomutase, binds cobalamin in the base-off/his-on conformation similar to the skeleton mutases. Lysine-5,6-aminomutase and Orthinine-4,5-aminomutase are used for the metabolism of lysine and ornithine, respectively, in *Clostridium sticklandii* (Brown, 2005; Stadtman, 1960). The final isomerase enzyme is the ribonucleotide reductase which catalyzes the conversion of nucleotides into deoxynucleotides (Banerjee and Ragsdale, 2003). The bacteria *Sinorhizobium meliloti* and *Lactobacillus leichmannii* utilize this cobalamin-dependent enzyme (Cowles and Evans, 1968; Lawrence and Stubbe, 1998).
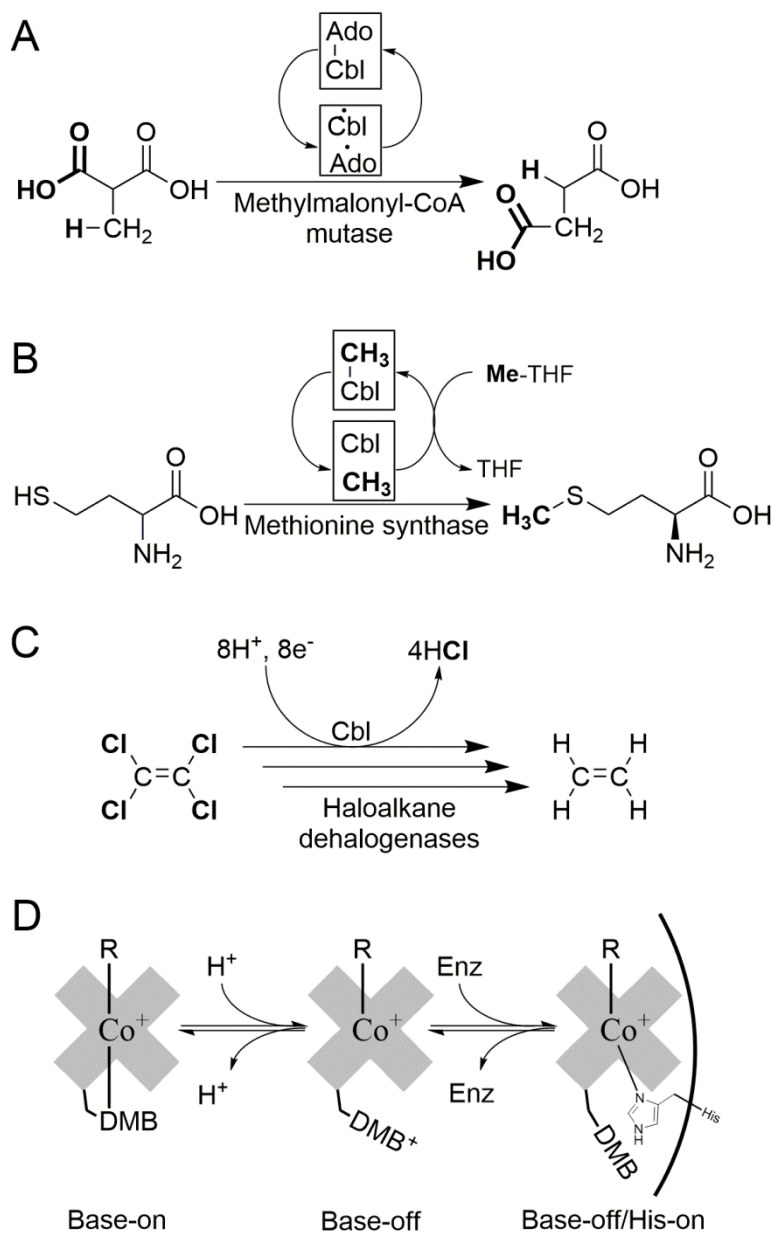
Figure 1.3 Broad classes of cobalamin-dependent reactions. A) The reaction catalyzed by methylmalonyl-CoA mutase, one of the cobalamin-dependent enzymes found in humans, is an example of a cobalamin-mediated isomerase reaction. The groups involved are shown in bold. (Cbl, cobalamin). B) Shown is a one-carbon transfer reaction mediated by methionine synthase, a second cobalamin-dependent enzyme found in humans. The transferred methyl group, originally donated by 5-methyltetrahydrofolate (Me-THF), is marked in bold. C) Shown is the cobalamin-dependent reductive dehalogenation of perchloroethylene to ethene. Only certain anaerobic bacteria are capable of performing these reactions. D) Illustration of three conformations of cobalamin: Base-on (DMB coordinated), Base-off (DMB not coordinated), Base-off/His-on (enzyme histidine coordinated).

The second broad class of reactions that require cobalamin involve one-carbon transfers (Figure 1.3B). Unlike the rearrangement reactions, these generally make use of methylcobalamin (Figure 1.2A), the cofactor form in which the upper ligand is composed of a CH_3 group. Also unlike some isomerases, the enzymes in this category all bind cobalamin in the base-off conformation

(Gruber et al., 2011). Generally, two methyl transferase enzymes are involved in cobalamin-dependent one-carbon transfers, the first to transfer the methyl group from the donor compound to cobalamin, and the second to transfer the methyl group from cobalamin to the substrate. During catalysis, methylcob(III)alamin donates its methyl group, formally as a carbocation, to a substrate, resulting in cob(I)alamin. A methyl donating molecule is then required to reproduce methylcob(III)alamin. In the case of methionine synthase the methyl group originates from 5-methyltetrahydrofolate (Figure 1.3B). Besides the synthesis of methionine, one-carbon transfers are also involved in methanogenesis and acetogenesis (Banerjee and Ragsdale, 2003; Ragsdale and Pierce, 2008). In these processes cobalamin takes on the role of a one-carbon shuttle, for example transferring methyl groups to factor F_{430} for reduction in methanogenesis (Gruber et al., 2011). Examples of methanogens and acetogens include *Methanobacterium thermoautotrophicum*, *Methanosarcina barkeri*, *Moorella thermoacetica*, and *Sporomusa ovata* (Kräutler et al., 1987; Pol et al., 1982; Stupperich et al., 1988; Wurm et al., 1980).

The final class of cobalamin-dependent enzymes is the reductive dehalogenase family (Figure 1.3C). These enzymes have not been as thoroughly characterized as the isomerases and methyltransferases, owing in part to the membrane localization of many of these proteins and slow growth of dehalogenating bacteria (Banerjee and Ragsdale, 2003). The form of cobalamin used in dehalogenating reactions is not known. One proposed mechanism begins with the oxidation of cob(II)alamin to cob(I)alamin in order to produce a radical that can attack a chlorinated substrate. Unlike in the isomerases, the radical is not proposed to regenerate cob(II)alamin (Banerjee and Ragsdale, 2003). A well-studied dehalogenating bacterium is *Dehalococcoides mccartyi* (He et al., 2007; Kräutler et al., 2003).

1.4   The biosynthesis of cobalamin

As previously mentioned, cobalamin, and cobamides in general, are only synthesized by prokaryotes. Two pathways have been described for the *de novo* synthesis of cobalamin, termed the aerobic and anaerobic pathways, each of which is comprised of approximately 30 genes (Roth et al., 1996). The majority of the enzymes of the aerobic pathway were determined in *Pseudomonas denitrificans* by the Crouzet group at Rhone-Poulec. This was accomplished through complementation of over 200 different cobalamin auxotrophs in three bacterial species using a plasmid library of the *P. denitrificans* genome. Once four complementation groups were discovered, analysis of individual genes and purified proteins could begin, resulting in the characterization of 22 genes (Blanche et al., 1995). However, the final step of the aerobic pathway was not resolved until 2006/2007 when the BluB enzyme of bacterium *Sinorhizobium meliloti* was found (by Michi Taga, Graham Walker and colleagues) to be responsible for the synthesis of 5,6-dimethylbenzimidazole (DMB), the lower ligand of cobalamin (Campbell et al., 2006; Taga et al., 2007). The function of BluB homologs in DMB biosynthesis was confirmed in other bacteria as well (Collins et al., 2013; Gray and Escalante-Semerena, 2007).

Meanwhile, the anaerobic pathway of cobalamin biosynthesis, with the exception of the anaerobic route to DMB, has also been solved. Studies by the Roth group in *Salmonella enterica* (Roth et al., 1993), the Scott group in *Propionibacterium freudenreichii* (Roessner et al., 2002), and multiple groups in *Bacillus megaterium* (Biedendieck et al., 2010; Brey et al., 1986; Moore
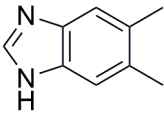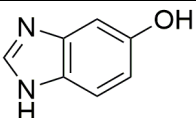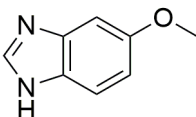
et al., 2013; Wolf and Brey, 1986) have all provided valuable insights into the assembly of cobalamin in the absence of oxygen. Briefly, the two pathways differ in their tolerance of or requirement for oxygen and at what step in the pathway the cobalt ion is inserted into the macrocycle. In both cases, 5-aminolevulinic acid is the building block to produce uroporphyrinogen III (UroIII), the shared precursor of all cyclic tetrapyrroles. In the aerobic route of cobalamin biosynthesis the modifications to UroIII involved in corrin ring synthesis (methylations, ring contraction, addition of amide groups) mostly occur prior to the insertion of the cobalt atom and require oxygen or are oxygen tolerant (Roth et al., 1996). In contrast, the anaerobic route inserts cobalt early in the pathway with most of the corrin ring modifications performed on the cobalt-containing form (Warren et al., 2002).
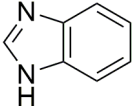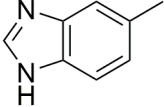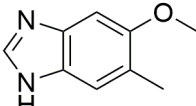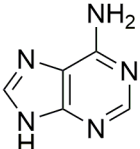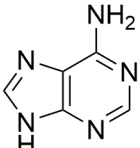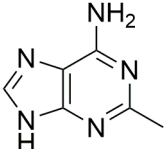
1.5   Cobalamin is only one representative of a large family of cobamide cofactors
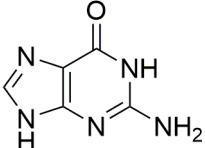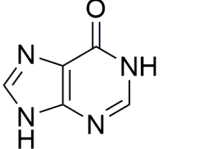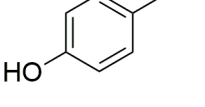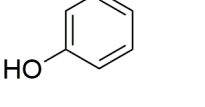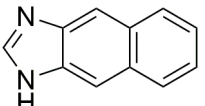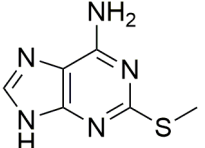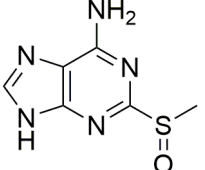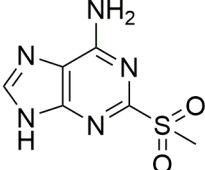
The human cobamide requirement appears to be specific to cobalamin. However, despite its importance to our health, cobalamin makes up only the minority of cobamides in our guts and several other environments (Allen and Stabler, 2008; Brown et al., 1955; Girard et al., 2009; Renz et al., 1987). The remainder of this chapter will focus on these "alternative" cobamides. Since the discovery of cobalamin, many additional cobamides have been found in nature, or they have been generated in the lab by feeding exogenous lower ligand bases to cobamide producing prokaryotes, a process known as guided biosynthesis (Perlman, 1959, 1971). In Table 1.1, I summarize the current state of knowledge of many of these cobalamin analogs. This table provides references for the discovery, structure, and physical characterization of known cobamides, and some physiological data on the organisms that produce them.

In addition to naturally producing a variety of cobamides, several species of bacteria are amenable to guided biosynthesis, mentioned earlier. This is the process in which an exogenous lower ligand can be introduced into the media of a culture and be taken up by a cobamide*-producing organism (Perlman, 1971). In high enough concentrations, the exogenous lower ligand base can out-compete the native lower ligand resulting in the production of a new cobamide. *Propionibacterium* species have been commonly used for the biosynthesis of a variety of cobamides containing benzimidazole, purine, and imidazole lower ligands (Kamikubo and Matsuura, 1969; Perlman and Barrett, 1958) though a variety of other organisms can also be used (Allen and Stabler, 2008; Gray and Escalante-Semerena, 2009a; Mok and Taga, 2013; Perlman, 1971). However, it should be noted that additional factors beyond availability of a lower ligand base can limit the utility of guided biosynthesis in producing cobamides, as will be discussed below.

Table 1.1 Characteristics of naturally occurring cobamides and the prokaryotes that synthesize them

Cobamide (Lower ligand in parentheses)	Lower ligand structure	Notable cobamide producers (* original source)	Associated metabolisms	Chemical characterization
Cobamides extracted from pure bacterial cultures				
Cobalamin (DMB) aka B$_{12}$		*Streptomyces griseus* (Rickes et al., 1948b), *Propionibacterium freudenreichii*, *Rhodopseudomonas protamicus*, *Propionibacterium shermanii*, *Pseudomonas denitrificans*, *Nocardia rugosa*, *Rhizobium cobalaminogenum*, *Micromonospora* sp., *Streptomyces olivaceus*, *N. gardneri*, *Butyribacterium methylotrophicum*, *Pseudomonas* sp., *Arthrobacter hyalinus*, (Martens et al., 2002) *Sinorhizobium meliloti* (Campbell et al., 2006), *Eubacterium limosum* (Lamm et al., 1982)	Various	X-ray crystal structure (Crowfoot-Hodgkin, 1965; Hodgkin et al., 1956; Rossi et al., 1985), NMR (Summers et al., 1986; Tollinger et al., 1999)
[5-OHBza]Cba (5-hydroxybenzimidazole) aka Factor III		*Methanobacillus omelianskii* (Lezius and Barker, 1965), *Methanosarcina barkeri*, *Methanobacterium thermoautotrophicum* and other methanogens (Kräutler et al., 1987; Pol et al., 1982; Stupperich and Kräutler, 1988)	Methanogenesis	NMR (Kräutler et al., 1987), UV-Vis (Lezius and Barker, 1965)
[5-OMeBza]Cba (5-methoxybenzimidazole) aka Factor IIIm		*Moorella thermoacetica* (Irion and Ljungdahl, 1965; Ljungdahl et al., 1965)	Acetogenesis	NMR (my work, Chapter 3), UV-Vis (Irion and Ljungdahl, 1965)

[Bza]Cba (benzimidazole)		*V. parvula* (my work)	Lactic acid fermentation	UV-vis (Barker et al., 1960b; Ladd et al., 1961)
[5-MeBza]Cba (5-methylbenzimidazole)		*Methanothrix soehngenii* (Kohler, 1988), *Desulfobulbus propionicus* and other sulfate-reducers, *Thermoplasma acidophilum* (Kräutler et al., 1988)	Sulfate-reduction and sulfur utilization	NMR, UV-Vis (Kräutler et al., 1988)
[5-OMe-6-MeBza]Cba (5-methoxy-6-methylbenzimidazole)		*Clostridium formicoaceticum* (Stupperich et al., 1988)	Acetogenesis	NMR, UV-vis (Stupperich et al., 1988)
[Ade]Cba (adenine) aka pseudocobalamin, pseudo-B_{12}		*Clostridium tetanomorphum* (Barker et al., 1958), *Salmonella enterica* (Keck and Renz, 2000), *Lactobacillus reuteri* (Santos et al., 2007), *Clostridium cochlearium* (Hoffmann et al., 2000), *Clostridium sticklandii* (Stadtman, 1960), *Methanococcus voltae* and other methanogens (Stupperich and Kräutler, 1988), *Synechocystis* sp. and other cyanobacteria (Tanioka et al., 2009)	Fermentation, methanogenesis, photosynthesis	X ray crystal structure (Kräutler et al., 2003), NMR (Hoffmann et al., 2000; Stupperich and Kräutler, 1988), UV-vis (Weissbach et al., 1960)
Nor[Ade]Cba (adenine) aka norspseudocobalamin, norpseudo-B_{12}		*Sulfurospirillum multivorans* (Kräutler et al., 2003)	Dehalogenation	X-ray crystal structure, NMR, UV-Vis (Kräutler et al., 2003)
[2-MeAde]Cba (2-methyladenine) aka Factor A		*S. enterica* (Keck and Renz, 2000), *C. cochlearium* (Hoffmann et al., 2000)	Fermentation	X-ray crystal structure (Kräutler et al., 2003), NMR (Hoffmann et al., 2000), UV-Vis

[Gua]Cba (guanine)	O, NH, N, NH, N, NH$_2$	*Nocardia* sp. (Barchielli et al., 1957), *Desulfovibrio vulgaris* (Guimarães et al., 1994)	Sulfate-reduction	NMR, UV-Vis (Guimarães et al., 1994)
[Hxa]Cba (hypoxanthine)	O, NH, N, N, N, H	*D. vulgaris* (Guimarães et al., 1994)	Sulfate-reduction	NMR, UV-Vis (Guimarães et al., 1994)
[Cre]Cba (*p*-cresol)	HO	*Sporomusa ovata* (Stupperich et al., 1988), *V. parvula* (my work, Chapter 2)	Acetogenesis, lactate fermentation	NMR, UV-Vis (Stupperich et al., 1988)
[Phe]Cba (phenol)	HO	*S. ovata* (Stupperich et al., 1989), *V. parvula* (my work, Chapter 2)	Acetogenesis, lactate fermentation	NMR, UV-Vis (Stupperich et al., 1989)
Cobamides found only as environmental isolates				
[Nza]Cba (naphthimidazole)	N, N, H	Sewage sludge (Renz, 1999)		
[2-SMeAde]Cba (2-methylthioadenine)	NH$_2$, N, N, N, N, H, S	Sewage sludge (Friedrich and Bernhauer, 1957)		UV-vis (Friedrich and Bernhauer, 1957)
[2-SOMeAde]Cba (2-methylsulfinyladenine)	NH$_2$, N, N, N, N, H, S, O	Sewage sludge (Renz et al., 1987)		UV-vis (Renz et al., 1987)
[2-SO$_2$MeAde]Cba (2-methylsulfonyladenine)	NH$_2$, N, N, N, N, H, S, O, O	Sewage sludge (Renz et al., 1987)		UV-vis (Renz et al., 1987)

1.6 Biochemical and genetic origins of cobamide diversity

One unique aspect of the cobamide family of cofactors is their structural diversity. The diversity in cobamides is limited to the lower ligand and to the presence or absence of carbon 177 in the nucleotide loop (Figure 1.1A). A given cobamide-producing organism will consistently synthesize the same type of cobamide when grown in pure culture, indicating that cobamide structure is genetically encoded. It is therefore possible to look at a limited subset of cobamide biosynthesis enzymes and their substrates as sources of cobamide structural diversity.

1.6.1   Norcobamides: B$_{12}$ analogs missing carbon 177

Though the majority of cobamide diversity is the result of diversity in the lower ligand base, the norcobamides represent a unique sub-class of cobamides. These cobamides differ from the majority of other cobamides in the absence of a methyl group on the nucleotide loop at position C177 (Figure 1.1A). Since cobamides are characterized by the identity of their lower ligand, in theory any cobamide could have a norcobamide doppelganger. In practice, only noradeninylcobamide (nor[Ade]Cba), the nor version of the cobamide with adenine (Ade) as its lower ligand, has been found in nature (Table 1.1). The organism responsible for the biosynthesis of this cobamide is *Sulfurospirillum multivorans*, which utilizes cobamides in its dehalogenating metabolism. Through guided biosynthesis, this same organism has been induced to synthesize norcobalamin and could possibly be used, in a similar manner, to produce a variety of other norcobamides (Keller et al., 2013).

While the exact mechanism by which *S. multivorans* produces nor[Ade]Cba is still unknown, the origin of C177 in other cobamides is. Labeling studies demonstrated that the amino acid L-threonine is the source of the aminopropanol portion of the nucleotide loop of cobamides (Krasna et al., 1957) (Figure 1.1B). It has been demonstrated genetically and biochemically in *S. enterica* that L-threonine is converted to L-threonine-phosphate via the PduX kinase (Fan and Bobik, 2008; Fan et al., 2009) and is then decarboxylated by the CobD enzyme to form aminopropanol-phosphate (Brushaber et al., 1998; Grabau and Roth, 1992). Finally, CbiB attaches aminopropanol-phosphate to cobyric acid to produce cobinamide-phosphate (Figure 1.1B) (Woodson et al., 2003; Zayas et al., 2007). Intriguingly, evidence has been given that when grown in the presence of ethanolamine-phosphate, *S. enterica* can produce norcobalamin and nor[Ade]Cba. Cobamides were produced under these conditions with the expected masses of cobalamin and [Ade]Cba minus a methyl group, implying that the ethanolamine-phosphate was incorporated rather than aminopropanol-phosphate through a type of guided biosynthesis (Zayas et al., 2007). It is open to speculation by what exact mechanism *S. multivorans* produces norcobamides. It is possible that serine rather than threonine is fed into the pathway (Kräutler et al., 2003), though other mechanisms have not been excluded. However it is achieved, the removal of carbon 177 impacts the chemical (*i.e.*, reduction potential) as well as structural nature (*i.e.*, base conformation) of cobamides (Butler et al., 2006; Kräutler et al., 2003).

1.6.2   The lower ligand: Site of most cobamide diversity

By far the most significant site of cobamide diversity is in the lower ligand. The identity of the lower ligand is dependent on multiple factors. First, I will review what is currently known about the biosynthetic origins of lower ligand bases and then I will discuss the processes by which lower ligand bases are incorporated into cobamides.

1.6.2.1 Benzimidazole origins

As the lower ligand of cobalamin, 5,6-dimethylbenzimidazole (DMB) has been most extensively studied. DMB is the only member to date of the benzimidazole lower ligand bases for which a biosynthetic origin has been determined. However, DMB is produced both aerobically and anaerobically, and only the aerobic pathway is known. This pathway consists of the single

enzyme conversion of reduced flavin mononucleotide (FMNH$_2$) to DMB (Figure 1.4) (Campbell et al., 2006; Taga et al., 2007). This reaction is catalyzed by the BluB enzyme, a member of the nitroreductase/flavin oxidoreductase enzyme family based on sequence homology (Campbell et al., 2006; Taga et al., 2007). The complete mechanism by which BluB accomplishes this has not been determined yet, but the first step is the formation of a C4a-peroxyflavin intermediate (Figure 1.4) from the reaction of FMNH$_2$ and molecular oxygen (Collins et al., 2013; Taga et al., 2007; Yu et al., 2012). Relatedly, Maggio-Hall *et al.* have reported on a possible route for the abiotic synthesis of DMB from riboflavin in the presence of oxygen (Maggio-Hall et al., 2003).
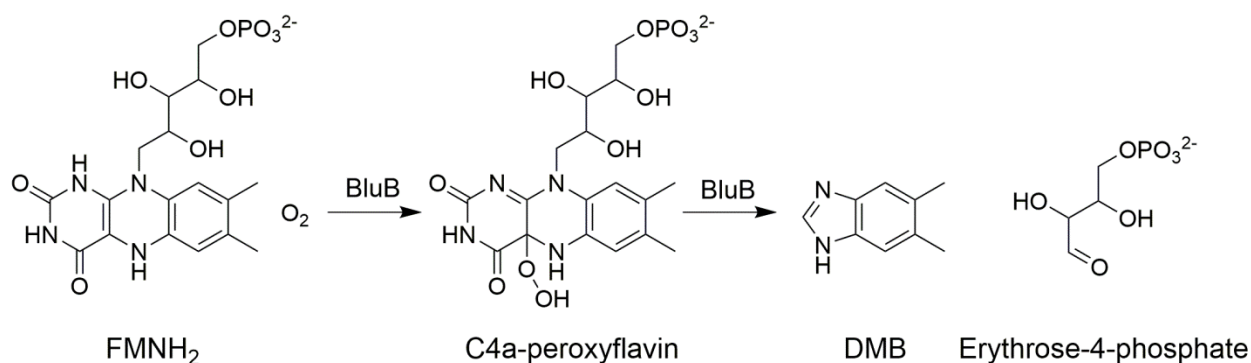
Figure 1.4 Aerobic biosynthesis of DMB, the lower ligand of cobalamin. The BluB enzyme converts reduced flavin mononucleotide (FMNH$_2$) to DMB, erythrose-4-phosphate, and an unknown co-product(s). The first intermediate of this reaction is a C4a-peroxyflavin generated by reaction with molecular oxygen.

Cobalamin and other cobamides with benzimidazole lower ligand bases (benzimidazolyl cobamides) have also been found in anaerobic environments and are made by a variety of anaerobic bacteria (Stupperich et al., 1990). The pathways by which these are synthesized are unknown, though meticulous labeling studies have shed light on the biosynthetic origins of the lower ligands. In particular, studies by Paul Renz and his laboratory in the bacterium *Eubacterium limosum* have identified the metabolic origins of each atom in DMB (Figure 1.5A). N1, C8, and C9 were found to originate from glycine (Lamm et al., 1982) and N3 from glutamine (Vogt and Renz, 1988). C2 comes from formate (Munder et al., 1992). The benzene ring is predominantly composed of carbons (4, 5, 6, and 7) derived from erythrose (Munder et al., 1992), though it appears this may not be universal to other benzimidazoles (Eisenreich and Bacher, 1991). Finally the methyl groups C10 and C11 both are derived from methionine (Lamm et al., 1980), a known methyl donor in central metabolism (Figure 1.5A). Interestingly, the DMB bound to cobalamin has been found to show regioselectivity (Hörig et al., 1978). Since the two nitrogen groups of DMB are chemically equivalent, it is an open question as to how it achieves consistent orientation. One possibility is that in this case CobT, the enzyme responsible for routing lower ligand bases into the cobamide biosynthetic pathway, acts on an asymmetric DMB precursor. Another possibility is that after its biosynthesis, DMB is shuttled by an enzyme to CobT, thus retaining a particular orientation. The CobT family of enzymes and their roles are discussed later.
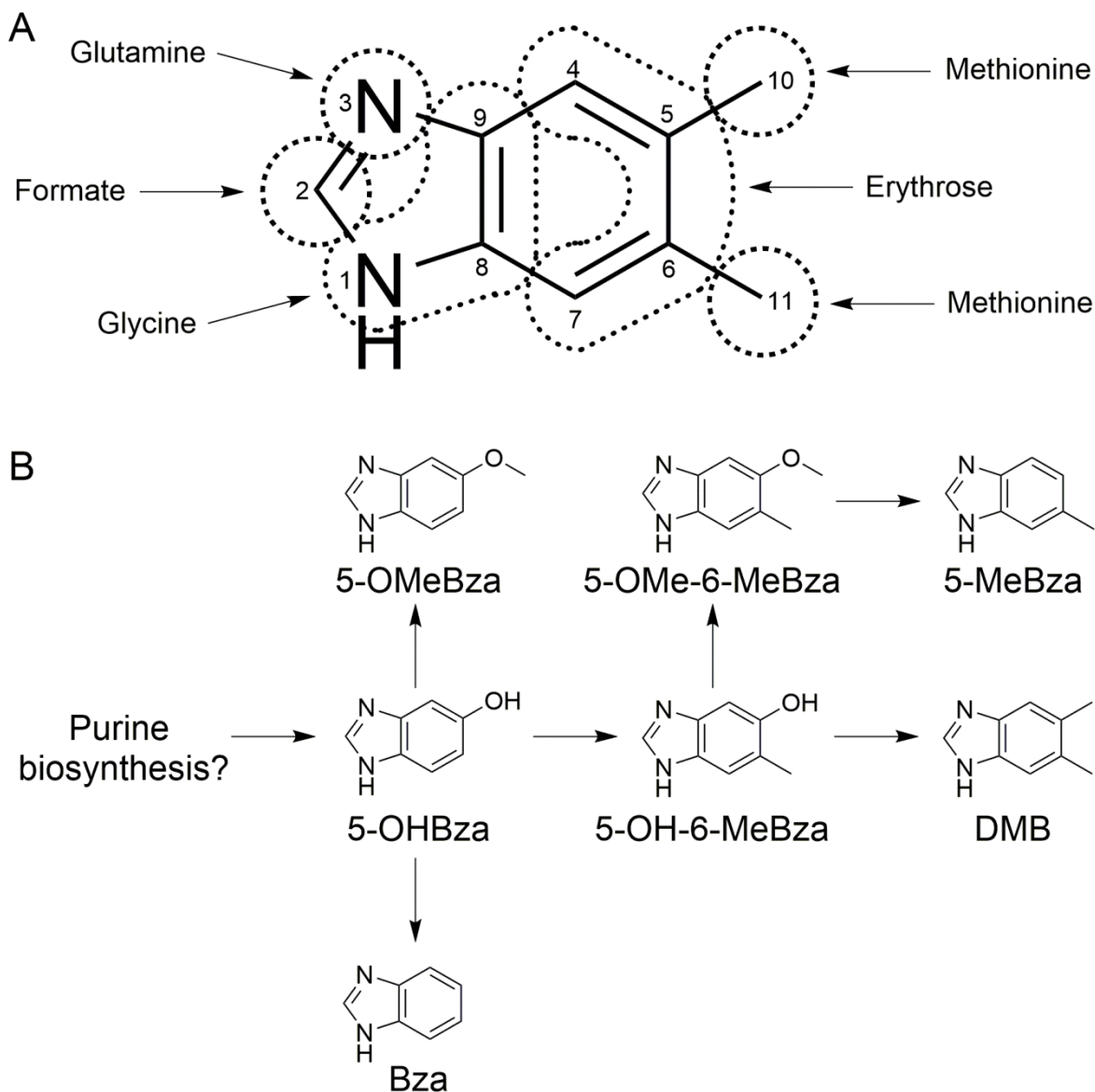
Figure 1.5 Anaerobic biosynthesis of DMB and other benzimidazole lower ligand bases. A) The anaerobic biosynthesis of DMB. The enzymes and mechanisms by which DMB is formed in the absence of oxygen are not known, but the metabolic precursors of DMB are. Glycine, formate, glutamine, erythrose, and methionine each contribute atoms as shown by the dotted lines. B) 5-hydroxybenzimidazole is thought to originate from the purine biosynthesis pathway and to be the precursor molecule to the other anaerobically produced benzimidazoles. Abbreviations: 5-hydroxybenzimidazole (5-OHBza); 5-hydroxy-6-methylbenzimidazole (5-OH-6-MeBza); DMB; 5-methoxybenzimidazole (5-OMeBza); 5-methoxy-6-methylbenzimidazole (5-OMe-6-MeBza); 5-methylbenzimidazole (5-MeBza); benzimidazole (Bza).

Several insights into the anaerobic biosynthesis of benzimidazole bases other than DMB have also been made (see Figure 1.5B and Table 1.1 for structures). For example, 5-hydroxybenzimidazole (5-OHBza) and 5-methoxybenzimidazole (5-OMeBza) have been shown to be formed in a similar manner as DMB from similar metabolites (Eisenreich and Bacher, 1991; Scherer et al., 1984; Wurm et al., 1980). Renz and colleagues again provided key insights

into the biosynthesis of the benzimidazoles and their apparent interconvertability (Figure 1.5B) by again taking advantage of the amenability of cobamide-producing bacteria for guided biosynthesis experiments. For example, it has been demonstrated that 5-OHBza is converted into 5-OMeBza in *Moorella thermoacetica* (Wurm et al., 1975), with the methyl group being provided by methionine (Wurm et al., 1980). 5-OHBza is additionally converted into 5-methoxy-6-methylbenzimidazole (5-OMe-6-MeBza) and DMB in *E. limosum* in what is thought to be a related process (Renz et al., 1993). Finally, while not definitively shown, it is likely that the methanogen *Methanothrix soehngenii* produces 5-methylbenzimidazole (5-MeBza) from 5-OHBza (Kohler, 1988). In contrast, when unsubstituted benzimidazole (Bza) was fed to cobamide-producing bacteria, no changes were observed, implying that Bza is not a precursor of other lower ligands (Scherer et al., 1984; Wurm et al., 1975). It has also been debated whether Bza isolated from the environment is of bacterial origin at all (Friedmann and Cagen, 1970). My experiments in Chapter 4 touch on this. To summarize these findings, it is proposed that 5-OHBza is the parent compound of the benzimidazoles, some of which are further converted into DMB (such as 5-OH-6-MeBza), while others represent metabolic dead ends (such as Bza). These findings are summarized in Figure 1.5B (Renz et al., 1993).

1.6.2.2 Purine origins

Besides benzimidazoles, cobamides are also found with purine-derived lower ligands and phenolic lower ligands. Adenine and [Ade]Cba (Dion et al., 1952), are perhaps the best studied lower ligand/cobamide pair after DMB/cobalamin. The biosynthesis of adenine, guanine, and hypoxanthine are already well characterized, as is their presence, in a variety of systems. However, what remains less clear is which specific sources of purines are utilized for cobamide biosynthesis.

One possible answer to this question is provided by a study in *S. enterica*, an organism that produces mostly [Ade]Cba in the absence of DMB (Keck and Renz, 2000). When DMB is excluded from the growth medium, *S. enterica* uses [Ade]Cba in place of cobalamin for all of its cobalamin-dependent processes (Anderson et al., 2008), though growth on different substrates requires different levels of cobalamin. Specifically, growth on 1,2-propanediol, using propanediol dehydratase, does not require the addition of DMB for the production of [Ade]Cba, though additional DMB is required for ethanolamine ammonia lyase dependent growth on ethanolamine. By selecting for growth of *S. enterica* on ethanolamine without exogenous DMB, Anderson *et al* isolated mutants that produced higher levels of [Ade]Cba. Interestingly, all the mutants found appeared to act by increasing the levels of free adenine, either by lowering the activity of adenine phosphoribosyltransferase or by increasing activity of AMP nucleosidase. The result of these mutations was an increase in free adenine in the cell which served as a source of lower ligands for the biosynthesis of [Ade]Cba. The presence of free adenine pools in another bacterium, *Sinorhizobium meliloti*, will be addressed in Chapter 2. The implication of these findings is that there is no requirement for a dedicated pool of adenine for it to be used as a lower ligand, though this pool may be limiting under conditions requiring high cobamide activity.

Adenine is not the only purine lower ligand found in cobamides. *S. enterica*, in addition to producing [Ade]Cba, also produces a cobamide termed Factor A ([2-methyladeninyl]Cba, [2-MeAde]Cba) (Table 1.1), in which the lower ligand base is 2-methyladenine (2-MeAde) (Dion et

al., 1954; Keck and Renz, 2000). [2-MeAde]Cba has long been known to contain this unique purine, but the origin of the lower ligand base is not well established (Dion et al., 1954). 2-MeAde has been found in ribonucleic acid extracts (Littlefield and Dunn, 1958) and 2-MeAde, as well as other purine derivatives, are known to be components of some tRNAs (Saneyoshi et al., 1972). One of these additional purines is 2-methylthioadenosine (Harada et al., 1968). When the sugar component of this molecule is removed it becomes 2-methylthioadenine (2-SMeAde) which has been reported as another cobamide lower ligand (Table 1.1) (Friedrich and Bernhauer, 1957). Renz *et al.* have also reported the extraction and purification of related cobamides with 2-methylsulfinyladenine and 2-methylsulfonyladenine as lower ligand bases (Renz et al., 1987) (Table 1.1). However, it is believed that these two cobamides actually originated as [2-SMeAde]Cba but were converted during the purification process (Renz, 1999).

The final purine containing cobamides that have been found to occur in nature come from the sulfate-reducing bacterium *Desulfovibrio vulgaris* (Guimarães et al., 1994) and from a *Nocardia* strain (Barchielli et al., 1957). *D. vulgaris* was found to produce both guanylcobamide ([Gua]Cba) and hypoxanthylcobamide ([Hxa]Cba in which guanine and hypoxanthine are the lower ligands respectively (Table 1.1). [Gua]Cba was the only cobamide reported in the *Nocardia* strain. Feeding of labeled guanine to *D. vulgaris* resulted in the production of both labeled [Gua]Cba and labeled [Hxa]Cba, implying that guanine is the source of hypoxanthine as a lower ligand (Guimarães et al., 1994).

Unlike the benzimidazoles, the purine lower ligands are not uniquely utilized by cobamide biosynthesis, but rather appear to be scavenged from unrelated metabolic functions (in the cases of adenine and guanine), or other cellular components (*e.g.* 2-MeAde and 2-SMeAde from tRNAs).

1.6.2.3 Phenolic origins

The final class of lower ligand bases that have been found in nature is the phenolics. These are represented by *p*-cresol and phenol (Table 1.1), both of which are naturally utilized as lower ligand bases by the bacterium *Sporomusa ovata* to produce *p*-cresolylcobamide ([Cre]Cba) and phenolylcobamide ([Phe]Cba) (Stupperich et al., 1988, 1989). Until recently, *S. ovata* was the only known producer of these cobamides, though now it has been shown that another *Sporomusa* species shares this ability (Yan et al., 2013). In addition, I found that another member of the *Veillonellaceae* family, *Veillonella parvula*, produces phenolyl cobamides as well (see Chapter 2). Feeding studies have been carried out in *S. ovata* to partially characterize the origin of the phenolic bases. Stupperich and Eisinger provided growing *S. ovata* cultures with labeled *p*-cresol, tyrosine, and glycine, as well as other compounds (Stupperich and Eisinger, 1989). Unsurprisingly, labeled *p*-cresol was readily incorporated into a cobamide as the lower ligand. As no label was found elsewhere, including in the protein fraction, it was concluded that *p*-cresol is not further transformed (Stupperich and Eisinger, 1989). On the other hand, labeled tyrosine was found both in the protein fraction as well as in the lower ligand portion of [Cre]Cba. Whether labeled [Phe]Cba was produced as well was not addressed in this study. Stupperich *et al* also performed a study in which *S. ovata* was provided 4-fluorophenol (Stupperich et al., 1993). In this study once again the label was only found associated with the cobamide, in this case *p*-

fluorophenolylcobamide. Together these studies point to tyrosine, and perhaps other aromatic amino acids, as the biosynthetic origin of the phenolic lower ligand bases. Furthermore, it appears that phenol is not converted into *p*-cresol, though it is unknown if the reverse is true. Like the purine lower ligand bases, the phenolic bases appear to be normal cellular metabolites that have been commandeered as cobamide lower ligands.
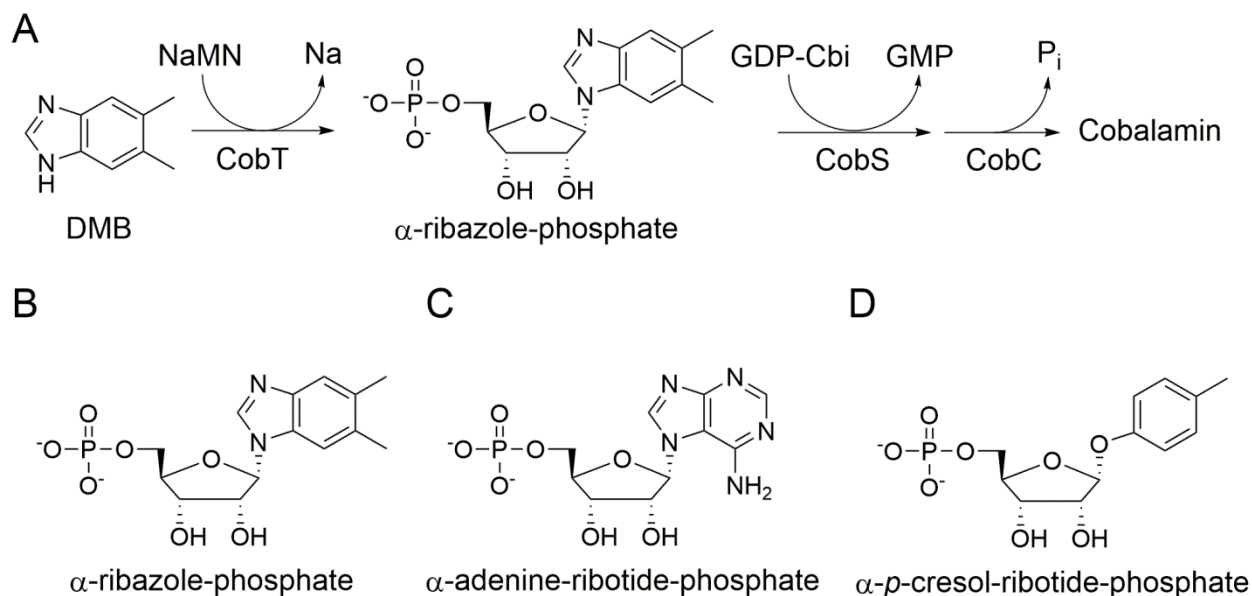
Figure 1.6 CobTSC reaction and α-ribotide-phosphate structures. A) The reactions catalyzed by CobT, CobS, and CobC. CobT attaches a phosphoribose moiety, donated by nicotinate mononucleotide (NaMN), to an unmodified lower ligand base (shown here as DMB). Nicotinic acid (Na) and an α-ribotide-phosphate are produced (shown here is α-ribazole-phosphate). CobS attaches the α-ribotide-phosphate group to GDP-cobinamide (Fig 1B), releasing GMP, and CobC removes the phosphate group to produce inorganic phosphate (P_i) and a complete cobamide (cobalamin in this example). B), C) and D) show structures of α-ribotide-phosphates from the benzimidazole (DMB), purine (adenine), and phenolic (*p*-cresol) classes of lower ligand bases, respectively.

Finally, a mention should be made about the divergent physical characteristics of the phenolic cobamides. A notable difference between the phenolic lower ligand bases and the benzimidazoles and purines is the presence of only one heteroatom. Furthermore, unlike the benzimidazoles and purines which contain nitrogen atoms, the phenolic lower ligand bases contain oxygen atoms in the form of hydroxyl groups. The absence of a second heteroatom in phenol and *p*-cresol means that there are no additional lone pair electrons available to coordinate with the cobalt atom in the corrin ring, meaning that cobamides containing these lower ligands are permanently in the base-off conformation (Figure 1.3D). Additionally, the lack of nitrogen atoms means that these lower ligand bases are covalently bound to the nucleotide loop through O-glycosidic bonds (Figure 1.6D) rather than N-glycosidic bonds (Figure 1.6BC) (Mok and Taga, 2013; Stupperich et al., 1988, 1989, 1990). The lack of lower ligand coordination makes phenolyl cobamides catalytically inactive in reactions that require base-on cobamides (see previous section on cobalamin dependent reactions).

1.6.3   Lower ligand activation and attachment

While there exist both an anaerobic and aerobic pathway for cobamide biosynthesis, these routes diverge during ring synthesis but converge again once synthesis of the corrin ring (Figure 1.1A) has been completed. The cobamide precursor at this point is most notably lacking a nucleotide loop (Figure 1.1A) (Roth et al., 1996). The nucleotide loop, which anchors the lower ligand base to the corrin ring, is assembled principally by the proteins CbiB, discussed above, and CobUTSC (as they are named in *S. enterica*) (Figure 1.6A). CobU, the adenosylcobinamide kinase and adenosylcobinamide-phosphate guanylyltransferase, has two functions in this pathway. It phosphorylates the aminopropanol tail of adenosylcobinamide and subsequently attaches a GMP group derived from GTP. In parallel CobT, the lower ligand phosphoribosyl transferase, attaches a phosphoribose moiety to an unmodified lower ligand base to make a lower ligand α-ribotide-phosphate. The products of CobU and CobT are combined by CobS, the cobalamin synthase, and CobC removes the phosphate group from the now attached phosphoribose to produce a cobamide (Figure 1.6A) (Maggio-Hall and Escalante-Semerena, 1999; Roth et al., 1996; Zayas and Escalante-Semerena, 2007).

1.6.3.1 CobS and CobC: The final steps of cobamide biosynthesis

Of all of the enzymes involved in cobamide biosynthesis, only CobT, CobS, and CobC interact with the lower ligand base. As CobT uses as its substrate the unmodified lower ligand base it has attracted the most attention for its role in determining cobamide structure (Friedmann, 1965) and will be examined in more detail below. The substrate specificities of CobS and CobC have been most carefully studied by the laboratory of Escalante-Semerena (Maggio-Hall and Escalante-Semerena, 1999; Zayas and Escalante-Semerena, 2007). In this case the focus was not on the identity of the lower ligand, but rather on whether CobS acted immediately following CobT, or after the dephosphorylation reaction of CobC. The authors found that the reaction kinetics of CobS were more favorable with α-ribazole-phosphate than with α-ribazole (Figure 1.6AB), implying that CobS acts before CobC (Blanche et al., 1995).

Other studies have indirectly measured the specificity of these two enzymes towards different lower ligand bases. *S. enterica* CobS and CobC can utilize both DMB-containing (Figure 1.6B) and adenine-containing (Figure 1.6C) α-ribotide-phosphates (Anderson et al., 2008). Others have found that, similarly, *S. enterica* has the genetic capability to utilize α-*p*-cresol-ribotide-phosphate (Fig 6D) and α-phenol-ribotide-phosphate, implying that its CobS and CobC proteins tolerate a wide range of substrates (Chan and Escalante-Semerena, 2011). Additional work of mine in Chapter 2 addresses this as well. Finally it should be mentioned that there are cases where CobS and CobC might play more of a role in determining cobamide structure. The proceeding chapters address some of these cases.

1.6.3.2 CobT: The lower ligand phosphoribosyl transferase

The reaction and substrate specificity of CobT have been studied for several decades. In the 1960's, the first α-phosphoribosyl transferase reactions (Figure 1.6A) were described and were immediately hypothesized to underlie the ability of certain organisms to incorporate diverse

lower ligand bases through guided biosynthesis (Friedmann, 1965; Friedmann and Fyfe, 1969; Friedmann and Harris, 1965; Fyfe and Friedmann, 1969). These studies, performed with cell extracts, were the first to demonstrate that CobT homologs may have different activities and may help determine cobamide structure. Specifically, the enzyme from *Propionibacter freudenreichii* was found to utilize both DMB and other benzimidazoles as substrates (Friedmann and Harris, 1965) but not adenine (Friedmann, 1965). This finding was in agreement with previous studies that had shown *P. freudenreichii* to produce cobamides with benzimidazole lower ligand bases but not [Ade]Cba. In contrast, *Clostridium sticklandii* produces [Ade]Cba and is capable of also producing cobamides with benzimidazole lower ligands. Friedmann and Fyfe demonstrated that the CobT enzyme isolated from this organism can utilize both adenine and benzimidazoles as substrates (Friedmann and Fyfe, 1969; Fyfe and Friedmann, 1969). The *cobT* gene was first isolated and described by Cameron *et al* from *Pseudomonas denitrificans*, termed *cobU* in the aerobic route to cobamide biosynthesis (Cameron et al., 1991), but *S. enterica* CobT has been studied in the most detail.

While earlier *in vitro* studies dealt predominantly with activation of DMB or adenine, the ability of *S. enterica* CobT to activate and/or bind a wide variety of substrates has been assayed by biochemical methods and X-ray crystallography (Cheong et al., 1999, 2001, 2002; Trzebiatowski and Escalante-Semerena, 1997). These studies indicated that *in vitro* the *S. enterica* CobT enzyme has very high substrate promiscuity, not only binding a variety of benzimidazoles and purines, but also apparently activating imidazole, 4,5-dimethyl-1,2-phenylenediamine, and histidine. The major implication of these findings is that CobT shows little substrate specificity. However, in the proceeding chapter I will present evidence that CobT homologs appear to have the ability to activate many different types of lower ligand bases *in vitro*, but under physiologically relevant conditions they constitute a strong barrier against the production of improper cobamides.

Recently an additional CobT homolog has been discovered and characterized (Chan and Escalante-Semerena, 2011; Newmister et al., 2012). Termed ArsAB by the Escalante-Semerena group, this enzyme belongs to the CobT enzyme family but, unlike canonical CobT enzymes which function as homodimers, ArsAB is a heterodimer. Interestingly, the only two organisms known to produce phenolyl cobamides, *S. ovata* and *V. parvula*, both carry homologs of *arsA* and *arsB* as tandem, overlapping open reading frames (Chan and Escalante-Semerena, 2011; Crofts et al., 2013; Gronow et al., 2010; Stupperich et al., 1988, 1989). Notably, phenol and *p*-cresol are the only two naturally occurring lower ligand bases for which no evidence exists of activation by the *S. enterica* CobT (Cheong et al., 2001; Crofts et al., 2013; Hazra et al., 2013). Expression of *S. ovata arsAB* in *S. enterica* was sufficient to allow it to produce phenolyl cobamides when provided the suitable lower ligand bases (Chan and Escalante-Semerena, 2011). It has been hypothesized by us and others that the unique genomic structure of these homologs may allow researchers to predict the presence of phenolic cobamides in a microbial community based on the presence of these genes (Chan and Escalante-Semerena, 2011; Crofts et al., 2013).

Finally, with a variety of CobT homologs sequenced and cloned, researchers have naturally begun to combine the knowledge of the structures of *S. enterica* CobT and *S. ovata* ArsAB with protein sequence alignments to predict substrate specificity of these enzymes. Some of these

mutants, dealing with the ability of *S. enterica* to produce cobalamin, are found in the proceeding chapter. Recently, Chan *et al.*, making similar mutations, revealed that mutating *S. enterica* CobT Ser80Tyr, Gln88Met, and Leu175Met together was sufficient to allow it to activate phenolic lower ligand bases *in vitro* (Chan et al., 2013). Further studies along these lines may allow for the development of strains of cobamide-producing bacteria with increased ability to make cobalamin or other cobamides of interest.

1.6.4   *Cobamide salvaging and remodeling*

An additional stage at which cobamide diversity is determined occurs after its biosynthesis. When some cobamide-producing bacteria are provided foreign cobamides they will convert it to a different cobamide by removing and replacing the lower ligand, a process called remodeling (Stupperich et al., 1987). More recently, corrinoid remodeling has been documented in the human gut microbial community. Following ingestion of high doses of labeled cobalamin, the stool of human subjects was analyzed for its cobamide content. Unexpectedly, the majority of the ingested cobalamin was converted in the gut to new cobamides, predominantly those containing a purine lower ligand (Allen and Stabler, 2008). In this context, remodeling of cobalamin led to an increase in the cobamide diversity of the community. In addition, our laboratory recently demonstrated the ability of the bacterium *Dehalococcoides mccartyi* to remodel a variety of cobamides, including those containing benzimidazole, purine, and phenolic lower ligands into cobalamin (Yi et al., 2012). It is likely in fact that cobamide salvaging and remodeling has a great effect on cobamide diversity in microbial communities. Indeed, it is known, based on genome sequencing data, that approximately 50% of the prokaryotes that require cobamides do not produce them and rely instead on cobamide import (Zhang et al., 2009).

The key step in cobamide remodeling is the removal of the lower ligand from the nucleotide loop (Figure 1.1A). The enzyme responsible for this process, the aminohydrolase CbiZ, was first studied in the archaea *Methanosarcina mazei* (Woodson and Escalante-Semerena, 2004). This enzyme was found to catalyze the removal of the nucleotide loop from cobamides to produce cobyric acid (Figure 1.1B). The removal of the nucleotide loop allows the organism to process cobyric acid through the cobamide biosynthesis pathway and install its own lower ligand. Although bioinformatics analysis of CbiZ points to an archaeal origin, it has been found in the genomes of a variety of bacteria (Gray and Escalante-Semerena, 2010; Gray et al., 2008; Yi et al., 2012). The role of CbiZ in remodeling cobamides is illustrated well by the cobamide-dependent metabolisms of *Rhodobacter sphaeroides* and *D. mccartyi*. *R. sphaeroides* requires cobamides when grown on acetate, with both cobalamin and [Ade]Cba being suitable. However, a *cbiZ* mutant strain, while still able to grow when provided cobalamin or Cbi (Figure 1.1B) and DMB, no longer grew using [Ade]Cba (Gray and Escalante-Semerena, 2009a). A functional *cbiZ* gene is required under these conditions in order to remove the adenine lower ligand and replace it with DMB. A similar dynamic likely takes place in *D. mccartyi*, a dechlorinating bacterium that our laboratory has shown to remodel a variety of cobamides from all three structural classes (Yi et al., 2012).

Finally, the ability of cobamide-producing organisms to take up lower ligand bases during guided biosynthesis and cobamide remodeling indicates that, for organisms that do not produce

their own lower ligand, environmentally available lower ligands are one more source of cobamide diversity. This is especially relevant for lower ligand bases, such as the benzimidazoles, that are not part of the core metabolism (in contrast to the purine and phenolic lower ligand bases which are produced endogenously in most cells). In fact, our lab and its collaborators have recently observed the presence of a variety of free benzimidazoles in a mixed microbial community that includes *D. mccartyi* (Y. J. Men, E. C. Seth, S. Yi, T. S. Crofts, R. H. Allen, M. E. Taga, and L. Alvarez-Cohen, submitted). I have also developed a bioassay for benzimidazoles that is the subject of Chapter 4. In addition to lower ligand bases, it appears likely that α-ribotides are also available in the environment. A pair of proteins from *Listeria innocua*, CblT and CblS, has been described to be an α-ribazole importer and an α-ribazole kinase respectively (Gray and Escalante-Semerena, 2010). This is especially significant in *L. innocua* as the genome of this organism encodes all the genes required for *de novo* cobamide biosynthesis except for *cobT*, which is otherwise required for the synthesis of α-ribazole-phosphate. It appears that *L. innocua* compensates by importing α-ribazole using CblT and routing it into the nucleotide loop assembly pathway by phosphorylating it with CblS.

1.7   Significance of cobamide variation

An unanswered question in the field is why do organisms use DMB as a lower ligand? When produced aerobically with BluB the price of one molecule of DMB is a molecule of flavin mononucleotide, a derivative of another vitamin, B_2 (Taga et al., 2007). When produced anaerobically it is estimated that biosynthesis of DMB would require several additional enzymes (Renz, 1999). Both of these routes presumably come with an energetic cost that would be lessened if adenine or some other non-specialized lower ligand were used. It is worth considering some features of DMB, and other benzimidazoles in some cases, that may provide insights into why they are produced. First, as noted above, DMB can be formed from flavin mononucleotide through the action of the BluB enzyme (Taga et al., 2007), but also apparently spontaneously, in the case of *S. enterica* among others, under aerobic conditions (Keck et al., 1998). A mechanism for the spontaneous formation of DMB from riboflavin under physiological conditions has been proposed (Maggio-Hall et al., 2003). An abiotic origin for DMB from existing cellular components makes its initial use as a lower ligand plausible but does not explain its staying power.

Another notable characteristic of DMB that separates it from almost all other lower ligand bases is the presence of an axis of symmetry. Besides DMB, only benzimidazole and the non-coordinating phenolic lower ligand bases share this characteristic. This could be of use in cobamide biosynthesis because lower ligand attachment would require less specificity in this case as both nitrogens in DMB are chemically identical (Ford et al., 1955; Friedmann and Fyfe, 1969). It has been observed in a variety of cobamide producing organisms that utilize asymmetric lower ligand bases that some percentage of cobamide, up to 20% in some cases, carries the lower ligand base in the opposite orientation (Ford et al., 1955; Friedmann and Fyfe, 1969; Friedrich and Bernhauer, 1958; Kohler, 1988; Kräutler et al., 1987, 1988; Kräutler, 1987; Perlman, 1959; Stupperich et al., 1993). Chapter 3 of this dissertation is dedicated to describing the biosynthesis and characteristics of these cobamide isomers.

Finally, some of the chemical characteristics of DMB differ from adenine. First, DMB is more hydrophobic, containing a benzene ring and only two heteroatoms, while the purine ring of adenine contains multiple heteroatoms. Second, the amino group of adenine provides it with the ability to form additional hydrogen bonds. It has been noted that when interacting with a cobamide-dependent enzyme the presence or lack of this amino group could greatly affect binding (Hoffmann et al., 2000). Finally, the tendency of the lower ligand to coordinate with the central cobalt atom differs, with [Ade]Cba more likely to assume a base-off conformation than cobalamin (Hoffmann et al., 2000; Ladd et al., 1961). As noted earlier in this chapter, different enzymes that require cobamides have differing requirements as to what conformations are needed.

As Table 1.1 demonstrates, there is a great diversity of cobamides, with differences in the lower ligand bases accounting for the vast majority of this diversity. The question remains though, why are DMB and cobalamin not used universally? One possibility is that even if DMB is the optimal lower ligand for many bacteria, other lower ligand bases might be sufficient and more readily obtainable. This is well exemplified by *S. enterica*, which has possibly taken the cost-saving route of using readily available adenine or 2-MeAde as its lower ligand bases and does not have the ability to synthesize DMB (Keck and Renz, 2000). However, when presented with DMB, from agar contamination (Anderson et al., 2008) or abiotic degradation of flavins (Keck et al., 1998), *S. enterica* will preferentially attach DMB rather than adenine to make cobalamin. This is consistent with the observation that cobalamin appears to be a better cofactor for *S. enterica* than [Ade]Cba (Anderson et al., 2008). If DMB is available freely in some environments, as mentioned earlier, some bacteria may have lost the ability to synthesize DMB and instead salvage it from the environment. When available, DMB could be used to make cobalamin, but when unavailable an organism could make do with an energetically less costly but also less effective lower ligand base. My studies with *S. meliloti* provide a possible example for the reverse of this. *S. meliloti* has maintained its ability to produce DMB while its CobT homolog effectively shields this organism from utilizing adenine as a lower ligand as is discussed in great detail in the next chapter.

The story may be somewhat different for anaerobic cobamide producers. As mentioned earlier, it appears that DMB is the final product of a series of steps during which 5-OHBza is transformed through various benzimidazole intermediates (Renz et al., 1993) (Figure 1.5B). In this case, a tradeoff may have been made, where utilization of an intermediate or side product on this pathway, such as 5-OMeBza, provides sufficient cofactor activity but at a lower anabolic cost. For either of these two possibilities, adoption of a new lower ligand would be met with selective pressure, likely including ones that would encourage the diversification of CobT homologs to better fit available lower ligand bases. Thus *S. enterica* may experience a dearth of benzimidazole lower ligand bases and so must have greater substrate promiscuity to utilize purine lower ligand bases (Anderson et al., 2008; Trzebiatowski and Escalante-Semerena, 1997) while *D. mccartyi* appears to have sufficient benzimidazoles available to it (Y. J. Men, E. C. Seth, S. Yi, T. S. Crofts, R. H. Allen, M. E. Taga, and L. Alvarez-Cohen, in prep) so that it does not attach adenine (Crofts et al., 2013; Hazra et al., 2013; Yi et al., 2012).

Another, not mutually exclusive, explanation for the observed diversity of cobamides lies in game theory. Cobamide biosynthesis requires near 30 gene products (Warren et al., 2002) and is likely an energetic investment based on its close regulation in *de novo* producers (Rodionov et al., 2003; Roth et al., 1996). It is possible that one benefit of producing a different cobamide is that competing organisms in the same community may be unable to utilize it as a cofactor. If this is the case, the ability of cobamide-utilizing organisms to salvage environmental cobamides and cobamide precursors potentially represents an energetic advantage. This is potentially the case with *D. mccartyi*, a bacterium that requires cobamides from its environment for its metabolism but is unable to use some forms, notably phenolyl and purinyl cobamides (Yi et al., 2012). Our observations of a trichloroethylene dechlorinating enrichment community emphasizes this as [Cre]Cba is abundant, though *D. mccartyi* is capable of cobamide remodeling and can grow in this environment nonetheless (Y. J. Men, E. C. Seth, S. Yi, T. S. Crofts, R. H. Allen, M. E. Taga, and L. Alvarez-Cohen, submitted).

1.8   Conclusion

Historically the majority of cobamide research has been directed at cobalamin because of its direct implications for human health. More recently, as other areas of research have turned to microbial communities as new sources of metabolic diversity, or for their applications to human and ecological health, the diversity of cobamides has become more important. In the following chapters I will add to this my own findings demonstrating the subtleties of CobT substrate specificity under physiological conditions, characterize cobamide isomers with lower ligand bases attached in an unusual orientation, and reveal experimental evidence that benzimidazoles are available in a variety of environments.

Chapter 2

Cobamide structure depends on both lower ligand availability and CobT substrate specificity

2.1 Summary

Cobamides are members of the vitamin B12 family of cofactors that function in a variety of metabolic processes and are synthesized only by prokaryotes. Cobamides produced by different organisms vary in the structure of the lower axial ligand. Here, we explore the molecular factors that control specificity in the incorporation of lower ligand bases into cobamides. We find that the cobT gene product, which activates lower ligand bases for attachment, limits the range of lower ligand bases that can be incorporated by bacteria. Furthermore, we demonstrate that the substrate specificity of CobT can be predictably altered by changing two active site residues. These results demonstrate that sequence variations in cobT homologs contribute to cobamide structural diversity. This analysis could open new routes to engineer specific cobamide production and understand cobamide-dependent processes.

2.2 Introduction

Many environmentally and industrially important microbial processes such as methanogenesis, acetogenesis, and reductive dechlorination are dependent on corrinoids, a class of cofactors produced solely by a subset of prokaryotes (Banerjee and Ragsdale, 2003; Roth et al., 1996). Corrinoids are modified tetrapyrroles that contain a centrally bound cobalt atom (Roth et al., 1996). Corrinoids that contain an upper and lower ligand are termed cobamides. Cobalamin **1**, also known as vitamin B$_{12}$ (Figure 2.1A), is the best studied cobamide and is an essential micronutrient for most animals including humans (Roth et al., 1996). While humans are thought to have a specific requirement for cobalamin **1**, 16 different cobamides with structural variability in the lower ligand (Figure 2.1B) have been described (Renz, 1999). Most microbes described to date produce only one to two different cobamides when grown in pure culture (Keck and Renz, 2000; Kräutler et al., 1988; Stupperich et al., 1989, 1990). While the factors that limit the range of cobamides produced by an organism remain unclear, previous studies have shown that different cobamides may not be functionally equivalent as cofactors for cobamide-dependent enzymes (Barker et al., 1960a, 1960b; Chan and Escalante-Semerena, 2011; Ford et al., 1955; Mok and Taga, 2013; Yi et al., 2012).

The *de novo* biosynthesis of a cobamide requires approximately 30 gene products (Warren et al., 2002). The attachment of the lower ligand base to the cobamide precursor GDP-cobinamide (GDP-Cbi **10**) is the last step in cobamide biosynthesis. This process has been studied most extensively in *Salmonella enterica* (*S. enterica*) (Roth et al., 1996). Prior to attachment, the lower ligand base is activated by the CobT enzyme by the transfer of a phosphoribose moiety from nicotinate mononucleotide (NaMN) or a related compound to form an α-glycosidic linkage (Figure 2.1C) (Cameron et al., 1991; Friedmann, 1965; Friedmann and Harris, 1965; Trzebiatowski and Escalante-Semerena, 1997). The phosphoribosylated base is subsequently attached to GDP-Cbi **10** by CobS, and the phosphate group is removed by CobC to form the cobamide (Figure 2.1C) (Maggio-Hall and Escalante-Semerena, 1999; Roth et al., 1996; Zayas and Escalante-Semerena, 2007). Homologs of *cobT* are present in nearly all bacterial genomes that contain the complete corrinoid biosynthesis pathway (Rodionov et al., 2003). However, in organisms that produce cobamides by an aerobic pathway, the *cobT* homolog is termed *cobU* (Roth et al., 1996). Two additional *cobT* homologs, termed *arsA* and *arsB*, have been described in the bacterium *Sporomusa ovata* (*S. ovata*). These genes together encode a heterodimeric enzyme responsible for the activation of phenolic lower ligand bases (Chan and Escalante-Semerena, 2011).

In this work, we investigate the molecular factors that limit the range of cobamides that can be produced by a single bacterial species. Two hypotheses are tested to address this question. First, the range of lower ligands that can be attached to form cobamides may be limited by the substrate specificity of the CobT enzyme. Alternatively, the cobamides produced by an organism may be limited solely by the availability of potential lower ligands. Evidence in favor of the first hypothesis comes from our previous observations of corrinoid biosynthesis in the *Sinorhizobium meliloti* (*S. meliloti*) *bluB* mutant. *S. meliloti bluB* is unable to synthesize 5,6-dimethylbenzimidazole (DMB, the lower ligand of cobalamin **1**) and instead produces the incomplete corrinoid GDP-Cbi **10** rather than incorporating intracellular adenine (Ade) to produce pseudo-B$_{12}$ **6** (adeninyl cobamide, [Ade]Cba) (Campbell et al., 2006). Similarly, *Lactobacillus reuteri* does not produce cobalamin **1** when DMB is provided and instead produces only [Ade]Cba **6** (Santos et al., 2007). The second hypothesis is supported by previous studies demonstrating the ability of several bacteria to attach a variety of exogenously provided lower ligands (Ford et al., 1955; Kamikubo and Matsuura, 1969; Maggio-Hall and Escalante-Semerena, 1999; Perlman and Barrett, 1958; Trzebiatowski and Escalante-Semerena, 1997). For example, *S. enterica* produces [Ade]Cba **6** and 2-methyladenylcobamide **7** ([MeAde]Cba), and *S. ovata* produces phenolylcobamide **8** ([Phe]Cba) and *p*-cresolylcobamide **9** ([Cre]Cba) when cultured without a lower ligand base, but both can produce benzimidazolyl cobamides when benzimidazole bases are provided (Anderson et al., 2008; Keck and Renz, 2000; Mok and Taga, 2013; Newmister et al., 2012; Stupperich and Eisinger, 1989; Stupperich et al., 1989, 1990). In addition, X-ray crystallography studies of *S. enterica* CobT and *S. ovata* ArsAB have demonstrated the ability to bind a variety of lower ligand bases in their active sites (Cheong et al., 1999, 2001, 2002; Newmister et al., 2012). Currently the literature as a whole does not support one hypothesis over the other.
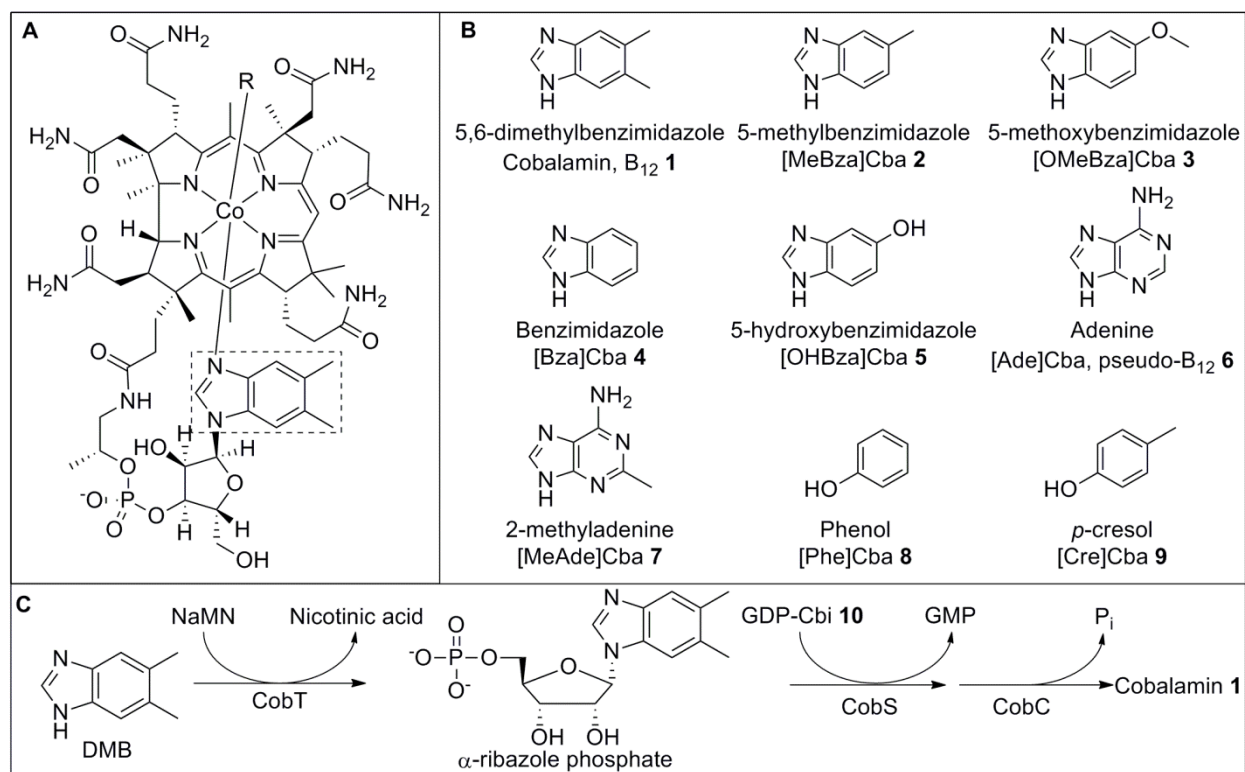
Figure 2.1 Structures of cobamides and lower ligands. (A) Structure of cobalamin **1**. The upper ligand, R, is a methyl or 5'-deoxyadenosyl group in the cofactor forms and a cyano group in the vitamin form. The lower ligand, DMB, is indicated by the box. (B) Structures of cobamide lower ligands. The name of each compound and the abbreviation for the corresponding cobamide are given below each structure. (C) The lower ligand attachment pathway. The pathway for the activation and attachment of DMB to form cobalamin is shown with the names of the *S. enterica* enzymes given. CobT catalyzes the activation of DMB by the attachment of a phospho-ribose moiety derived from nicotinate mononucleotide (NaMN) to form α-ribazole phosphate. The CobS and CobC enzymes catalyze the attachment of the activated base to the cobamide precursor GDP-cobinamide (GDP-Cbi **10**) and de-phosphorylation of the product to form cobalamin, respectively. Abbreviations: DMB, 5,6-dimethylbenzimidazole; MeBza, 5-methylbenzimidazole; OMeBza, 5-methoxybenzimidazole; Bza, benzimidazole; OHBza, 5-hydroxybenzimidazole; Ade, adenine; MeAde, 2-methyladenine; Phe, phenol; Cre, p-Cresol; NaMN, nicotinic acid mononucleotide; GDP, guanosine diphosphate; GMP, guanosine monophosphate; Pi, inorganic phosphate.

Here, we test these hypotheses by examining the ability of five phylogenetically diverse bacteria (Figure 2.2A) to attach a variety of exogenously supplied lower ligand bases. In addition to *S. meliloti*, *L. reuteri*, and *S. enterica*, which are discussed above, we have investigated lower ligand attachment in *Veillonella parvula*, a sequenced relative of *S. ovata* (Gronow et al., 2010), and found that *V. parvula* also produces [Cre]Cba **9**. Using *S. meliloti* as a genetic host, we examine the role of *cobT* homologs in determining the range of lower ligands that can be attached. We also include *Dehalococcoides mccartyi* (*D. mccartyi*) in this study, which we previously found is capable of attaching some of the benzimidazoles investigated here, despite its inability to synthesize a corrinoid or lower ligand base *de novo* (Yi et al., 2012). Our results show that the range of lower ligands that can be attached is controlled by both the availability of lower ligand bases and substrate specificity of CobT.

A

Odoribacter splanchnicus

Rhodomicrobium vannielii

Cellvibrio japonicus

Vibrio cholerae

Colwellia psychrerythraea

Shewanella oneidensis

Kosmotoga olearia

Thermosipho melanesiensis

Desulfotalea psychrophila

Desulfobulbus propionicus

Pelobacter carbinolicus

Hippea maritima

Denitrovibrio acetiphilus

Desulfatibacillum alkenivorans

Acetobacterium woodii

Eubacterium limosum

Thermincola potens

Moorella thermoacetica

Dehalococcoides mccartyi CobT

Heliobacterium modesticaldum

Tepidanaerobacter acetatoxydans

Clostridium acetobutylicum

Clostridium thermocellum

Treponema primitia

Spirochaeta smaragdinae

Veillonella parvula CobT

Clostridium saccharolyticum

Clostridium phytofermentans

Eubacterium rectale

Clostridium cellulovorans

Clostridium perfringens

Syntrophobotulus glycolicus

Desulfurispirillum indicum

Alkaliphilus metalliredigens

Lysinibacillus sphaericus

Escherichia coli

Salmonella enterica CobT

Lactobacillus reuteri CobT

Tolumonas auensis

Veillonella parvula ArsA

Actinoplanes sp.

Geodermatophilus obscurus

Thermobispora bispora

Pseudonocardia dioxanivorans

Rubrobacter xylanophilus

Acidimicrobium ferrooxidans

Mycobacterium tuberculosis

Nocardioides sp.

Nakamurella multipartita

Frateuria aurantia

Allochromatium vinosum

Bradyrhizobium japonicum

Agrobacterium tumefaciens

Sinorhizobium meliloti CobU

Hirschia baltica

Parvibaculum lavamentivorans

Puniceispirillum marinum

Novosphingobium aromaticivorans

Segniliparus rotundus

Hyphomonas neptunium

Veillonella parvula ArsB

Paludibacter propionicigenes

0.2

100
100
69
77
100
98
97
90
96
100
100
96
76
33
47
93
50
59
94
12
53
85
85
66
99 33
93
63
76
11
100
18
39
29
59
100
58
89
79
62
98
99
100
90
59
100
50
46
43
81
26
100

B

```
Dialister invisus              56   -DG-ENKTGGKTSLTEVQMVS
Veillonella parvula ArsB       58   -DGPQNDQHGSESYAAIKRFN
Veillonella parvula CobT       53   MEG-YVGYNYEVTQKQSRNML
Streptococcus sanguinis        62   AEG-VSANPQETTYTVARNIL
Fusobacterium varium           63   EEG-ISSCPLEYTRIVSEAML
Slackia heliotrinireducens     65   AQG-VTQCGQEVSTLVADSVA
Bacteroides fragilis           58   DEG-VSLSPKEITWQQISNFL
Leptotrichia goodfellowii      64   REN-VSKSKRVITQYVVEAML
Clostridium acetobutylicum     62   EEN-VSSCPKSVTASVTRNFM
Salmonella enterica            64   DEG-VAVSPKIVTAIQAANMT
Lactobacillus reuteri          60   HEG-VSASPQKITAIQAVNMM
Sinorhizobium meliloti         69   RQG-VTPFPASVTAQMVENFA
Veillonella parvula ArsA       62   DMG-VSAYPKETTVGMTQNYL
Dehalococcoides mccartyi       63   DEK-FHNWPKEVTAQMLQNFA
Pseudomonas putida             62   EEG-ISAYPQAVTGQMLCNFV
Laribacter hongkongensis       55   CEG-VSAYPAAVTVEMVHNFT
Mycobacterium tuberculosis     60   RSG-VSAYPPEVTAQMVANID
Herpetosiphon aurantiacus      62   EYG-VSAYPSSVTAQMVANFL
Acidimicrobium ferrooxidans    69   DEG-VTPWPRDVTAQMVANFL
Rubrobacter xylanophilus       75   SRG-VSPWPQEVTAEMVRNFC
Desulfotignum phosphitoxidans  64   EEG-VSLFPAEVTPQMVLNFA
Halobacteroides halobius       63   AEG-VSAFPQEVTTQMVYNFL
Caldilinea aerophila           72   AES-VSAYPSEVTAQMVLNFL
Paenibacillus mucilaginosus    56   AEG-VSAFPAEVTPQMVLNFL
Chloroflexus aggregans         63   RQG-VSAFPAEVTPQMVLNFL
Ochrobactrum anthropi          70   AKN-ITPFPPSVTAQMVENFA
Fibrella aestuarina            55   AEG-VSAYPAEVTYQMVQNFI
Aquimarina agarilytica         58   KDNEVNAYPQEVTAQMVYNFL
Niastella koreensis            50   KTGLVNAYPQEVTAQMVLNFA
Leptospira biflexa             61   EEP-VSLYPKDVTWQMVFNFL
Treponema primitia             49   TEG-VSLYPQEVTRQMVLNIL
```

Figure 2.2 Phylogenetic analysis of CobT homologs. (A) Boot-strapped nearest neighbor phylogenetic tree of CobT protein homologs. The tree was assembled as described in Experimental Procedures. Proteins from organisms appearing in this study are indicated in bold. (B) Alignment of CobT proteins from phylogenetically diverse prokaryotes. The region shown corresponds to the region shown in Figure 2.6. Proteins from organisms appearing in this study are indicated in bold.

2.3   Results

2.3.1   Guided biosynthesis in four bacteria shows that a limited set of lower ligand bases can be incorporated into cobamides.

The bacteria *S. meliloti, S. enterica, L. reuteri,* and *V. parvula* were chosen as model organisms to investigate the range of lower ligands that can be attached to form cobamides. The cobamides produced by *S. meliloti*, *S. enterica*, and *L. reuteri* have previously been identified as cobalamin **1**, [Ade]Cba **6** and [MeAde]Cba **7**, and [Ade]Cba **6**, respectively (Campbell et al., 2006; Keck and Renz, 2000; Santos et al., 2007). To confirm that these cobamides are synthesized by our laboratory strains, corrinoids were extracted and analyzed from cultures of each organism. High performance liquid chromatography (HPLC) (Figure 2.3) and liquid chromatography tandem

mass spectrometry with multiple reaction monitoring (LCMS) (data not shown) demonstrated that the expected cobamides were present.
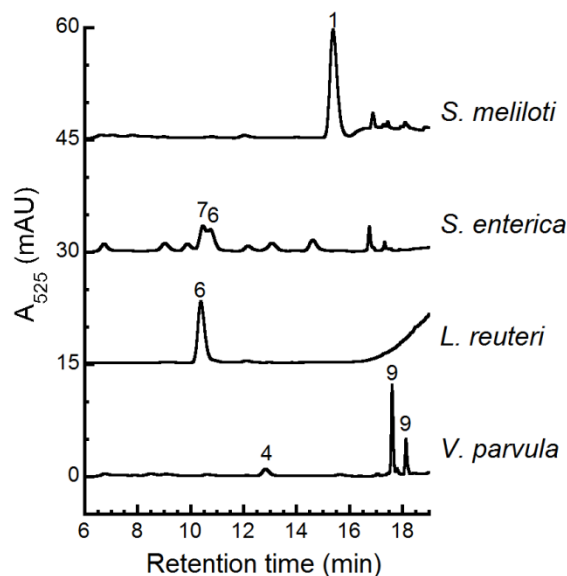

Figure 2.3 Native cobamides produced by bacteria in this study. HPLC analysis of corrinoid extracts from cultures of the indicated bacteria is shown. Numbers correspond to the compounds shown in Figure 2.1. The identity of each numbered peak was determined by LCMS.

S. ovata is known to produce phenolyl cobamides due to the activity of the ArsAB enzyme (Chan and Escalante-Semerena, 2011; Stupperich and Eisinger, 1989; Stupperich et al., 1989). As the genome sequence of *S. ovata* was not available at the time, we tested whether *V. parvula*, a sequenced (Gronow et al., 2010) relative of *S. ovata*, might also produce phenolyl cobamides. HPLC analysis of corrinoid extracts from *V. parvula* show that this bacterium produces a pair of corrinoid species that co-elute with [Cre]Cba **9** extracted from *S. ovata* cultures (Figure 2.3). LCMS analysis confirmed these to be [Cre]Cba **9** by comparison of both retention time and m/z (data not shown). Surprisingly, an additional corrinoid in the extract from *V. parvula* was identified as benzimidazolylcobamide **4** ([Bza]Cba) by LCMS and comparison to a standard prepared by guided biosynthesis in *S. enterica*. As such, *V. parvula* is the only organism known to synthesize both phenolyl and benzimidazolyl cobamides.

Table 2.1 Cobamides produced by guided biosynthesis in five bacteria

| | Lower ligand provided | | | | | | | |
	DMB	MeBza	OMeBza	Bza	OHBza	Ade	Phe	Cre
S. meliloti	+[a]	+	+	+	+	-[b]	-	-
S. enterica	+	+	+	+	+	+	-	-
L. reuteri	-	-	-	-	-	+	-	-
V. parvula	+	+	+	+	+	-	+	+
D. mccartyi[c]	+	+	+	+	n.d.[d]	n.d.	n.d.	n.d.

[a] + indicates corrinoid detectable by HPLC
[b] - indicates corrinoid peak not detectable by HPLC or LCMS
[c] Data reported by Yi *et al.* (2012) based on cultures supplemented with cobinamide and a lower ligand base
[d] n.d., not determined.

We next examined the ability of these organisms to incorporate lower ligand bases provided in the culture medium, a process known as guided biosynthesis. Each of the lower ligand bases shown in Figure 2.1B (excluding 2-methyladenine) was tested for incorporation into cobamides in each bacterium by culturing in media containing the lower ligand base followed by corrinoid extraction and analysis. A representative HPLC analysis of corrinoids extracted from *S. meliloti* is shown in Figure 2.4. An *S. meliloti bluB* mutant was used for the guided biosynthesis experiment because this strain does not produce the lower ligand DMB (Campbell et al., 2006; Taga et al., 2007). HPLC (Figure 2.4) and LCMS analysis (data not shown) demonstrated that *S. meliloti bluB* attached each of the five benzimidazoles to form the corresponding benzimidazolyl cobamides. However, no cobamides were detected when Ade, Phe, or Cre was provided. LCMS analysis revealed that the only corrinoids present in these extracts were mono- and di-cyano forms of GDP-Cbi **10**, a cobamide precursor lacking a lower ligand that we previously detected in cultures of *S. meliloti bluB* (Campbell et al., 2006).
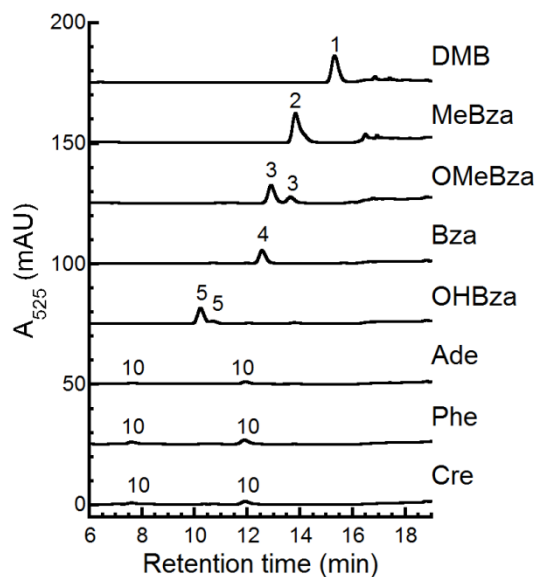
Figure 2.4 Corrinoids produced by guided biosynthesis in S. meliloti bluB. HPLC analysis of corrinoid extracts of an *S. meliloti bluB* mutant cultured in the presence of the indicated lower ligand base is shown. Numbered peaks were verified by LCMS.

The lower ligand incorporation profiles for *S. enterica*, *L. reuteri*, and *V. parvula* were also analyzed. The results of these experiments, as well as previously published results for the bacterium *Dehalococcoides mccartyi* (Yi et al., 2012), are summarized in Table 2.1. *S. enterica* is capable of incorporating adenine in addition to each of the five benzimidazoles, but was not capable of incorporating the phenolic compounds, in agreement with previous findings (Chan and Escalante-Semerena, 2011; Cheong et al., 2001). *L. reuteri* could incorporate only adenine, while *V. parvula* incorporated all of the compounds tested with the exception of adenine (Table 2.1). The failure of *L. reuteri* to incorporate other lower ligand bases does not appear to be due to an inability of the molecules to enter the cell, as we detected an average of 105 ± 2 pmol/O.D.$_{600}$ of free DMB from the cell pellet fraction of three *L. reuteri* cultures grown with 5 μmol DMB. Together, these results indicate that, with the exception of *L. reuteri*, each bacterium can

incorporate multiple lower ligands to form cobamides, though only *S. enterica* and *V. parvula* were capable of incorporating multiple classes of lower ligands.

2.3.2 The S. meliloti cobU mutant can be complemented by heterologous expression of cobT homologs

The results of the guided biosynthesis experiments suggest that the range of lower ligands that can be attached by a bacterium is not controlled solely by the availability of different lower ligand bases in the environment. Next, we examined the influence of *cobT* on lower ligand selectivity, as *cobT* acts first in the lower ligand attachment pathway (Roth et al., 1996). Specifically, we tested whether expression of *cobT* homologs from different bacteria could influence the range of lower ligands that can be attached. *cobT* homologs were expressed in an *S. meliloti* strain with a mutation in *cobU* (the *cobT* homolog in *S. meliloti*). The *cobU* mutant constructed for this study was unable to grow on LB medium without cobalamin **1** supplementation due to the cobalamin **1** requirement of the ribonucleotide reductase enzyme (Campbell et al., 2006; Taga and Walker, 2010). This cobalamin **1** auxotrophy of the *cobU* strain and the *bluB cobU* double mutant strain could be rescued by the addition of cobalamin **1** to the growth medium or by expression of the *S. meliloti cobU* gene on a plasmid. To test whether *cobT* homologs from other bacteria could function in *S. meliloti*, *cobT* homologs from *S. meliloti*, *D. mccartyi*, *S. enterica*, *L. reuteri*, and *V. parvula* were expressed on a plasmid in an *S. meliloti bluB cobU* strain in the presence of DMB. These *S. meliloti* strains are indicated hereafter by the shorthand *Sm cobT$_{Dm}^{+}$*, in which two-letter abbreviations for each organism are used (*S. meliloti*, *Sm*; *D. mccartyi*, *Dm*; *S. enterica*, *Se*; *L. reuteri*, *Lr*; *V. parvula*, *Vp*). All of the *S. meliloti* strains expressing *cobT* homologs from *D. mccartyi*, *S. enterica*, and *L. reuteri* were viable and produced cobalamin **1** when DMB was provided, indicating that these *cobT* homologs were functional. The ability of *Sm cobT$_{Lr}^{+}$* to grow in the presence of DMB is considered later.

Table 2.2 Pairwise identity matrix of CobT homologs

	D. mccartyi CobT	*S. enterica* CobT	*L. reuteri* CobT	*V. parvula* CobT	*V. parvula* ArsA	*V. parvula* ArsB
S. meliloti CobU	35.9	30.0	27.5	25.8	25.2	22.8
D. mccartyi CobT		40.7	29.0	23.7	33.6	26.1
S. enterica CobT			38.0	26.7	30.2	24.2
L. reuteri CobT				29.1	29.3	22.3
V. parvula CobT					25.6	18.4
V. parvula ArsA						20.3

Percent identity was calculated for each individual CobT sequence using the BioEdit sequence alignment editor.

An examination of the *V. parvula* genome revealed the presence of three different *cobT* homologs, Vpar_1456, Vpar_1457, and Vpar_1602, whereas a single *cobT* homolog was present in each of the other bacterial genomes we examined (except *D. mccartyi*, which contains two *cobT* genes with identical sequences) (Gronow et al., 2010; Seshadri et al., 2005). However, only one of the three *cobT* homologs from *V. parvula*, Vpar_1602, complemented the cobalamin **1** auxotrophy of the *S. meliloti bluB cobU* strain. Vpar_1456 and Vpar_1457 are located adjacent to one another, similar to the recently identified *S. ovata cobT* homologs *arsA* and *arsB* (Chan

and Escalante-Semerena, 2011). To test whether Vpar_1456 and Vpar_1457, like *S. ovata arsA* and *arsB*, function when expressed together, both genes were cloned on a single plasmid. Co-expression of these genes rescued the cobalamin **1** auxotrophy of the *bluB cobU* strain in media containing DMB. Based on the function of these genes as described below, and because the predicted protein sequences share 46.7% and 32.6% identity with *S. ovata arsA* and *arsB*, we refer to Vpar_1457 and Vpar_1456 as *arsA* and *arsB*, respectively. The degree of relatedness among the *cobT* homologs examined in this study is shown in the pairwise identity matrix in Table 2.2 and the phylogenetic analysis in Figure 2.2A.

2.3.3   *Expression of cobT homologs alters the lower ligand attachment profile of S. meliloti*

To test whether the expression of *cobT* from other bacteria could alter the range of lower ligand bases that can be incorporated into cobamides, guided biosynthesis assays were performed in the *S. meliloti bluB cobU* strains expressing each *cobT* homolog. A representative HPLC analysis of this assay performed with *Sm cobT*$_{Se}$$^+$ and *Sm arsAB*$_{Vp}$$^+$ is shown in Figure 2.5. *Sm cobT*$_{Se}$$^+$ was found to incorporate the five benzimidazoles as well as adenine, but was unable to attach phenol or cresol (Figure 2.5A), while *Sm arsAB*$_{Vp}$$^+$ could attach all of the bases provided except adenine (Figure 2.5B). These results match the results of the guided biosynthesis study for *S. enterica* and *V. parvula*, respectively (Table 2.1), and demonstrate that these *cobT* homologs are sufficient to specify the range of lower ligands that can be attached.
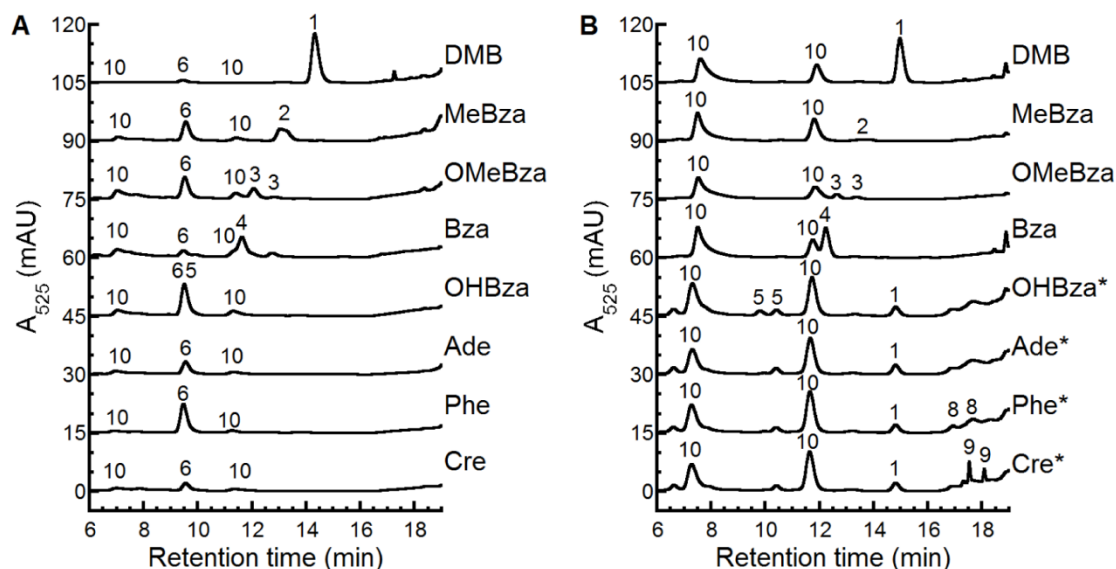
Figure 2.5 Guided biosynthesis in S. meliloti bluB cobU expressing cobT homologs. HPLC analysis of corrinoids extracted from (A) *Sm cobT*$_{Se}$$^+$ and (B) *Sm arsAB*$_{Vp}$$^+$ containing the indicated lower ligand bases. The identities of numbered peaks were verified by LCMS. Asterisks indicate addition of 0.1 μM cobalamin **1** to cultures to support growth.

In analyzing the results above, we observed differences in the amount of incorporation of each substrate, as variable levels of GDP-Cbi **10** were present in the corrinoid extracts (Figure 2.5). We reasoned that these differences likely reflect both the overall degree of complementation of the heterologously expressed *cobT* genes as well as differences in the ability of CobT homologs

31

to activate or attach each lower ligand. The amount of each cobamide produced by guided biosynthesis as a percentage of the total extracted corrinoids was used as an indication of the level of incorporation of each lower ligand (Table 2.3). We found that the lower ligand bases that could be incorporated by each bacterium (Table 2.1) mirrored the lower ligand incorporation profile of *S. meliloti* expressing their respective *cobT* homologs (Table 2.3). An exception to this trend was observed in *L. reuteri*. The inability of *L. reuteri* to incorporate benzimidazoles was not transferred to *Sm cobT$_{Lr}$*$^+$, since the *Sm cobT$_{Lr}$*$^+$ strain was able to attach all five benzimidazoles such that GDP-Cbi **10** made up less than half of the total corrinoids in each culture (Table 2.3). Although adenine was the only lower ligand incorporated by *L. reuteri*, it appears to be incorporated less efficiently than the benzimidazoles in *Sm cobT$_{Lr}$*$^+$ (Table 2.3). Interestingly, both *Sm cobT$_{Se}$*$^+$ and *Sm cobT$_{Lr}$*$^+$ were capable of producing [Ade]Cba **6** without the exogenous addition of adenine, indicating that intracellular adenine pools are available for use as lower ligands in *S. meliloti* as has previously been observed in *S. enterica* (Anderson et al., 2008). This is illustrated in Figure 2.5A where [Ade]Cba **6** was detected in all corrinoid extracts from *Sm cobT$_{Se}$*$^+$.

The two *cobT* homologs of *V. parvula* exhibited distinct lower ligand incorporation profiles when expressed individually in *S. meliloti*. While both *Sm cobT$_{Vp}$*$^+$ and *Sm arsAB$_{Vp}$*$^+$ were capable of incorporating all of the benzimidazole lower ligands, *Sm arsAB$_{Vp}$*$^+$ was the only strain that also incorporated the phenolic substrates *p*-cresol and phenol (Figure 2.5B, Table 2.3). *Sm cobT$_{Vp}$*$^+$ also incorporated adenine, though to a lesser extent than either *Sm cobT$_{Se}$*$^+$ or *Sm cobT$_{Lr}$*$^+$, even though adenine incorporation was not observed in *V. parvula* (Table 2.1). Together, these results indicate that the range of lower ligands that can be incorporated into cobamides in a given organism is heavily influenced by *cobT*.

Table 2.3 Cobamides produced by guided biosynthesis in S. meliloti bluB cobU expressing cobT homologs

	Lower ligand provided[a]							
	DMB	MeBza	OMeBza	Bza	OHBza	Ade	Phe	Cre
Sm cobU$_{Sm}$ $^+$	100	100	2.0	27	4.3	0	0	0
Sm cobT$_{Dm}$ $^+$	64	14	40	33	0.61	0	0	0
Sm cobT$_{Se}$ $^+$	96	38	22	40	n.a.[b]	56	0	0
Sm cobT$_{Lr}$ $^+$	100	76	59	75	n.a.	9.7[c]	0[c]	0[c]
Sm cobT$_{Vp}$ $^+$	63	94	100	100	41	1.0	0	0
Sm arsAB$_{Vp}$	42	4.6	11	30	3.9[c]	0[c]	1.0[c]	11[c]

[a] Numbers represent the level of incorporation of each lower ligand base as a percentage of the total corrinoids present
[b] n.a., not available. Percentage is not available due to co-elution of [OHBza]Cba **5** and [Ade]Cba **6**
[c] Cultures were grown with cobalamin to achieve higher growth, but cobalamin present in the corrinoid extracts was not used in calculations

2.3.4   *Site-directed mutagenesis of S. enterica cobT alters lower ligand attachment specificity*

The results presented above demonstrate that expression of *cobT* homologs from different bacteria in *S. meliloti* allowed it to produce cobamides that wild type *S. meliloti* is incapable of producing. This finding led us to investigate the sequence variations among *cobT* homologs that may contribute to the observed differences in substrate specificity. A multiple sequence

alignment of the CobT homologs revealed that the residues at positions equivalent to Ser80 and Gln88 in *S. enterica* CobT co-vary with the ability to attach adenine (Figure 2.6A). CobT enzymes with Ser and Glu at these positions allow the incorporation of adenine, while those that do not incorporate adenine have Phe/Tyr/Trp and Met, respectively (Figure 2.6A). These residues were previously implicated in the stabilization of adenine in the active site of *S. enterica* CobT by interacting with the amino group (N10) and the N3 ring nitrogen of adenine (Cheong et al., 2001).
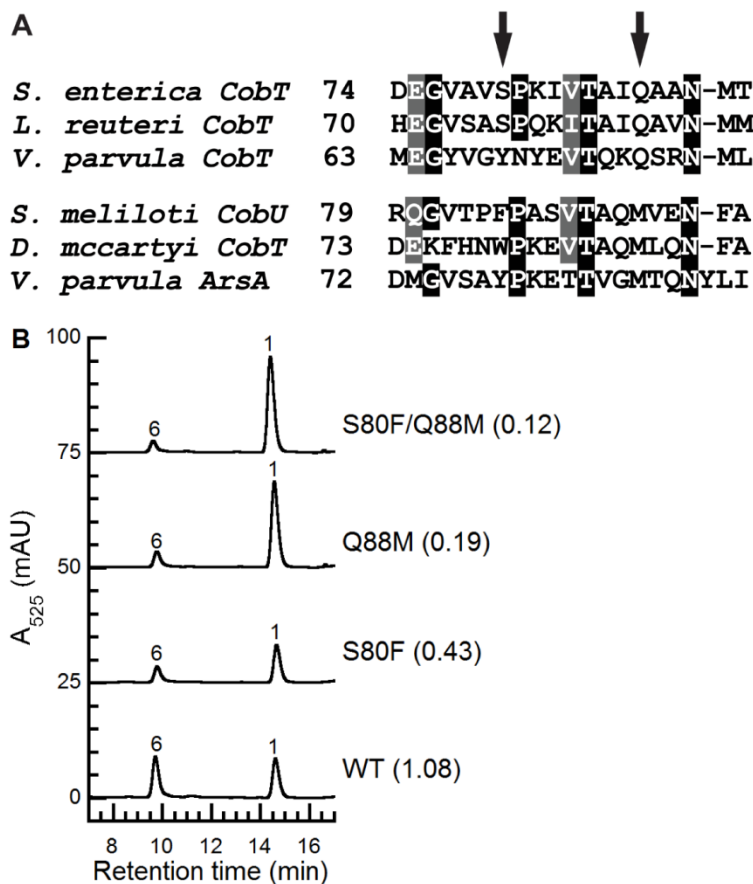
Figure 2.6 Site-directed mutagenesis of S. enterica CobT leads to altered lower ligand incorporation. (A) Multiple sequence alignment of CobT homologs that are capable of activating adenine when expressed in *S. meliloti* (*S. enterica* CobT, *L. reuteri* CobT, and *V. parvula* CobT) and those that do not activate adenine (*S. meliloti* CobU, *D. mccartyi* CobT, and *V. parvula* ArsA). Only the region surrounding the two amino acid residues targeted for site-directed mutagenesis (arrows) is shown. ArsB is not included because it lacks a true active site (Newmister et al., 2012). (B) HPLC analysis of corrinoids extracted from *Sm cobT*$_{Se}$$^{+}$ (WT), *Sm cobT*$_{Se}$ S80F (S80F), *Sm cobT*$_{Se}$ Q88M (Q88M), and Sm *cobT*$_{Se}$ S80F/Q88M (S80F/Q88M) cultures grown in the presence of 0.5 µM DMB. The ratio of [Ade]Cba **6** to cobalamin **1** is shown in parentheses. See also Figure 2.2.

To test the hypothesis that residues Ser80 and Gln88 of *S. enterica* CobT influence the ability to activate adenine, we performed site-directed mutagenesis on *S. enterica cobT* to change these residues to Phe and Met, respectively, which are found at equivalent positions in *S. meliloti* CobU (Figure 2.6A). Plasmids containing *cobT*$_{Se}$ with the S80F, Q88M, or S80F/Q88M mutations were introduced into the *S. meliloti bluB cobU* strain. These strains were cultured in

33

the presence of a limiting amount of DMB to compare the level of attachment of DMB and intracellular adenine. HPLC analysis of corrinoid extracts from these cultures showed that *Sm cobT$_{Se}$*$^+$ produced approximately equal amounts of [Ade]Cba **6** and cobalamin **1** under these conditions (Figure 2.6B). *Sm cobT$_{Se}$* S80F produced 2.7-fold less [Ade]Cba **6** than *Sm cobT$_{Se}$*$^+$ but produced a similar level of cobalamin **1**, after normalizing for culture density (Figure 2.6B). The Q88M mutation impacted both DMB and adenine incorporation, since 2.4-fold less [Ade]Cba **6** and 2.3-fold more cobalamin **1** was produced by *Sm cobT$_{Se}$* Q88M compared to *Sm cobT$_{Se}$*$^+$ (Figure 2.6B). The effects of the S80F and Q88M mutations appeared to be additive, as the *Sm cobT$_{Se}$* S80F/Q88M double mutant produced 3.5-fold less [Ade]Cba **6** and 2.6-fold more cobalamin than *Sm cobT$_{Se}$*$^+$ (Figure 2.6B). These results provide further evidence that differences in *cobT* sequence are responsible for the observed differences in lower ligand incorporation. The active site residues Ser and Gln are not the sole determinants of adenine activation, however, since *S. meliloti* strains expressing *cobU* genes with the reciprocal mutations were unable to incorporate adenine (data not shown).

2.3.5   *Altered lower ligand specificity results in a loss of viability in S. meliloti*

In the course of our experiments, we observed that the growth of the *S. meliloti* strains varied according to the lower ligand base provided, and supplementation with cobalamin **1** was required for growth when certain strains were supplied with adenine or phenolic compounds (Figure 2.5). This result suggested some of the cobamides produced did not support growth in *S. meliloti*. To measure the effect of different cobamides on growth, the O.D.$_{600}$ of *Sm cobT$_{Se}$*$^+$ cultures was measured following growth to stationary phase in minimal media containing each lower ligand base that could be incorporated into a cobamide. We found that DMB, 5-methylbenzimidazole (MeBza), and benzimidazole (Bza) were equivalent in supporting the growth of *S. meliloti*, while cultures supplied with 5-methoxybenzimidazole (OMeBza), 5-hydroxybenzimidazole (OHBza), or adenine reached a lower final culture density (Figure 2.7A). This result indicates that cobalamin **1**, [MeBza]Cba **2**, and [Bza]Cba **4** support the growth of *S. meliloti* to a greater extent than [OMeBza]Cba **3**, [OHBza]Cba **5**, or [Ade]Cba **6**.

To verify that the phenotype observed when lower ligand bases were added to cultures was due to the effect of the cobamides produced rather than differences in the efficiency of incorporation of the lower ligands, an *S. meliloti cobD bluB* strain, which is incapable of producing corrinoids or DMB, was cultured in the presence of a representative purified cobamide from each structural class. Cultures grown with [Ade]Cba **6** or [Cre]Cba **9** led to 3.7- and 13-fold lower final culture densities, respectively, than those grown with cobalamin **1** (Figure 2.7B). These results demonstrate that different cobamides are not functionally equivalent in *S. meliloti*.
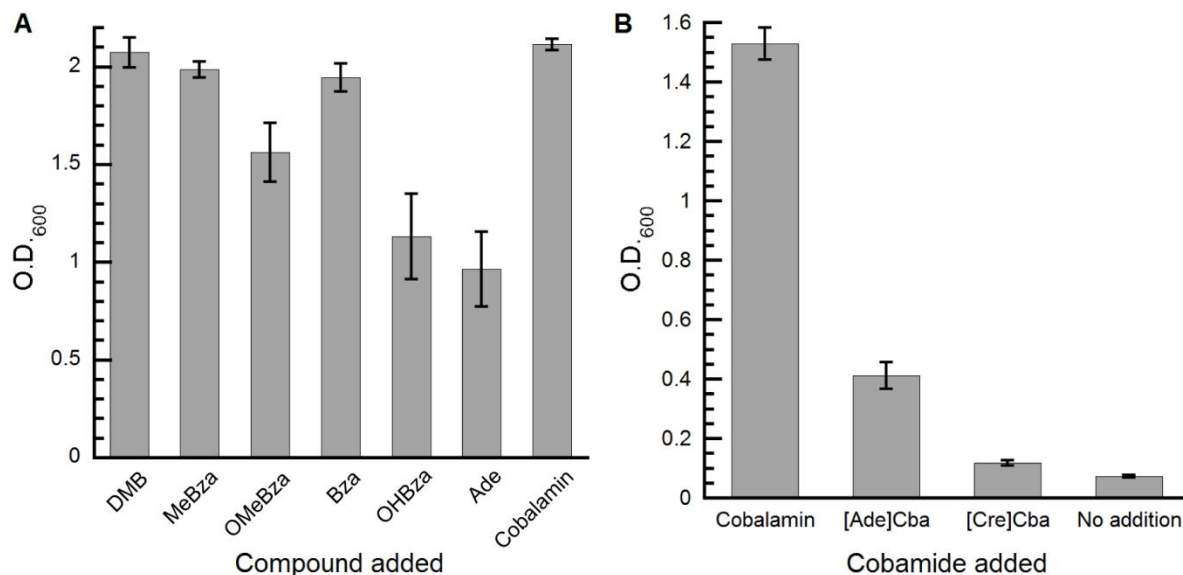
Figure 2.7 Biosynthesis of non-native corrinoids affects growth of S. meliloti. The O.D.600 of *S. meliloti* cultures following growth to stationary phase was measured for (A) *Sm cobT$_{Se}^{+}$* cultures containing 5 µM of the indicated compounds and (B) *S. meliloti cobD bluB* cultures containing 1 µM of the indicated purified cobamides. Error bars represent the standard error of three samples.

2.4   Discussion

It has historically been recognized that different prokaryotes produce a variety of different cobamides, and that the structural differences among these cofactors are almost entirely limited to the lower ligand (Brown et al., 1955; Ford et al., 1955). Although diverse cobamides have been observed in environments such as the human intestine, bovine rumen, and contaminated groundwater, the microbes that produce cobamides have been observed to synthesize only one or two different cobamides when grown in pure culture (Brown et al., 1955; Kräutler et al., 1988; Stupperich et al., 1989, 1990; Keck and Renz, 2000; Allen and Stabler, 2008; Girard et al., 2009; Men, Seth, Yi, Crofts, Allen, Taga and Alvarez-Cohen unpublished). This study addresses the molecular basis of lower ligand selectivity in a representative set of bacteria. We have identified the *cobT* gene as being responsible for limiting the range of lower ligands that can be incorporated. Specificity in CobT activity could be a mechanism of limiting the production of cobamides to those that function as cofactors for a particular organism.

Our genetic system in *S. meliloti* allowed us to examine the influence of substrate specificity by CobT on cobamide synthesis in the absence of factors that may be unique to each organism. We found that the range of lower ligands that could be attached by the *S. meliloti* strains expressing *cobT* homologs mirrored that of the originating bacteria in nearly all cases. The fact that multiple lower ligand bases can be incorporated by guided biosynthesis suggests that the structure of the cobamide produced by each organism is governed in part by the availability of lower ligand bases in the cell. Additionally, our results show that the range of lower ligands that can be incorporated is limited by substrate specificity in *cobT*, and that *cobT* homologs from different organisms have distinct specificity profiles. A notable exception to this pattern was *L. reuteri*, which was incapable of incorporating benzimidazoles by guided biosynthesis, while expression

of *L. reuteri cobT* enabled the incorporation of benzimidazoles in *S. meliloti*. It is possible that in *L. reuteri* another enzyme, such as CobS, prevents the incorporation of benzimidazoles into cobamides. This hypothesis could be addressed experimentally by expressing the *L. reuteri cobS* gene in an *S. meliloti cobS* mutant.

Our results showed that the major differences in substrate specificity among the *cobT* homologs are in the ability to activate adenine and phenolic compounds, while the activation of benzimidazoles is common to all of the *cobT* homologs tested. Differences in the preferences of each CobT enzyme for each lower ligand substrate are investigated in more detail in the accompanying paper by Hazra *et al*. The activation of phenolic compounds appears to be a specialized function of the *arsA* and *arsB* gene products, since none of the other CobT homologs in this or previous studies is capable of activating phenolic compounds (Chan and Escalante-Semerena, 2011; Cheong et al., 2001, 2002). This is likely due to significant differences in the reactivity required for the formation of an O-linked rather than an N-linked glycosidic bond. In addition, consistent with previous studies of *S. ovata* ArsAB, *V. parvula* and *Sm arsAB$_{Vp}^+$* can produce both benzimidazolyl and phenolyl cobamides by guided biosynthesis (Chan and Escalante-Semerena, 2011; Newmister et al., 2012). Given that *V. parvula* naturally produces both [Cre]Cba **9** and [Bza]Cba **4**, and that *Sm arsAB$_{Vp}^+$* is capable of activating benzimidazoles in addition to phenolic compounds, it is puzzling that *V. parvula* possesses an additional CobT homolog that activates only benzimidazoles (and adenine, to a limited extent). A possible explanation is that *V. parvula cobT* and *arsAB* are differentially expressed in order to produce varying ratios of [Cre]Cba **9** and [Bza]Cba **4** under different environmental conditions.

We found that the ability to activate adenine for the production of [Ade]Cba **6** in *S. meliloti* is limited to only three of the *cobT* homologs. In the case of *S. enterica* and *L. reuteri*, this is consistent with their production of [Ade]Cba **6**. Our observation that *Sm cobT$_{Vp}^+$* produces a small amount of [Ade]Cba **6** is unexpected considering the absence of [Ade]Cba **6** in the *V. parvula* corrinoid extracts. This may be explained either by an inability to detect low levels of [Ade]Cba **6** that may be present in the *V. parvula* extract or an insufficient concentration of free adenine in *V. parvula*. The production of a relatively low level of [Ade]Cba **6** by *Sm cobT$_{Vp}^+$* may be due to the presence of Tyr and Gln in *V. parvula* CobT at positions equivalent to residues S80 and Q88 in *S. enterica* CobT. Previously it was proposed that hydrophilic residues are important for stabilizing adenine in the binding pocket and that the hydrophobic residues Y79 and M87 at analogous positions in *S. ovata* ArsA are important for stabilizing phenolic lower ligand bases (Cheong et al., 2001; Newmister et al., 2012). If additional CobT enzymes are found that share this pattern, it may be possible to use this analysis in the bioinformatic prediction of cobamide structures based on *cobT* sequences. Indeed, analysis of these positions in several CobT homologs reveals a co-occurrence of aromatic residues (Phe/Tyr/Trp) with Met, or H-bonding residues (Ser/Thr/Cys) with Gln or Val (Figure 2.2B).

We speculate that the ability to exclude certain lower ligand bases from incorporation into cobamides is a mechanism of preventing an organism from producing a cobamide that it cannot use. Based on this hypothesis, specificity in *cobT* would be particularly important for the exclusion of adenine, which is present intracellularly but is used as a lower ligand only by a subset of bacteria. This idea is supported by our observation that the production of [Ade]Cba **6** in

Sm cobT$_{Se}^{+}$ leads to poor growth (Figure 2.7A). A counterexample to this case is *S. ovata*, which can incorporate benzimidazoles as cobamide lower ligands but is incapable of using benzimidazolyl cobamides for the metabolism of certain carbon sources (Mok and Taga, 2013; Stupperich et al., 1990). The absence of a mechanism of excluding benzimidazoles from incorporation into cobamides in *S. ovata* suggests that *S. ovata* does not produce benzimidazoles or encounter them in the environment.

The results of this study suggest that diversity in cobamide structure is achieved in nature by a combination of the biosynthesis or availability of various lower ligand bases and the molecular specificity in the lower ligand activation gene *cobT*. Because the structure of cobamides has been shown to affect cofactor function in many cases, specificity in *cobT* likely ensures that bacteria produce only cobamides that can be used for their metabolic processes. However, why such diversity in cobamide structure exists in nature remains unclear. The use of lower ligand bases that also function in other cellular process, such as purines, or are products of amino acid catabolism, such as phenolic compounds, is relatively common among bacteria (Anderson et al., 2008; Stupperich and Eisinger, 1989). It is therefore puzzling that some bacteria have dedicated pathways for the biosynthesis of benzimidazoles that, to the best of our knowledge, only function as cobamide lower ligands. Substrate specificity in CobT could be explained by the requirement of certain metabolic processes for specific cobamides (Mok and Taga, 2013; Yan et al., 2012, 2013; Yi et al., 2012). In addition, diversity in cobamide structure may be driven by a need for cobamide-producing organisms to limit the cobamide remodeling or salvaging activity of other organisms. Further biochemical and ecological studies will be necessary to explore these possibilities.

2.5   Significance

Cobamides, which include vitamin B$_{12}$, are essential cofactors for many organisms, both eukaryotes and prokaryotes, in a variety of metabolic pathways but are produced only by a subset of prokaryotes. Cobamides are distinguished by the structure of the lower ligand, and different cobamides are not necessarily functionally equivalent as cofactors. While 16 structurally distinct cobamides have been described, organisms typically synthesize only one or two different cobamides when grown in pure culture. Here we sought to identify the molecular factors that determine which cobamides are produced by bacteria. The final steps in cobamide biosynthesis involve activation of a lower ligand base and attachment to a cobamide precursor. We utilized guided biosynthesis assays in four bacteria and found that a distinct set of lower ligand bases is incorporated by each organism. Heterologous expression of *cobT* homologs from these bacteria and from *Dc. mccartyi* in *Sinorhizobium meliloti* demonstrated that CobT is responsible for preventing the incorporation of certain lower ligand bases. Additionally, the *S. enterica cobT* gene was engineered by site-directed mutagenesis to lower its affinity for adenine relative to DMB, the lower ligand of vitamin B$_{12}$. We also report on the potential for prediction of the cobamide biosynthetic capability of microbial communities based on metagenomic analysis of *cobT* homologs. Our results show that both the availability of lower ligand bases and the substrate specificity of CobT limit the lower ligands that can be attached to form cobamides. Finally, we found that growth is inhibited in *S. meliloti* strains that have been engineered to synthesize alternate cobamides. Our results suggest that substrate specificity in CobT protects

bacteria from producing cobamides that do not support their metabolism. It is important to understand the molecular determinants that underlie cobamide structural diversity due to the integral role played by cobamides in several important microbial communities.

2.6 *In* vitro studies of CobT substrate specificity

In addition to the work presented here I also contributed to the published manuscript by Amrita B. Hazra (Hazra et al., 2013) on which I appear as third author. The results of this paper complement those described above. CobT homologs from each organism in my study, as well as the homolog from *Desulfovibrio vulgaris*, were provided with the lower ligand substrates shown in Figure 2.1B. The resulting α-ribotide-phosphate products were analyzed and characterized by HPLC, LCMS, and ^{1}H-NMR. Each combination of paired lower ligand substrates was also provided to each CobT enzyme in competition reactions to determine relative substrate preference. It was found that, *in vitro*, CobT enzymes display greater substrate promiscuity than *in vivo* and that there exists a widespread preference for DMB as a substrate. This was confirmed to be physiologically relevant by competing benzimidazole lower ligands *in vivo* using *S. meliloti bluB*. I contributed this last experiment, as well as general research design and critical reading of the manuscript.

2.7 Materials and Methods

2.7.1 *Bacterial strains and growth conditions*

S. meliloti strains were grown at 30 °C with aeration on LB agar plates containing 2.5 mM $MgSO_4$ and 2.5 mM $CaCl_2$ (LBMC) or in M9 sucrose medium containing 10 µg/L biotin, 10 µM cobalt chloride and 1 mg/ml methionine (Maniatis et al., 1982). When necessary, media were supplemented with antibiotics at the following concentrations (µg/ml): tetracycline, 10; spectinomycin, 100; gentamicin, 50; and streptomycin, 500. Cobalamin auxotrophs of *S. meliloti* were cultured with 10 µM cyanocobalamin when needed. For extraction of corrinoids from *S. meliloti*, 5 ml of M9 was inoculated with a colony of *S. meliloti* grown on LBMC and incubated for 40-50 h. These cultures were diluted to an O.D.$_{600}$ of 0.02 to 0.04 in 250 ml M9 and incubated for 48-60 h.

For extraction of native corrinoids from *S. enterica* serovar Typhimurium strain LT2, a 500 ml culture was grown anaerobically under an N_2 atmosphere in a minimal medium with ethanolamine as a nitrogen source (Keck and Renz, 2000). For guided biosynthesis experiments, *S. enterica* cultures were grown aerobically in NCE medium supplemented with 1 µM dicyanocobinamide and the indicated lower ligand bases with 1,2-propanediol as the carbon source (Gray and Escalante-Semerena, 2009b). *L. reuteri* DSM 20016 was cultured anaerobically in vitamin B$_{12}$-free Assay medium (Difco) at 37 °C under an atmosphere of 95% N_2 and 5% CO_2 at pH 6.3. *V. parvula* DSM 2008 was cultured anaerobically in a modified *Veillonella* medium at 37 °C under 95% N_2 and 5% CO_2 at pH 7.0 in which 1.5 g/L casamino acids were substituted for individual amino acids (Lopes and Cruz, 1976). For all guided biosynthesis experiments, the media were supplemented with 5 µM lower ligand base unless otherwise noted. O.D.$_{600}$ measurements were recorded using a BioTek Synergy 2 multiwell plate reader. Corrinoids were

purified for feeding experiments from cultures of *S. enterica* or *S. ovata* and quantified as described (Gray and Escalante-Semerena, 2009b; Yi et al., 2012).

To determine the ability of DMB to enter *L. reuteri* cells, three 10 ml cultures of anaerobic MRS medium containing 5 μmol DMB were inoculated with *L. reuteri* to an O.D.$_{600}$ of 0.03. Following 24 h of growth, cells were harvested by centrifugation at 9,000 g, washed in 1 ml of 0.85% saline, and the cell pellet was resuspended in 1ml methanol. The lysate was applied to a Sep-Pak C$_{18}$ cartridge (Waters), and the cartridge was washed with 6 ml of 20% methanol, eluted with 2 ml of 80% methanol. The extract was dried under vacuum and resuspended in 0.25 ml deionized H$_2$O. The DMB in the extract was quantified by HPLC as described below.

2.7.2   Genetic and molecular techniques

To construct the *S. meliloti cobU*::Spc^R strain, a 1.3 kb genomic fragment containing the *S. meliloti cobU* gene and flanking region was amplified by PCR and cloned into pUC19 (Yanisch-Perron et al., 1985) at the *Eco*RI and *Pst*I restriction sites. A Spc^R fragment obtained by BamHI digestion of pHP45Ω (Fellay et al., 1987) was cloned into a BglII site located within the *cobU* open reading frame. The EcoRI-PstI fragment containing *cobU*::Spc^R was transferred to the suicide plasmid pK19 *mob sacB* (Schäfer et al., 1994). The resulting plasmid was introduced onto the *S. meliloti* chromosome by triparental mating (Leigh et al., 1985). Screening for the *cobU* gene replacement and loss of the pK19 *mob sacB* plasmid were performed as described (Schäfer et al., 1994; Taga and Walker, 2010). Double mutant strains were constructed by M12 phage transduction (Finan et al., 1984).

For expression of *cobT* homologs in *S. meliloti*, the *S. enterica trp* promoter sequence was introduced at the BglII and EcoRI restriction sites in the vector pMP220 to produce pMP220+P$_{trp}$. To construct plasmids expressing *cobT* homologs from *S. meliloti*, *S. enterica*, and *D. mccartyi*, genomic regions including the *cobT* open reading frame (ORF) and 20 bp of upstream sequence were amplified by PCR from purified genomic DNA (Chen and Kuo, 1993) and cloned into pMP220+P$_{trp}$ at the XbaI and PstI restriction sites. To avoid potential codon usage incompatibility, the *L. reuteri cobT* ORF was synthesized (GeneArt) to contain codons optimized for expression in *S. meliloti* (Puigbò et al., 2007, 2008) and was fused with the 25 bp sequence directly upstream of the *S. meliloti cobU* ORF (Figure 2.8) and cloned into pMP220+P$_{trp}$. A genomic sequence spanning the *V. parvula arsA* and *arsB* ORFs and 20 bp upstream of *arsA* was amplified by PCR from the *V. parvula* genome and cloned into pMP220+P$_{trp}$. The *S. enterica cobT* S80F and Q88M mutations were constructed by PCR-mediated site-directed mutagenesis (Weiner et al., 1994). All plasmids were introduced into *S. meliloti cobU*::Spc^R *bluB::gus* Gm^R by triparental mating (Leigh et al., 1985).

```
CGAGTTCACATTGCAAGGATGATCCATGCGCGGCAAAATTAAGATGGAATTCGACGAAGTCGAA
CTGCGCCCGCTGAACGAAGAAAAGATGAACGCGATGCAGCAGAAGATTGACGACCTCGCCAAGC
CGGTCAAGGGCCTTGGCCGCCTTGAGGAACTTGCCGACCACCTGGCCGGCATCTACCGCACCAC
CGACTTCAACATCACCCCGCGCAAGTGCCTTGTTTTCGCGGCCGACAACGGCGTCGTCCACGAG
GGCGTCTCGGCTTCGCCGCAGAAGATCACTGCTATCCAGGCTGTTAACATGATGAAGGGCCACA
CTACCGTTGCGGCTCTCGCCAAGGCGTTCAACTGTGACCTCCAGGTCTATGACATCGGCATCAA
GACCGACCTCAACTTCCCGTTGGTCCGTGTTCACAAGATCCGCCACGAAACCGAAGATATGCTC
CACAAGCCGGCCATGTCGATCGCGGAAGCCAGAGAAGCGGTCCAATATGGCTTCGATATGGGAG
TCAAGGCGTTCGAAGAAGGCAACCAGGTTATCGCCACGGGCGAACTCGGCATGGGTAATACGAC
AGCGGCCTCGGCCATCATCGCCGCCCTGCTGAACAAGACCGCTGCGGAAGTAGTGGGCCGCGGC
TCGAACATCAGCGACGAACGTCTCAAGCACAAGATCGATGTCGTCAACCGCTCGCTGGAACGCG
CCAACCTGAAAGAAAACGAGTCGCCGGACCCGCTCATCGTGCTGTCTGAGGTTGGCGCCCTCGA
GCTGGGCGCCATGGCAGGCGCCATGCTGTCGGCGGGCGCGATGAACAAGCCGGTCCTTCTCGAC
GGCTTCCTGTCGTATAGTGCTGCCCTCCTTGCCAATTCGATCAAGCCGGGCGTGGTCAACTACA
TGATCCCGACGCATAAGTCGAAGGAAAAGGGTTCGCGCATCGTCCTCGATGCCCTCGGCCTGGA
CCCGTACATCGACATCAACATGTGCGTTGGCGAAGGGTCGGGCGCCATGATGCTGCTGCCGTGG
CTGGATGGTGTTAAGGCCATGCTCACCAACATGAACACCCTCCAGGAAATGGATTTCAACTTCA
TCCCGTAA
```

Figure 2.8 Sequence of *L. reuteri cobT* after codon optimization for expression in *S. meliloti*. The gene was optimized for expression in *S. meliloti* as described in Experimental Procedures. The 25 bp region directly upstream from the *S. meliloti* cobU ORF is underlined and is followed by the optimized *L. reuteri cobT* ORF.

2.7.3 Phylogenetic analysis of CobT sequences

The CobT representative proteome alignment with 15% co-membership threshold was downloaded from PFAM online resource (Chen et al., 2011) on June 24, 2013. The alignment was uploaded into the BioEdit program (Hall, 1999) for manual removal of gaps and sequences. A bootstrapped nearest neighbor phylogenetic tree based on this alignment was obtained from the CIPRES Science Gateway Resource (Miller et al., 2010) using the RAxML algorithm (Stamatakis, 2006; Stamatakis et al., 2008) and was visualized using MEGA5 (Tamura et al., 2011). Alignments were produced using the MUSCLE program (Edgar, 2004) following retrieval of CobT sequences from the NCBI database on June 18, 2013. The pairwise identity matrix comparing CobT homologs was constructed using the BioEdit program (Hall, 1999).

2.7.4 Corrinoid extraction and analysis

Corrinoids were extracted from cell pellets essentially as described (Yi et al., 2012). Briefly, cell pellets were resuspended in methanol with 0.1% potassium cyanide. Following cell lysis, clarification, and removal of solvent, the extract was applied to a Sep-Pak C_{18} cartridge (Waters) and eluted in 1.8 ml of 75% methanol. The sample was dried at 45 °C under reduced pressure, resuspended in 0.5 ml distilled water, and filtered through a 10,000 MWCO filter (Pall). Samples were stored at -20° C prior to analysis.

An Agilent 1200 series HPLC system equipped with a UV-diode array detector was used to analyze the extracted corrinoids. Samples were injected onto an Agilent SB-Aq 4.5 x 150 mm column with 5 μm pore size at a flow rate of 1 ml/min with mobile phases of A, 0.1% formic acid in water and B, 0.1% formic acid in methanol. The column was maintained at 30 °C. Corrinoids were eluted with a gradient of 25% solvent B for 2 min, 25 to 34% solvent B over 11 min, and 34% to 70% solvent B over 3.5 min. LCMS analysis was performed using the above LC method on an Agilent 6410 liquid chromatograph-triple quadrupole mass spectrometer with multiple reaction monitoring (Yi et al., 2012).

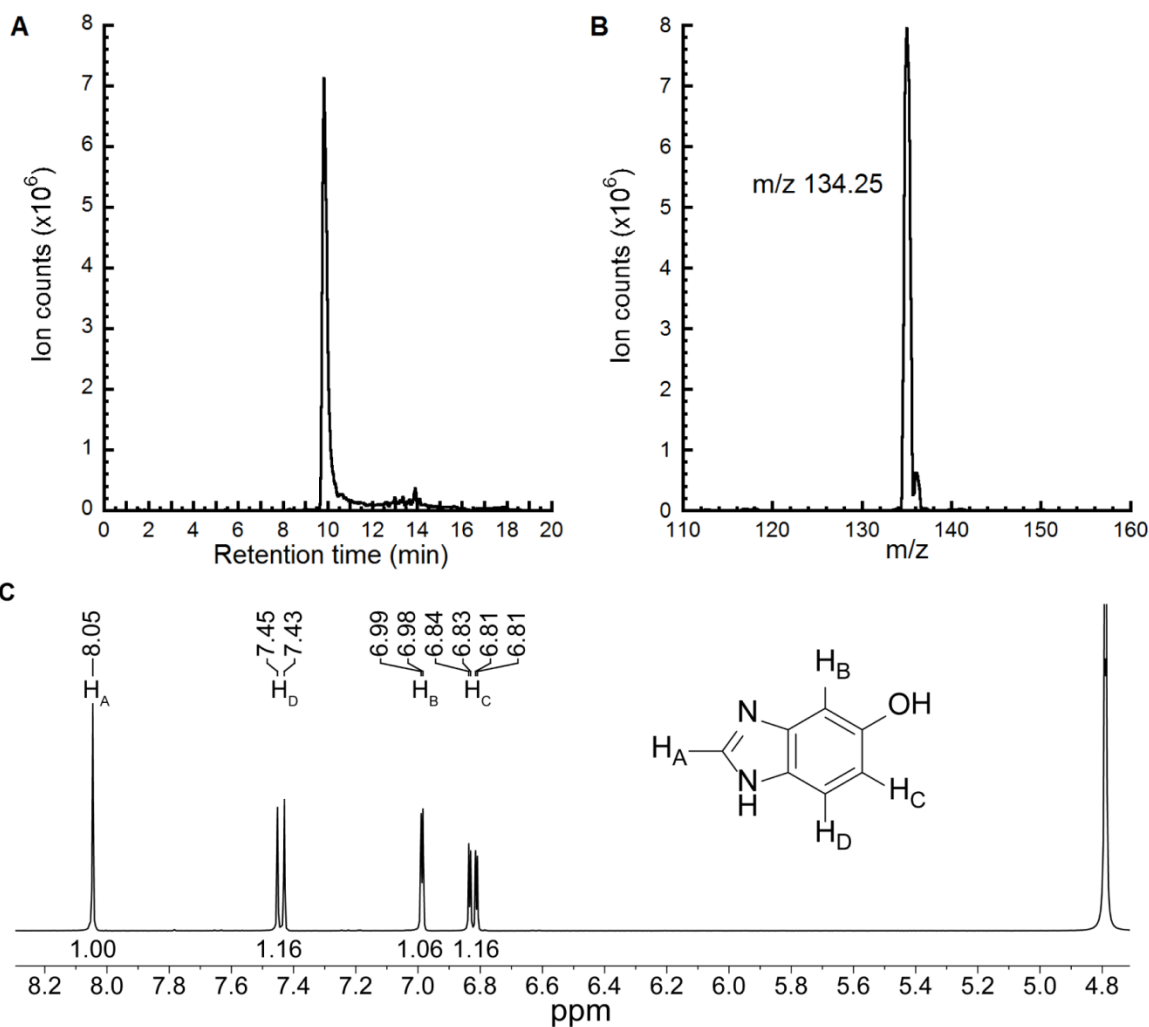### 2.7.5   Synthesis and purification of 5-hydroxybenzimidazole

Figure 2.9 Synthesis and purification of OHBza. (A) LCMS analysis of purified OHBza and (B) mass of peak at 10 min. (C) 1H NMR spectroscopy of the product in D2O. The peak at 4.8 ppm corresponds to water. 1H NMR (400 Mz, D2O) 8.05 (1H, s), 7.44 (1H, d), 6.99 (1H, d), 6.82 (1H, dd).

5-OHBza was synthesized from OMeBza as previously described (Renz et al., 1993) with some modifications and purified as follows. A round bottom flask was charged with 600 mg of OMeBza and 27 ml of 40% hydrobromic acid and heated under reflux at approximately 105 ℃ for 5 h. The product was precipitated by neutralization with sodium hydroxide on ice. The crude product was filtered and dried under vacuum. Final purification was achieved by loading the crude product on a silica column equilibrated with 100% chloroform and eluted with 18:2:0.1 chloroform:methanol:triethylamine. Fractions were collected and analyzed by HPLC. Fractions containing pure OHBza were combined and dried under vacuum. The final product weighed 543 mg and was more than 95% pure as determined by HPLC, with no detectable OMeBza. Analysis of the product by LCMS (Figure 2.9A and 2.9B) and ^{1}H nuclear magnetic resonance spectroscopy (Figure 2.9C) confirmed the structure and purity of the product.

Chapter 3

Cobamide lower ligand orientation can vary and is specified by CobT

3.1   Summary

Cobamides, which include vitamin B_{12} (also known as cobalamin), are a class of cofactors that aid in a variety of biological processes, including fermentation, methanogenesis, amino acid biosynthesis, and bioremediation. Cobamides contain a modified porphyrin ring and are structurally related to other tetrapyrroles such as heme, Factor F_{430} and chlorophyll. Cobamides differ from these other compounds by the presence of a central cobalt atom, the loss of one carbon atom from the tetrapyrrole macrocycle, and the presence of upper and lower axial ligands. While the upper ligand defines the catalytic capabilities of the cofactor, the lower ligand is the site of structural variability and can take the form of a benzimidazole, purine, or phenolic group. The lower ligand is covalently linked to the cobamide through a ribose moiety with either an α-N- or O-glycosidic bond and, in many cases, also forms a coordination bond to the central cobalt atom. The cobalamin biosynthesis enzyme CobT is responsible for the creation of this glycosidic bond and uses an unmodified lower ligand base and a ribose donor as its substrates to produce α-ribosides. Asymmetric benzimidazole and purine bases contain two potential sites on which CobT could act to generate either of two products. Here we demonstrate that CobT enzymes can catalyze the *in vitro* formation of α-riboside structural isomers using 5-methoxybenzimidazole, 5-hydroxybenzimidazole, and adenine as substrates. We provide UV-Vis and ^{1}H-NMR characterization of these molecules and show that different CobT homologs catalyze the production of different quantities of each isomer. Cobamide producing bacteria are also shown to be capable of synthesizing cobamide structural isomers that differ in lower ligand orientation. Cobamide isomers containing 5-methoxybenzimidazole and 5-hydroxybenzimidazole were similarly characterized by spectroscopic means. Finally, we demonstrate that the *cobT* gene provides a genetic basis for cobamide isomer production and that active site mutations in the enzyme can alter the ratio of isomers produced. These findings demonstrate that the CobT enzyme not only provides specificity during substrate attachment but also determines substrate orientation.

3.2   Introduction

As mentioned in previous chapters, CobT and its homologs are the enzymes responsible for activating lower ligand bases for attachment to produce cobamides (Figure 3.1). While performing experiments in Chapter 2 to determine the ability of cobamide-producing bacteria to incorporate exogenous lower ligand bases, I observed that in some cases two cobamide species were produced instead of just one (*e.g.* see Figure 2.4). Other groups have observed that a single cobamide can appear as two species by HPLC due to adoption of differing cyanation states (Anderson et al., 2008; Stupperich et al., 1989). However, I observed that double peaks only occurred when the exogenous lower ligand did not contain an axis of symmetry, for example in the case of 5-methoxybenzimidazle (5-OMeBza **1**, Figure 3.2). In parallel, Amrita Hazra and Jennifer Tran were studying CobT substrate specificity *in vitro* (Hazra et al., 2013). They similarly observed that when provided with an asymmetric substrate the CobT enzyme from

certain organisms catalyzed the formation of products that appear as two species on HPLC (Figure 3.3). We hypothesized that these two observations are evidence for the formation of structural isomers of cobamides and α-ribotides, respectively.
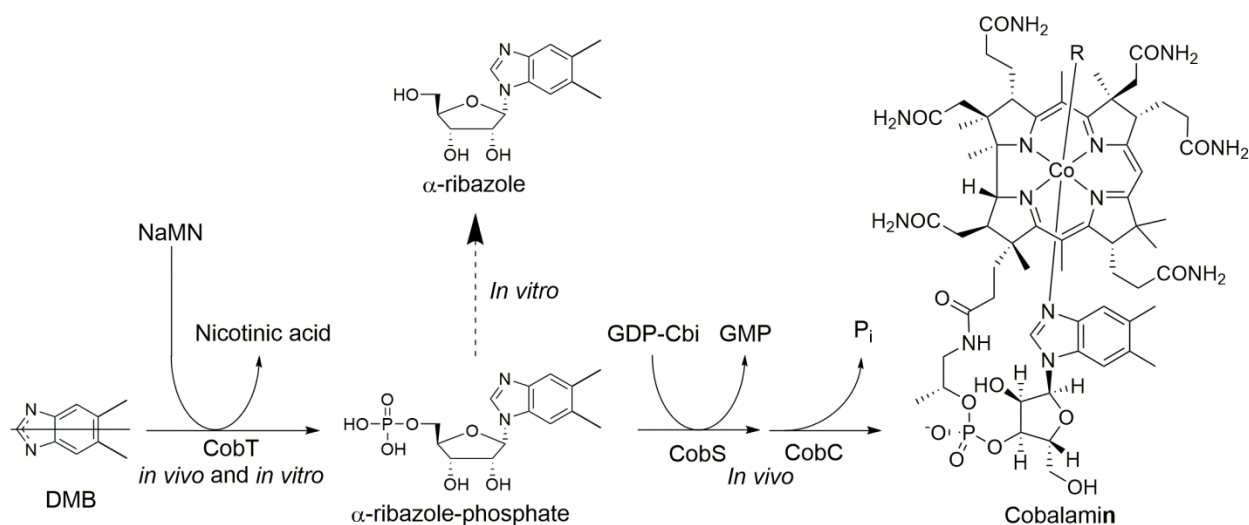

Figure 3.1 Lower ligand attachment pathway and symmetry of DMB. Lower ligand attachment pathway and cobalamin structure. 5,6-dimethylbenzimidazole (DMB, shown in resonance hybrid with a line indicating its axis of symmetry) is phosphoribosylated by CobT using nicotinate mononucleotide (NaMN) as the ribose donor to produce α-ribazole-phosphate. α-ribazole-phosphate is attached by CobS to the cobamide precursor guanosine diphosphate-cobinamide (GDP-Cbi) and is dephosphorylated by CobC to produce cobalamin. In the cofactor forms, the upper ligand "R" group is an adenosyl or methyl group. During purification, the upper ligand is replaced with a cyano group to improve stability. The labile phosphate groups of α-ribotide-phosphate species (e.g. "lower ligand"-RP) were removed enzymatically in vitro to produce α-ribotides (e.g. "lower ligand"-R) prior to further characterization.

Our hypothesis is supported by previous findings. It has been noted that some cobamides contain lower ligand bases that, unlike DMB, do not contain an axis of symmetry (Figure 3.1) (Ford et al., 1955; Friedmann and Fyfe, 1969). It has also been noted that CobT could theoretically activate benzimidazole or purine lower ligand bases by attaching a phosphoribose moiety to either nitrogen of the imidazole ring to produce structural isomers (Cheong et al., 2001; Ford et al., 1955; Friedmann and Fyfe, 1969).

After cobalamin, adeninylcobamide ([Ade]Cba) was the second naturally occurring cobamide to have its structure determined by X ray crystallography and has been extensively characterized and found to be [N7-Ade]Cba **18** (Dion et al., 1952; Hoffmann et al., 2000; Kräutler et al., 2003; Stupperich and Kräutler, 1988). [N7-Ade]Cba **18** and all other purinyl cobamides have only been reported to contain lower ligands covalently bound through the N7 nitrogen. It has been suggested that the N9 isomer is not be possible due to steric constraints (Anonymous, 1955). On the other hand, it has frequently been reported that when asymmetric benzimidazole lower ligand bases are incorporated, either naturally or through guided biosynthesis, two cobamides are produced, with one in much higher abundance. These have been proposed to be structural isomers, but evidence supporting this hypothesis is lacking (Friedrich and Bernhauer, 1958; Kohler, 1988; Kräutler et al., 1987; Perlman, 1959; Stupperich et al., 1993). In at least one case the new isomers were estimated to represent 15-20% of the total cobamide produced (Kohler,

1988). In this chapter, we confirm the hypothesis that isomers of cobamides are produced by guided biosynthesis in bacteria, and that isomers of α-ribotides are produced by CobT homologs *in vitro*. This work demonstrates that the biosynthesis of cobamide isomers is genetically encoded by the *cobT* gene, and that CobT therefore provides specificity both in its choice of substrate and its orientation.
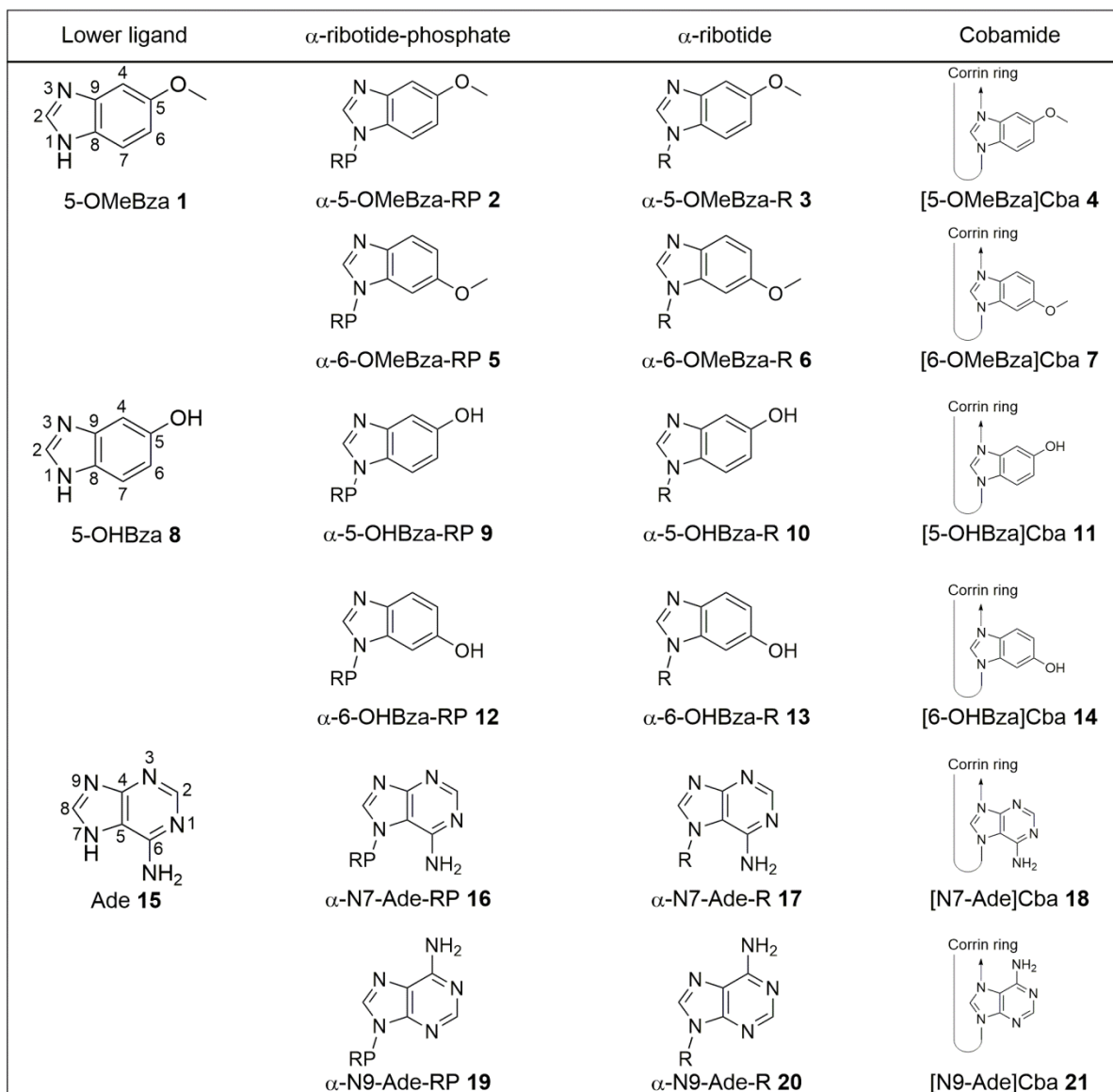
Figure 3.2 Structures of asymmetric lower ligand bases, their CobT products, and corresponding cobamides. Structures of the asymmetric lower ligand bases studied in this chapter as well as the corresponding (phospho)ribosylated CobT products and cobamides are shown. Cobamides are represented schematically to emphasize lower ligand orientation. Abbreviations: methoxybenzimidazole (OMeBza), hydroxybenzimidazole (OHBza), adenine (Ade **15**), ribotide-phosphate (RP), ribotide (R), cobamide (Cba).

3.3 Results

3.3.1 *CobT enzymes react with asymmetric substrates to form two products*

To begin testing the hypothesis that the appearance of multiple HPLC peaks is due to the formation of different α-ribotide-phosphate products, Amrita Hazra and Jennifer Tran analyzed the *in vitro* reactions of asymmetric lower ligands with CobT homologs by HPLC. The lower ligand bases we investigated were 5-OMeBza **1**, 5-OHBza **8**, and Ade **15** (Figure 3.2). The *in vitro* activation of 5-methylbenzimidazole by CobT homologs also resulted in the synthesis of two products, but these products could not be separated well by HPLC, so 5-methylbenzimidazole was not included in our further analyses (data not shown). The CobT homologs that we choose for this study were *V. parvula* and *S. meliloti* for the study of 5-OMeBza **1** and 5-OHBza **8** and *S. enterica* and *L. reuteri* for the study of Ade **15**. The results of *in vitro* experiments using the CobT enzymes from these organisms were of particular interest with regard to the formation of structural isomers (Figure 3.3). The results of these *in vitro* experiments supported our hypothesis. When 5-OMeBza **1** was provided as a substrate, two products were observed, with the relative proportions of the two products varying between CobT homologs (Figure 3.3A). *L. reuteri* CobT produced essentially only the first product, while *S. enterica* CobT and *S. meliloti* CobU (the name for the *S. meliloti* CobT homolog) each produced a small quantity of the second product. *V. parvula* CobT on the other hand produced a majority of the second peak. The inclusion of *S. enterica* in our study provided a useful control, as it has been established that the *S. enterica* CobT enzyme binds asymmetric benzimidazoles in the orientation that would produce 5-substituted α-ribotide-phosphates (Cheong et al., 2001). Based on this we tentatively assigned the first peak as α-5-OMeBza-RP **2** and the second as α-6-OMeBza-RP **5**. The results using 5-OHBza **8** as a substrate, though not as easy to interpret due to poor chromatographic performance, similarly showed two products in each reaction in different proportions (Figure 3.3B). Based on later analyses, we assigned peak 1 as α-5-OHBza-RP **9** and peak 2 as α-6-OHBza-RP **12**.

Lastly, the *in vitro* reactions using Ade **15** as a substrate provided unexpected results. While the literature contains examples of benzimidazolyl cobamide isomers being produced in small amounts, to date there has been no example of a cobamide or CobT product that attaches any purine through the N9 nitrogen (see Figure 3.2 for numbering). We observed that, when provided with Ade **15** as a substrate, only one of two products was formed in each reaction (Figure 3.3C). Based on later analysis, the product associated with *S. enterica* CobT was assigned as α-N7-Ade-RP **16** while the product of *L. reuteri* CobT was assigned as α-N9-Ade-RP **19**.
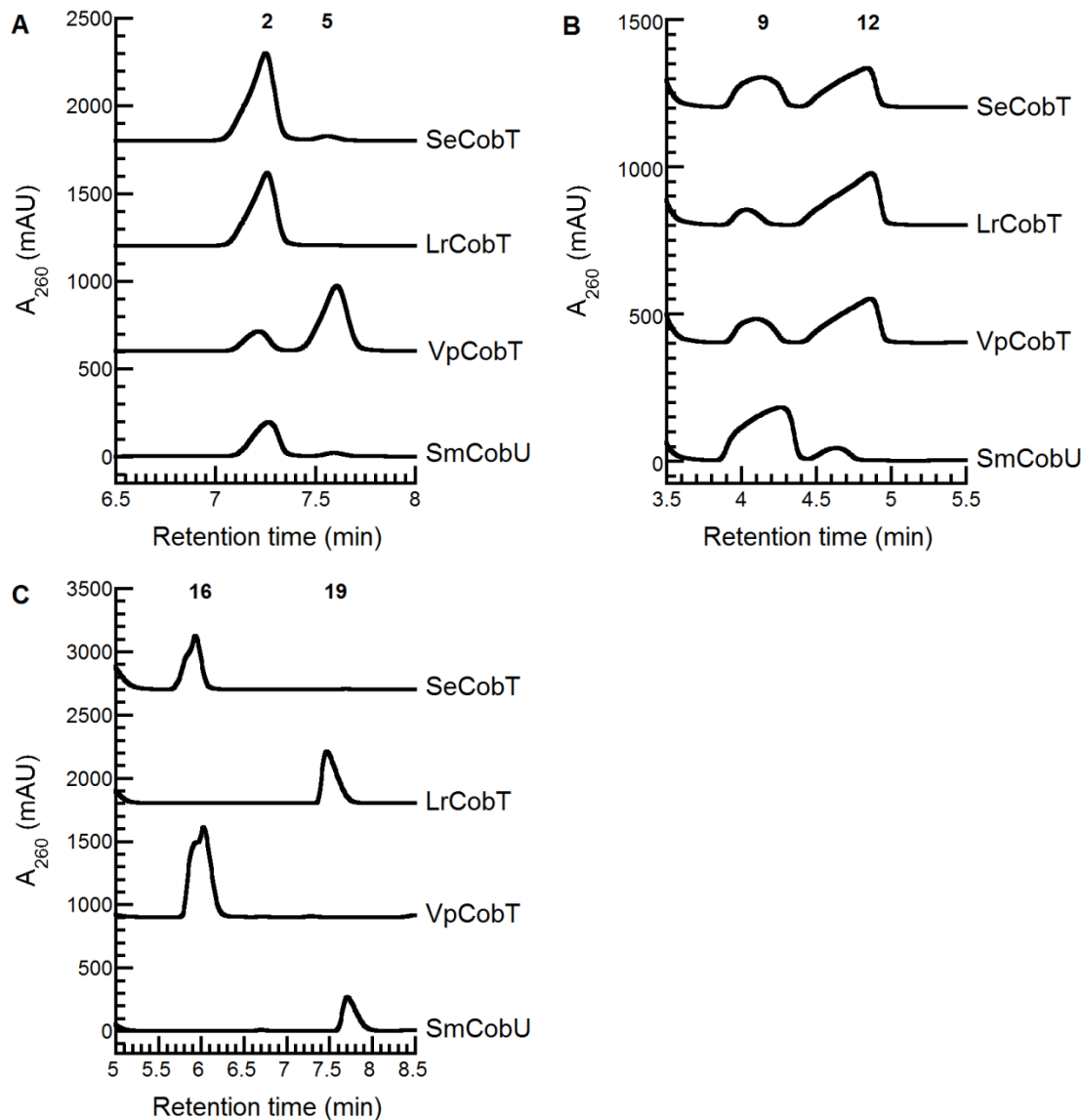
Figure 3.3 CobT homologs can catalyze the formation of structural isomers from asymmetric lower ligand bases *in vitro*. Lower ligand bases were used as substrates in reactions with CobT enzymes from *S. enterica* (SeCobT), *L. reuteri* (LrCobT), *V. parvula* (VpCobT), or *S. meliloti* (SmCobU). HPLC analysis of *in vitro* CobT reactions with (A) 5-OMeBza **1**, (B) 5-OHBza **8**, or (C) Ade **15** as the substrate. For each reaction the α-ribotide-phosphate products are labeled according to Figure 3.1.

3.3.2 Characterization of α-ribotide-phosphate structural isomers produced in vitro by CobT enzyme homologs

In order to confirm the identities of the α-ribotide-phosphate products, a representative CobT reaction was analyzed by LC-MS/MS. Figures 3.4A and 3.4B show the results of the reaction with *S. meliloti* CobU and 5-OMeBza **1**. When monitored at 361.07 m/z (corresponding to α-5(6)-OMeBza-RP), two chromatographically different species are seen (Figure 3.4A). The individual mass spectra of the two species (Figure 3.4B) show that they have the same m/z as

47

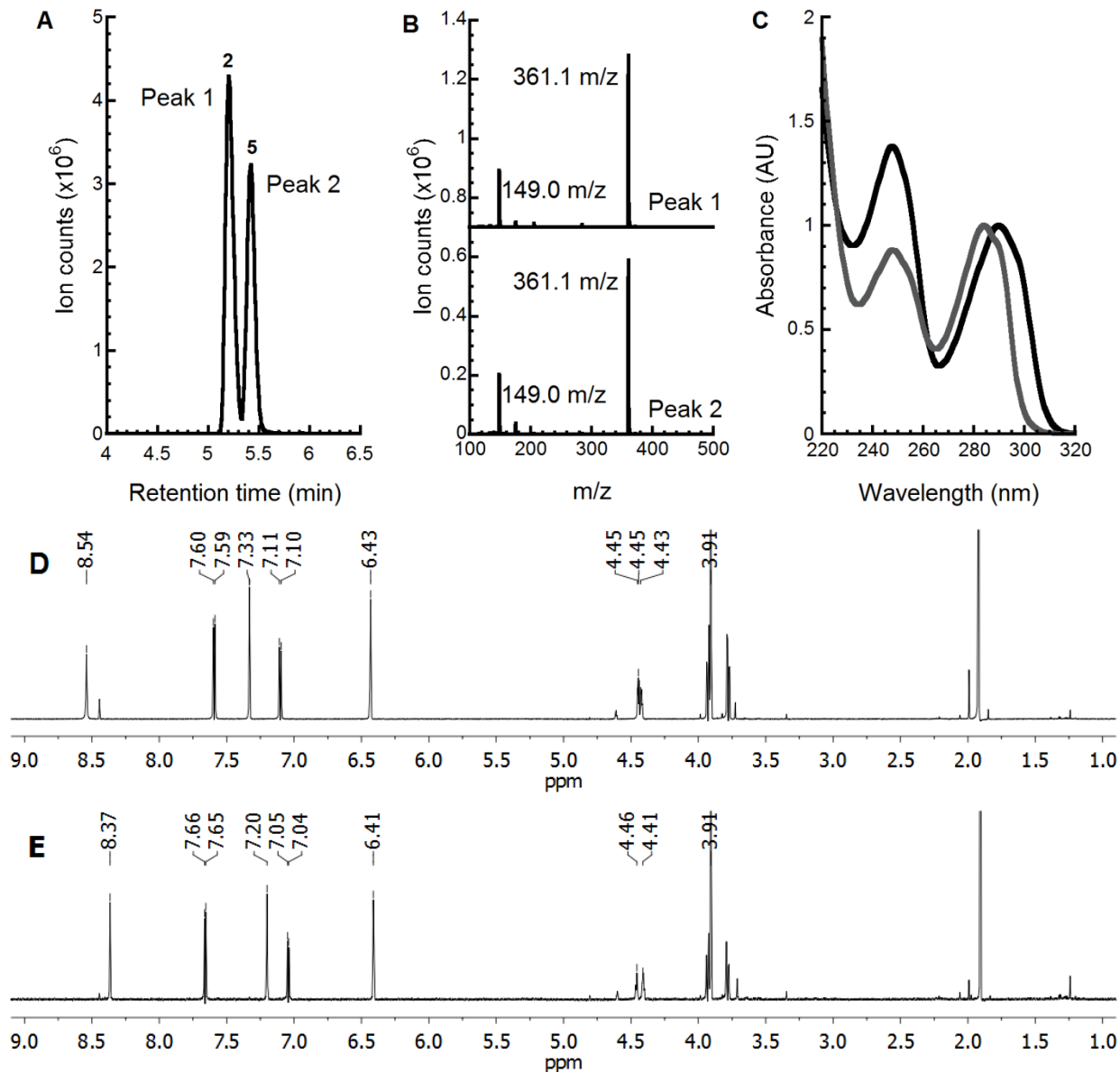**Figure 3.4 Characterization of α-5(6)-OMeBza-RP structural isomers.** A) LC-MS showing representative total ion counts of the α-ribotide-phosphate products purified by HPLC from the *S. meliloti* CobU reaction with 5-OMeBza **1** as the substrate. Two species were produced in this reaction. LC-MS was monitored at 361.07 m/z. B) Masses of peaks 1 and 2 from A), assigned as α-5-OMeBza-RP **2** and α-6-OMeBza-RP **5** respectively, both with an expected mass of 361.07 m/z. C) UV-Vis spectra of α-5-OMeBza-RP **2** (in black) α-6-OMeBza-RP **5** (in gray) normalized to the maximum absorbance near 280 nm. D) ^{1}H-NMR spectra of α-5-OMeBza-R **3** and E) α-6-OMeBza-R **6**. Major differences are found in the low field region from 8.5 to 7 ppm which account for the protons of 5-OMeBza **1** aromatic system (the peak at 3.91 corresponds to the methoxy protons).

each other and the expected m/z of a α-5(6)-OMeBza-RP to within error of the instrument. Also apparent in the mass spectrum are peaks at 149.0 m/z which match closely with the expected m/z of the singly charged 5-OMeBza **1** lower ligand by itself. Similarly, the products formed by *S. meliloti* CobU with 5-OHBza **8** were found to have the same, predicted, m/z of 347.06 within error (Figures 3.5AB). Finally, in reactions containing adenine, *S. enterica* CobT and *L. reuteri*

CobT produced compounds with differing retention times but with the m/z of 348.1, expected for α-N7(9)-Ade-RP (Figures 3.6AB).
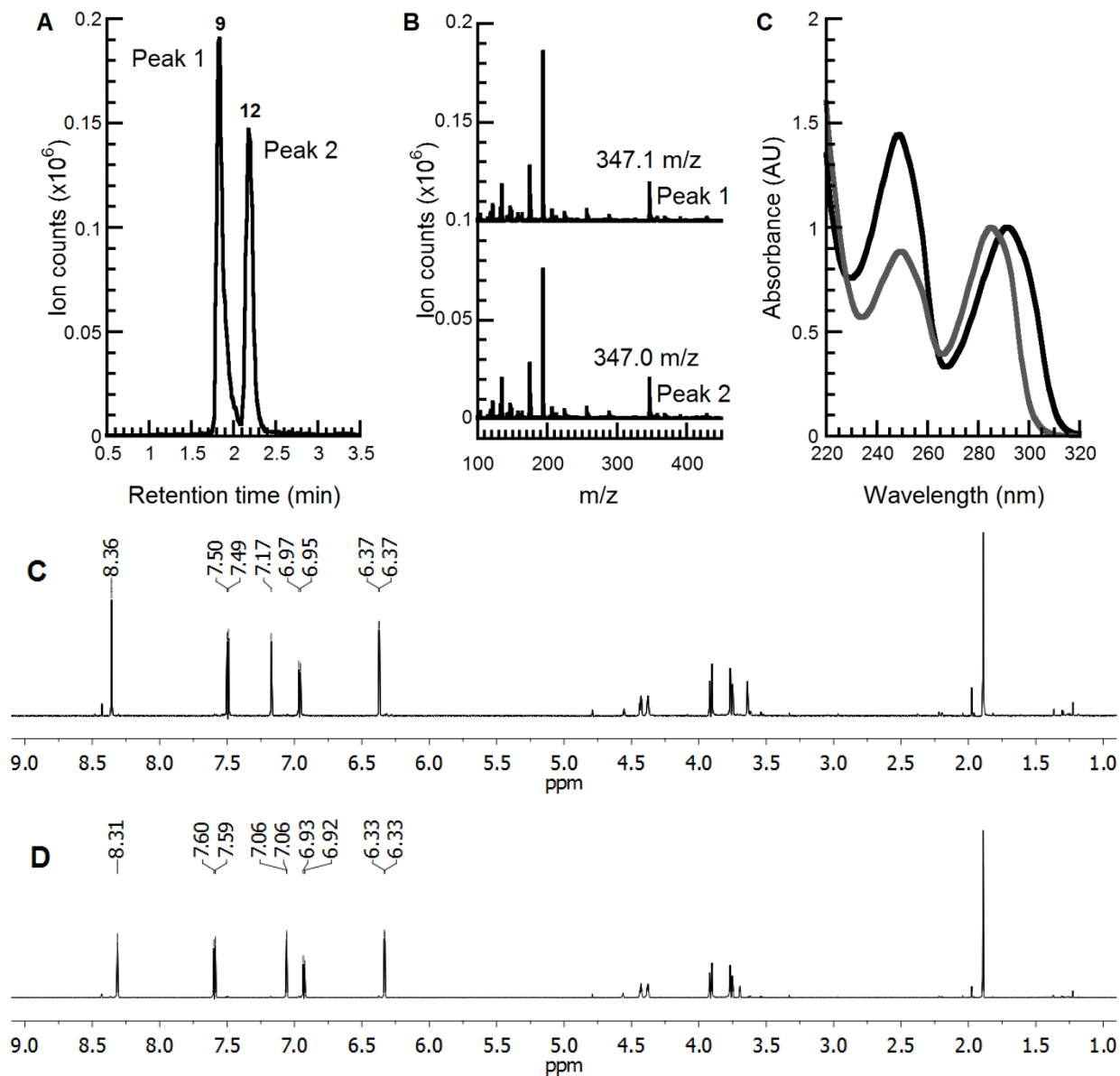
Figure 3.5 *V. parvula* CobT produces two products with 5-OHBza. A) LC-MS showing representative total ion counts of the α-ribotide-phosphate products from the *S. meliloti* CobU reaction with 5-OHBza **8**. Two species were produced in this reaction. LC-MS was monitored at 347.06 m/z. B) Masses of peaks 1 and 2 from A), identified as α-5-OHBza-RP **9** and α-6-OHBza-RP **12** respectively. C) UV-Vis spectra of α-5-OHBza-RP **9** (in black) α-6-OHBza-RP **12** (in gray) normalized to the maximum absorbance near 280 nm. D) 1H-NMR of α-5-OHBza-R **10** and E) of α-6-OHBza-R **13**. Major differences are found in the downfield region from 8.5 to 6.5 ppm which correspond to the aromatic protons of 5-OHBza **8**.

To further investigate our hypothesis we analyzed the putative isomers by UV-Vis and ¹H NMR spectroscopy. By these methods the two compounds clearly differed. For example, the UV-Vis spectra of the putative α-5-OMeBza-RP **2** and α-6-OMeBza-RP **5** compounds differ

substantially. α-5-OMeBza-RP **2** has local maxima at 248 and 290 nm and local minima at 232 and 266 nm. In contrast α-6-OMeBza-RP **5** has maxima at 248 and 284 nm and minima at 234 and 265 nm, and lower absorbance at 250 nm. The UV-Vis spectra of α-5-OHBza-RP **9** versus α-6-OHBza-RP **12** (Figure 3.5C) on the one hand and α-N7-Ade-RP **16** and α-N9-Ade-RP **19** (Figure 3.6C) on the other also differ significantly.

In analyzing the ^{1}H-NMR spectra of each α-ribotide-phosphate product, we found that the ribose phosphate group from the compounds was partially hydrolyzed, leading to extra peaks that interfered with the analysis. We therefore treated the purified α-ribotide-phosphate compounds with alkaline phosphatase to remove the phosphate group, and re-purified the resulting α-ribotides by HPLC. Figures 3.4D and 3.4E show the ^{1}H-NMR spectra of α-5-OMeBza-R **3** and α-6-OMeBza **6** respectively. These spectra differ significantly in the low field end of the spectrum where the protons on the benzimidazole rings are found. Notably, the chemical shifts of the CH_3 protons associated with the methoxy group, found at 3.91 ppm, do not differ. The low field shifts of α-5-OHBza-R **10** and α-6-OHBza **13** (Figure 3.5DE) also differ significantly. Finally, the ^{1}H-NMR spectrum of peak we assigned as α-N7-Ade-R **17** (Figure 3.6D) is consistent with the previously published spectra (Hazra et al., 2013) and differs significantly from the peak designated α-N9-Ade-R **20** (Figure 3.6E) in the low field region as well.

3.3.3 Characterization of cobamide structural isomers produced in bacteria by guided biosynthesis

As with the *in vitro* reactions, we and other groups previously observed that corrinoid extracts from bacteria grown with asymmetric benzimidazole lower ligand bases show two, often partially overlapping, peaks when analyzed by HPLC (*e.g.* Figure 2.4) (Kohler, 1988; Kräutler et al., 1987; Stupperich et al., 1993). *V. parvula* CobT produced a large quantity of both α-5-OMeBza-RP **2** and α-6-OMeBza-RP **5** *in vitro*. I therefore chose to use guided biosynthesis to investigate the production of [5-OMeBza]Cba **4** and [6-OMeBza]Cba **7** in *V. parvula*. Following growth with exogenous 5-OMeBza **1**, the corrinoid extract from *V. parvula* was analyzed by HPLC and compared to similar extracts from *S. enterica* and *S. ovata*, an acetogenic bacterium related to *V. parvula*, grown with 5-OMeBza (Figure 3.7A). The *V. parvula* extract was found to contain two peaks, while the extracts from *S. enterica* and *S. ovata* each contained only one. As the *S. enterica* extract only contained the first species, we tentatively assigned this as [5-OMeBza]Cba **4** and the second species, found only in the *V. parvula* extract, as [6-OMeBza]Cba **7**. Parallel experiments with 5-OHBza **8** were also performed. Interestingly, when provided this lower ligand, *V. parvula* produced a lower level of the second species as measured by absorbance on HPLC (Figure 3.8A). As with [5(6)-OMeBza]Cba, the first species was tentatively designated as [5-OHBza]Cba **11** and the second as [6-OHBza]Cba **14**. Representative LC-MS/MS analyses of [5(6)-OMeBza]Cba (monitored at 679.3 and 149.1 m/z corresponding to the doubly charged cobamide and the 5-OMeBza **1** lower ligand) and [5(6)-OHBza]Cba (monitored at 672.3 and 135.1 m/z, corresponding to the doubly charged cobamide and the 5-OHBza **8** lower ligand) are shown in Figures 3.7B and 3.8B.
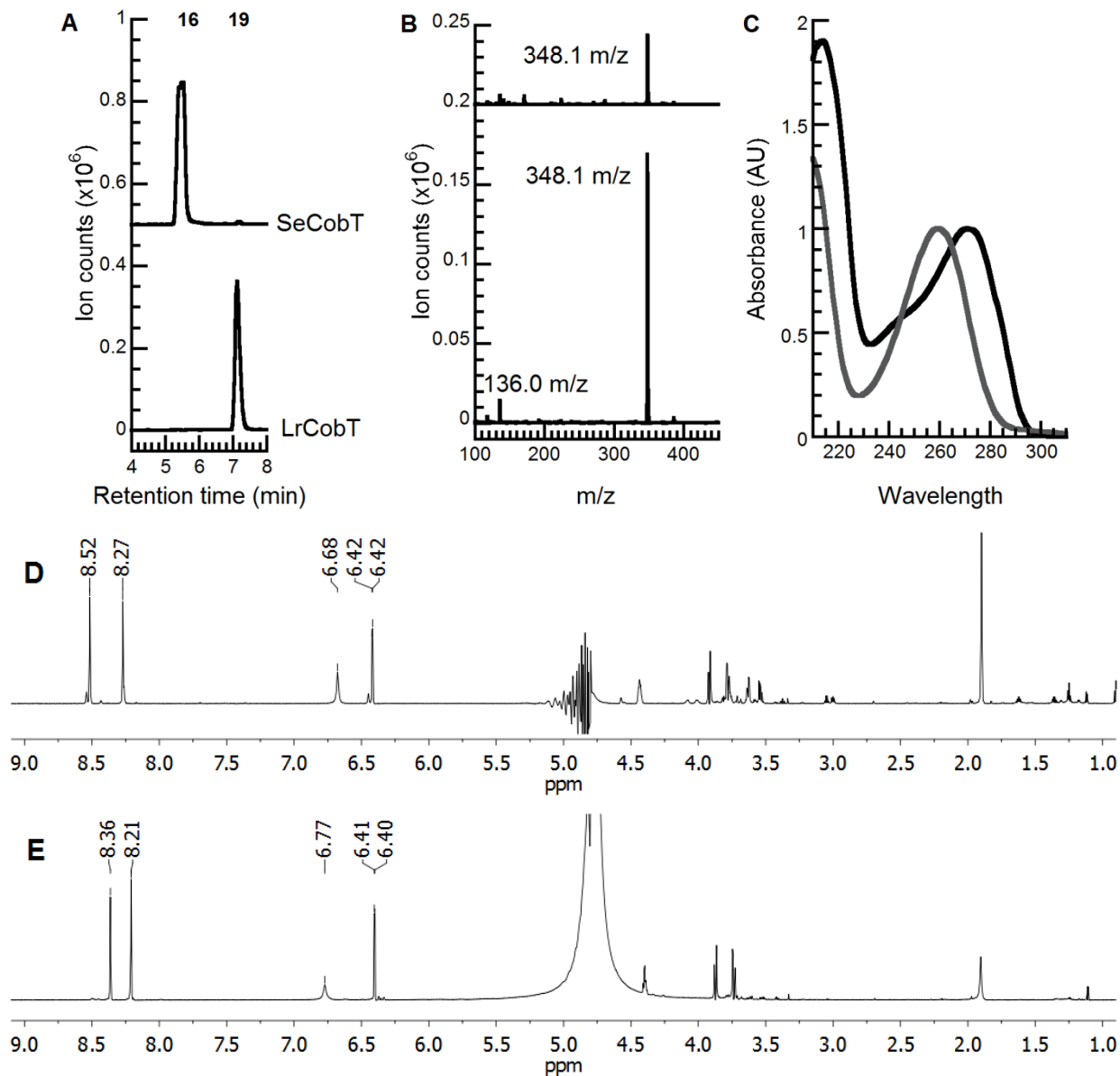
Figure 3.6 CobT reactions with *S. enterica* and *L. reuteri* CobT with Ade lead to different products. A) LC-MS showing total ion counts for α-ribotide-phosphate products of reactions with either *S. enterica* or *L. reuteri* CobT with Ade **15**. LC-MS was monitored at 348.06 m/z. B) Masses of *S. enterica* (top) and *L. reuteri* (bottom) products from A), identified as α-N7-Ade-RP **16** and α-N9-Ade-R **19** respectively. C) UV-Vis spectra of α-N7-Ade-RP **16** (in black) α-N9-Ade-RP **19** (in gray) normalized to the maximum absorbance near 280 nm. D) ^{1}H-NMR of α-N7-Ade-R **17** and E) of α-N9-Ade-R **20**. Major differences are found in the downfield region from 9 to 8 ppm which corresponds to the aromatic protons of Ade **15**.
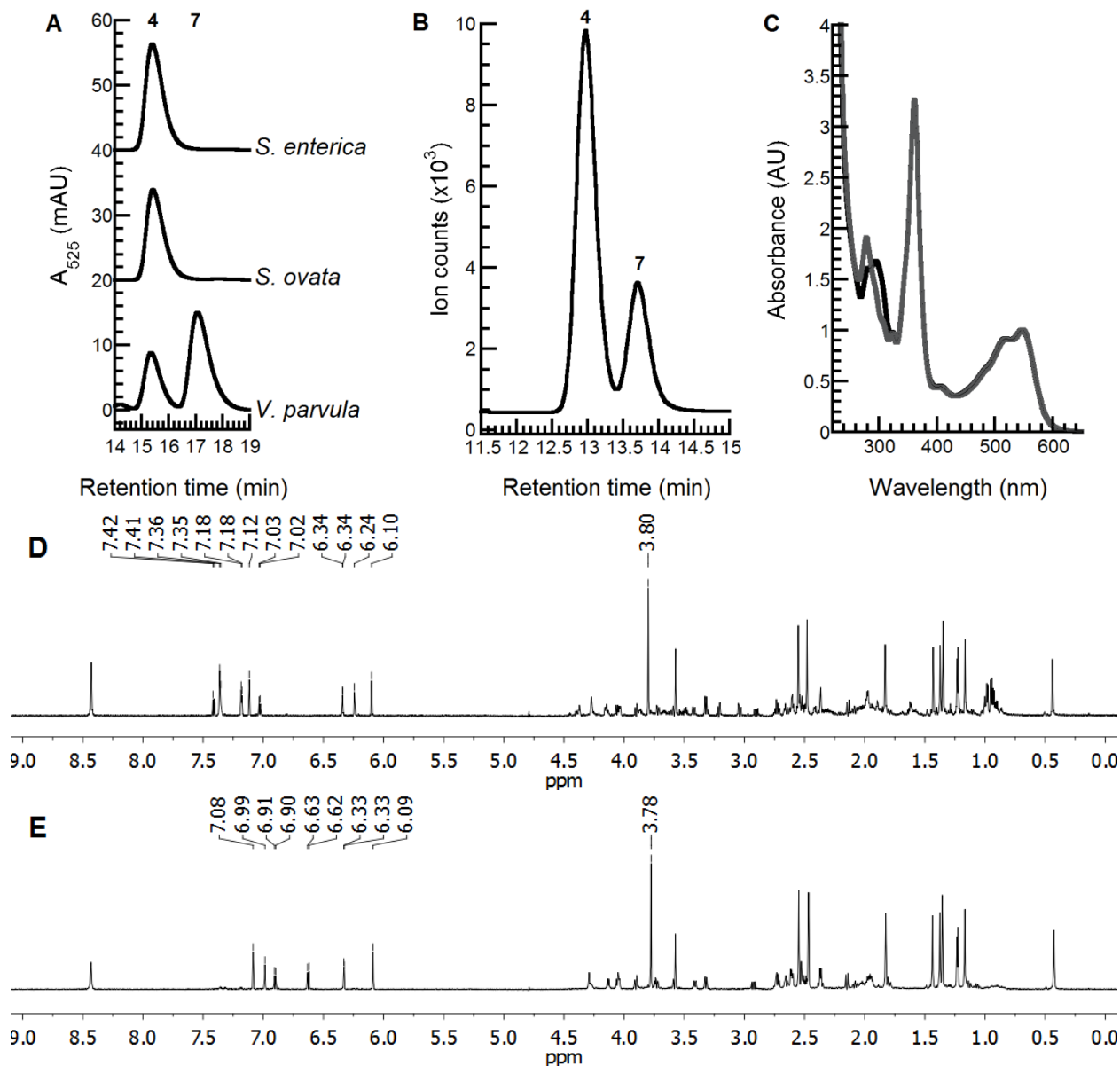
Figure 3.7 Analysis of [5-OMeBza]Cba and [6-OMeBza]Cba produced by guided biosynthesis. A) HPLC of corrinoid extracts of *S. enterica*, *S. ovata*, and *V. parvula* cultures grown with 5-OMeBza 1. B) LC-MS/MS of a representative corrinoid extract from *S. meliloti bluB* grown with 5-OMeBza **1**. Transitions were monitored at 679.3 and 149.1 m/z, corresponding to doubly charged [5(6)-OMeBza]Cba and the 5-OMeBza **1** lower ligand. C) UV-Vis spectra of [5-OMeBza]Cba **4** (in black) and [6-OMeBza]Cba **7** (in gray) normalized to the absorbance at 550 nm. D) 1H-NMR of [5-OMeBza]Cba **4** and E) of [6-OMeBza]Cba **7**. Major differences are found in the low field region from 7.5 to 6 ppm which include the protons of 5-OMeBza **1** (the peaks at 3.80 and 3.78 ppm correspond to the methoxy protons).

Each of these cobamides was purified for further chemical analysis. It has been shown that some bacteria can convert 5-OHBza **8** to 5-OMeBza **1** (Wurm et al., 1975) and our own stock of 5-OHBza **8** was synthesized from 5-OMeBza **1**. To verify that the masses determined by LC-MS/MS were correct (the instrument accuracy is within hundreds of ppm) we analyzed the cobamides by TOF-MS to obtain exact masses. The [5(6)-OMeBza]Cba isomers showed masses

of 679.2808 and 679.2806 m/z, respectively, which are within 2 and 0 ppm of the predicted mass of the doubly charged species. The [5(6)-OHBza]Cba isomers similarly had recorded masses of 672.2728 and 672.2756 m/z which are within 1.3 and 4.2 ppm of the predicted mass of the doubly charged species. With this evidence we were confident about the assigned identities of the cobamides and characterized them further by spectroscopy.
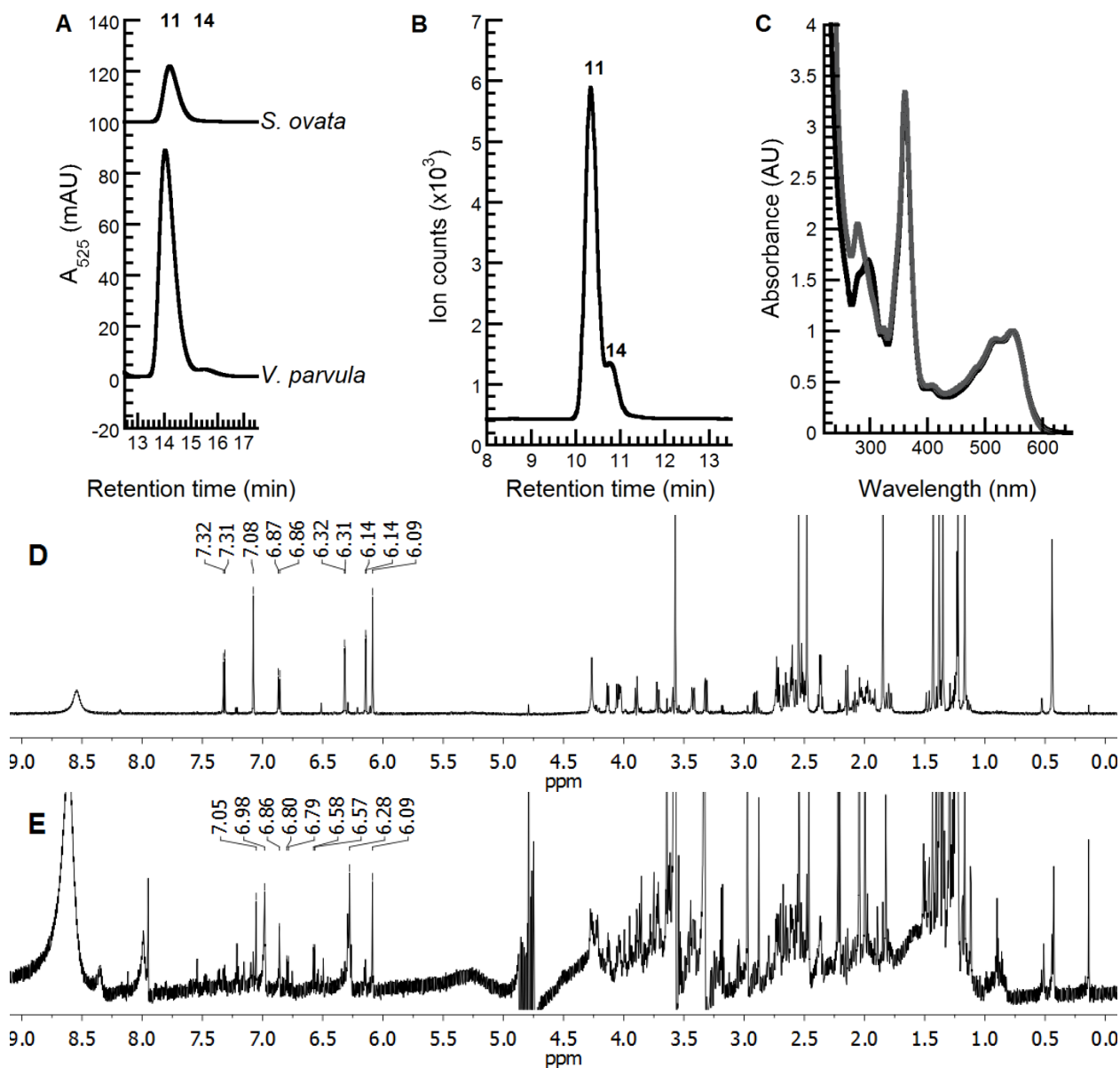
Figure 3.8 *V. parvula* incorporates 5-OHBza to produce two cobamides. A) HPLC of corrinoid extracts of *S. ovata* and *V. parvula* cultures grown with 5-OHBza **8**. B) LC-MS of a representative 5-OHBza **8** corrinoid extract from *S. meliloti bluB*. Transitions were monitored at 672.3 and 135.1 m/z. C) UV-Vis spectra of [5-OHBza]Cba **11** (in black) or [6-OHBza]Cba **14** (in gray) normalized to the absorbance at 550 nm. D) 1H-NMR of [5-OHBza]Cba **11** and E) [6-OHBza]Cba **14**. Major differences are found in the downfield region from 7.5 to 6 ppm which include to the aromatic protons of 5-OHBza **8**.

During HPLC purification, the UV-Vis spectra of putative [5-OMeBza]Cba **4** and [6-OMeBza]Cba **7** were recorded. These spectra, normalized to absorbance at 550 nm, are shown in Figure 3.7C. The spectra are indistinguishable at wavelengths higher than 350 nm, a region dominated by the absorbance properties of the corrin ring. Between 250 and 350 nm the two spectra differ significantly. As expected, these differences mirror the different spectra observed for the corresponding α-ribotide-phosphate compounds (Figure 3.4C). Similar results were observed for [5-OHBza]Cba **11** and [6-OHBza] **14** (Figure 3.8C). Each pair of cobamides was also analyzed by ^{1}H-NMR. When analyzing these spectra, we began with the putative [5-OHBza]Cba **11** because a published ^{1}H-NMR spectrum exists for comparison (Kräutler et al., 1987). Our spectrum (Figure 3.8D) and the published spectrum match very well, confirming our assignment of [5-OHBza]Cba **11** and [6-OHBza]Cba **14** (Figure 3.8E). The low field ^{1}H-NMR spectrum of [5-OHBza]Cba **11** is more similar to the spectrum of the compound we designated as [5-OMeBza]Cba **4** (Figure 3.7D) than the compound designated as [6-OMeBza]Cba **7** (Figure 3.7E). This result, in combination with similarities in the UV-vis spectra of α-5-OMeBza-RP **2** and [5-OMeBza]Cba **4**, indicates that our assignments are correct. As seen with the ^{1}H-NMR spectra of these α-ribotide-phosphates the spectra of the cobamide isomers containing 5-OBeBza **1** differ substantially in the low field region but not in the up field region. Notably, the peaks at 3.80 and 3.78 ppm, corresponding to the methoxy CH_3 group, align closely with eachother (Figure 3.7DE).

3.3.4 In Lactobacillus reuteri corrinoid extracts, only the canonical [N7-Ade]Cba 18 is detected

Several prokaryotes, including *L. reuteri*, have been previously described as synthesizing [Ade]Cba and other purinyl cobamides such as 2-methyladeninylcobamide, 2-methioadeninylcobamide, guaninylcobamide and hypoxanthinylcobamide (Barchielli et al., 1957; Guimarães et al., 1994; Hoffmann et al., 2000; Keck and Renz, 2000; Kräutler et al., 2003; Renz, 1999; Santos et al., 2007; Stadtman, 1960; Stupperich and Kräutler, 1988). However, neither the presence nor the enzymatic production of an α-N9-linked purine has never been reported in the corrinoid literature, and all previously reported CobT enzymes produced α-N7-Ade-R **16** (Cheong et al., 2001; Friedmann and Fyfe, 1969; Fyfe and Friedmann, 1969; Hazra et al., 2013). The production of α-N9-Ade-RP **19** is therefore unique to the *L. reuteri* CobT and the *S. meliloti* CobU (Figure 3.3C), though *cobU* does not appear to have this function *in vivo* (see chapter 2).

In light of this, we set out to determine whether [N9-Ade]Cba **21** did exist in *L. reuteri*. Since *L. reuteri* and *S. enterica* produce [Ade]Cba in the absence of an added lower ligand base, cobamides from cultures of these bacteria were purified by HPLC and subsequently analyzed. The extracts were purified twice using different HPLC methods to remove any other corrinoids, such as 2-methyladeninylcobamide in *S. enterica*, with similar retention times.

Unlike the α-ribotides produced by their respective CobT enzymes, the cobamides from *S. enterica* and *L. reuteri* had identical retention times (Figure 3.9A) and UV-Vis spectra (data not shown). Similarly, the two cobamides had identical masses determined by TOF-MS (Figure 3.10). Finally, analysis by ^{1}H-NMR of the two cobamides found no significant differences in the

proton spectra, and notably the low field spectra were identical (Figure 3.9BC). Based on these results, we conclude that the two cobamides are identical. Explanations for the divergent results of our *in vitro* and *in vivo* experiments are addressed in the discussion.
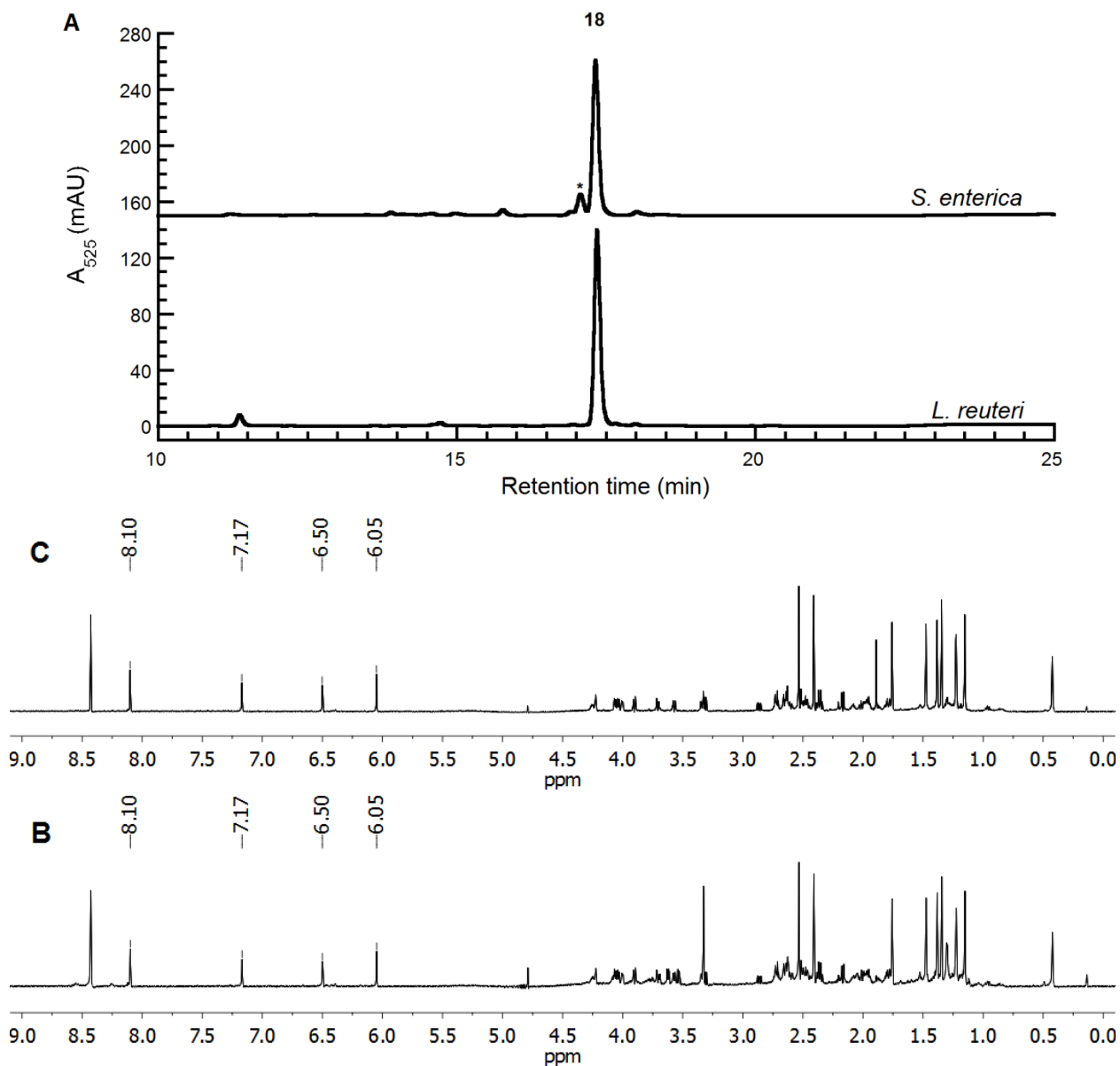
Figure 3.9 *L. reuteri* does not produce [N9-Ade]Cba as its major cobamide in vivo. A) HPLC of purified [Ade]Cba extracts from *S. enterica* and *L. reuteri* cultures. The peak marked as (*) corresponds to [2-methyladenine]Cba. B) 1H-NMR of [N7-Ade]Cba **18** extracted from *S. enterica* and C) 1H-NMR of [N7-Ade]Cba **18** extracted from *L. reuteri*. The downfield peaks are from protons associated with the Ade **15** lower ligand and the nucleotide loop ribose.
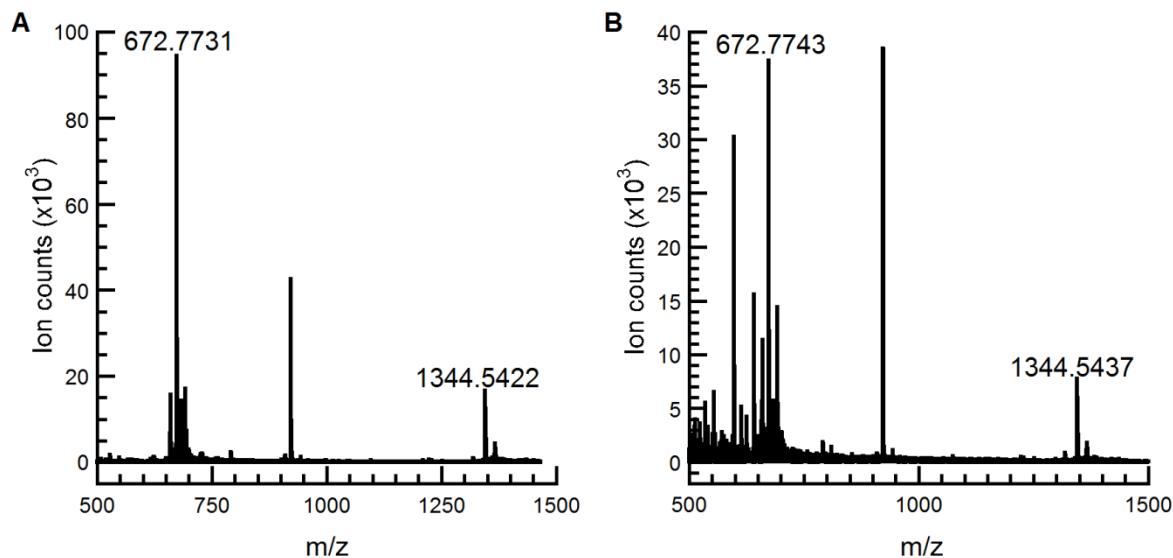
Figure 3.10 *S. enterica* **and** *L. reuteri* **both produce [N7-Ade]Cba as their major cobamides.** A) Exact mass TOF-MS of the purified cobamide from *S. enterica* from Figure 3.9 and B) of the purified cobamide from *L. reuteri* from Figure 3.9

3.3.5 Heterologous expression and mutation of cobT genes demonstrates their role in maintaining substrate orientation

In Chapter 2, I demonstrated that *cobT* homologs differ in their lower ligand substrate specificity. The *in vivo* results presented earlier in this chapter provide evidence that asymmetric lower ligand bases can be used to produce cobamide isomers, and the *in vitro* results point to a role for CobT in determining the ratio of the two isomers. To determine the influence of CobT on cobamide isomer formation *in vivo*, we utilized the *S. meliloti* strains from Chapter 2 to study CobT activity in a heterologous host. The *S. meliloti cobU bluB* double mutant with *cobU* complemented with *cobT* homologs from *S. enterica*, *L. reuteri*, *V. parvula*, and *S. meliloti* (referred to in the text as *Sm cobT$_{Se}$$^+$*, *Sm cobT$_{Lr}$^{opt}*, *Sm cobT$_{Vp}$$^+$*, and *Sm cobU$_{Sm}$$^+$* respectively) were grown in media containing exogenous lower ligand bases. This guided biosynthesis experiment with 5-OMeBza **1** showed that each strain produced different ratios of [5-OMeBza]Cba **4** and [6-OMeBza]Cba **7** (Figure 3.11A). The relative levels of each isomer in *Sm cobT$_{Se}$$^+$*, *Sm cobU$_{Sm}$$^+$*, and *Sm cobT$_{Vp}$$^+$* are similar to those seen in *S. enterica*, *S. meliloti*, and *V. parvula*, respectively, though *V. parvula* and *V. parvula* CobT synthesized both isomers while only one was detected in *Sm cobT$_{Vp}$$^+$* (compare Figures 3.3A, 3.7A, and 3.11A). The results from *Sm cobT$_{Se}$$^+$*, *Sm cobU$_{Sm}$$^+$* (Figure 3.11A) matched well with our *in vitro* experiments (Figure 3.3A), with both making a majority of [5-OMeBza]Cba **4** and a small amount of [6-OMeBza]Cba **7**. I previously reported some of these results in Chapter 2 but the presence of structural isomers was not explored.

In Chapter 2, I described a series of *S. meliloti cobU bluB* strains expressing *S. enterica* CobT containing point mutations in two active site residues. These strains, abbreviated as *Sm cobT$_{Se}$$^+$*, *Sm cobT$_{Se}$^{S80F}*, *Sm cobT$_{Se}$^{Q88M}*, and *Sm cobT$_{Se}$$^{SF, QM}$*, were constructed to investigate CobT specificity with regard to using Ade **15** as a substrate. Here I have examined these strains to

investigate specificity with regard to substrate orientation. Since *S. enterica* produces mostly α-5-OMeBza-RP **2** and *V. parvula* produces mostly α-6-OMeBza-RP **5** *in vitro* (Figure 3.3A) I hypothesized that the *Sm cobT$_{Se}$*^{S80F} mutant would produce the highest levels of [6-OMeBza]Cba **7**. Unexpectedly, it was the Q88M mutation that resulted in the increased production of [6-OMeBza]Cba **7** (Figure 3.11B). The *Sm cobT$_{Se}$*^{S80F} strain and the *Sm cobT$_{Se}$*$^{SF, QM}$ both produced overall more cobamide, in agreement with our finding in chapter 2 that increasing the hydrophobicity of the *S. enterica* CobT binding pocket favored DMB of Ade **15**.

Figure 3.11 Heterologous expression of cobT homologs in *S. meliloti* results in different ratios of [5-OMeBza]Cba and [6-OMeBza]Cba. A) HPLC of corrinoid extracts of guided biosynthesis experiments with *Sm cobT$_{Se}$*$^+$, *Sm cobU$_{Sm}$*$^+$, and *Sm cobT$_{Vp}$*$^+$ grown with 5-OMeBza **1**. B) HPLC of corrinoid extracts of *Sm cobT$_{Se}$*$^+$, *Sm cobT$_{Se}$*^{S80F}, *Sm cobT$_{Se}$*^{Q88M}, and *Sm cobT$_{Se}$*$^{SF, QM}$ grown with 5-OMeBza **1**.

Finally, we explored the possibility that expression of the *L. reuteri cobT* in *S. meliloti* would result in the production of [N9-Ade]Cba **21**. In chapter 2, I demonstrated that, while *S. meliloti* cannot produce [Ade]Cba naturally, it is capable of producing this cobamide when expressing *cobT* genes from either *S. enterica* or *L. reuteri*. Notably however, production of [Ade]Cba was associated with decreased growth and cell viability in *S. meliloti*. The major cobamide of *Sm cobT$_{Lr}$*^{opt} has an identical retention time to [N7-Ade]Cba produced by *S. enterica* and *L. reuteri* (Figures 3.12AB) to the extract from *S. enterica*. ^{1}H NMR analysis confirmed that this cobamide was again [N7-Ade]Cba **25** (Figure 3.12C). Together, these results show that *L. reuteri* CobT is able to produce α-N9-Ade-RP **19** *in vitro*, but in both *L. reuteri* and *S. meliloti* [N7-Ade]Cba is produced.

Figure 3.12 Expression of *L. reuteri cobT* in *S. meliloti* allows for the biosynthesis of [N7-Ade]Cba. A) HPLC chromatogram of the crude corrinoid extracts from *Sm cobT*$_{Lr}$^{opt} and from *S. enterica*. Only the [Ade]Cba peak (boxed) was collected. Peaks marked with (*) indicate species that appear to not be present in the *S. enterica* extract. (+) Indicates exogenous cobalamin that was added as a growth requirement. B) HPLC chromatogram of the [Ade]Cba peak purified from A) using an isocratic method. The shoulder marked with (*) indicates the potential presence of an additional cobamide. The *S. enterica* peak marked (+) corresponds to [2-methyladenine]Cba. The boxed peaks were further purified. C) The [Ade]Cba purified from B) was analyzed by 1H-NMR and found to consist of mostly [N7-Ade]Cba **18**.

3.4   Discussion

While undertaking our previous study of CobT specificity and guided biosynthesis in diverse cobamide-producing bacteria, we observed that sometimes both the products of *in vitro* CobT reactions and the cobamides produced by bacteria contained more than one α-ribotide-phosphate and cobamide respectively. In these cases, just one of the products matched a known compound by retention time. The appearance of a second compound depended on the symmetry of the lower ligand, either as a substrate for the CobT reaction or as an exogenous base for guided biosynthesis. Specifically, if the provided lower ligand contained an axis of symmetry, as in DMB, then only one peak was observed. When the provided lower ligand was not symmetric, two compounds were produced (*e.g.* Figures 3.3 and 3.4). This also depended on the identity of the cobamide producing bacterium, with *S. enterica* producing little of the additional peak, but both *S. meliloti* and *V. parvula* producing significant quantities. This led us to hypothesize that

the two compounds represented structural isomers, and that CobT, in addition to its role in substrate specificity, also regulates substrate orientation. All of the lower ligands for which these observations were made contain an imidazole ring, which could theoretically form N-glycosidic bonds with ribose using either of the two nitrogen atoms (Ford et al., 1955). In fact, the asymmetric benzimidazoles have been seen to do this on a variety of occasions, though in all cases the new isomer represented a minor portion of the total cobamides (Kohler, 1988; Kräutler et al., 1987; Stupperich et al., 1993). Here I have provided evidence that the ability to form two cobamide isomers is characteristic of a variety of cobamide-producing organisms and that the pattern of isomer formation is genetically encoded in the *cobT* gene. With collaborators, I have provided the most detailed characterization of these isomers in the scientific literature. The results are discussed in more detail below.

The results that were most unexpected are those associated with Ade **15**. To date, only cobamides with purines covalently bound through the N7 nitrogen have been described. Our *in vitro* results demonstrate that the CobT enzyme from *L. reuteri* has the biochemical capability of producing α-N9-Ade-RP **19** (Figures 3.3C and 3.6), the necessary substrate for the synthesis of [N9-Ade]Cba **21**. However, the major cobamides produced by *L. reuteri* (Figure 3.9) and *Sm cobT$_{Lr}$*^{opt} (Figure 3.12), are clearly identical to the [N7-Ade]Cba **18** produced by *S. enterica*. There are a few possible reasons the *in vitro* results differ from the *in vivo* results. First is the possibility that our *in vitro* system does not reflect the complexity of the *in vivo* reaction. We have attempted to control for this by repeating the Ade **15** *in vitro* experiments with *S. enterica* and *L. reuteri* CobT using a variety of different ribose donors (including nicotinate mononucleotide and nicotinamide mononucleotide as well as their analogous adenine-containing dinucleotides), and by changing the pH and reaction conditions. None of these manipulations altered the orientation of the α-ribotide-phosphate products.

Second, there is some indirect evidence that the *L. reuteri* CobT may be producing some quantity of α-N9-Ade-R **19** *in vivo*. In order to express heterologous *cobT* genes in *S. meliloti* under a native promoter, the *S. meliloti cobU* gene was swapped with the *cobT* homolog from *S. enterica* and the codon optimized *cobT* from *L. reuteri*. When grown on solid media with cobalamin, DMB, and no supplementation *Sm bluB cobU::cobT$_{Se}$*$^{+}$ and *Sm bluB* both grow without a noticeable defect (Figure 3.13ABC). However, when *Sm bluB cobU::cobT$_{Lr}$*^{opt} was cultured without supplementation the strain failed to grow (Figure 3.13C). This is not due to failure to express the *cobT* gene as supplementation with DMB rescues growth (Figure 3.13B). I hypothesized that this was evidence for the production of [N9-Ade]Cba **21**.

We sought to obtain more direct evidence for the presence of [N9-Ade]Cba **21** by extracting corrinoids from a strain of *S. meliloti* expressing the *L. reuteri cobT*. We were unable to collect a concentrated sample of the cobamide and as a result our ^{1}H NMR analysis was performed on a dilute sample in which the major component appears to be [N7-Ade]Cba **18** (Figure 3.12C). It is possible that, while the majority of cobamide present is [N7-Ade]Cba **18**, an additional quantity of [N9-Ade]Cba **21** is present below the limit of detection. While purifying the cobamide extract on HPLC I observed that the purified [N7-Ade]Cba **18** peak had a small shoulder, possibly indicating the presence of an additional species (Figure 3.12B). In this case, the minority cobamide may be acting as a competitive inhibitor of the cobamide-dependent enzymes of *S.*

meliloti, explaining the growth defect, while remaining below the detection limit of ^{1}H-NMR. It is also possible that [N9-Ade]Cba **21** is present in larger quantities but that it was not collected during purification. When analyzing corrinoid extracts from bacterial cultures it is possible to overlook base-off corrinoids as being cobamide precursors such as GDP-Cbi. It could be that both *L. reuteri* and *Sm cobT$_{Lr}$*[opt] can produce [N9-Ade]Cba **21**, but that this cobamide takes a base-off conformation in solution and so was not collected. Some of the peaks that were not collected are labelled in Figure 3.12A. We are planning on analyzing some of these base-off corrinoids (Figure 3.12AB) by LC-MS/MS to determine whether they contain [N9-Ade]Cba **21**.
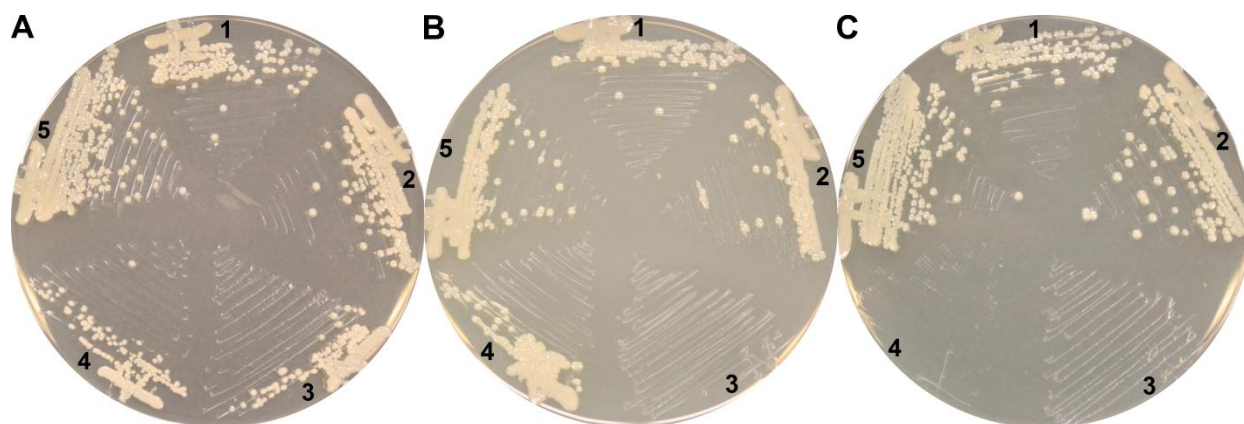

Figure 3.13 S. meliloti shows a growth defect when grown using [Ade]Cba produced using L. reuteri cobT. *S. meliloti* strains were cultured in a NrdJ-dependent condition on agar plates containing (A) 10 μM cobalamin, (B) 10 μM DMB, and (C) no addition. Each plate was streaked with five strains: (1) S. meliloti WT, (2) S. meliloti bluB, (3) *S. meliloti bluB cobU*, (4) *S. meliloti bluB cobU::cobT$_{Lr}$^{opt}*, and (5) *S. meliloti bluB cobU::cobT$_{Se}$$^{+}$*. The DMB growth requirement for the *L. reuteri* expressing strain but not the *S. enterica* or *S. meliloti* expressing strains may indicate the presence of [N9-Ade]Cba **21**.

A final possible explanation for our results is the involvement of downstream enzymes or other factors. The product of CobT is attached to GDP-Cbi by the CobS enzyme (Figure 3.1). It is possible that the *L. reuteri* CobT produces α-N9-Ade-RP **19** in addition to α-N7-Ade-RP **16** but that CobS enzyme can only use α-N7-Ade-RP **16** as a substrate. We proposed a similar hypothesis in chapter 2 to explain why *L. reuteri* does not produce cobalamin when provided exogenous DMB despite the ability of its CobT enzyme to activate DMB. The substrate specificity of CobS has not been thoroughly investigated yet and it is an open question as to whether it has any control at all over what lower ligand bases can be incorporated by cobamide producing organisms.

The results of our guided biosynthesis experiments in *V. parvula* and on its CobT enzyme are the first to demonstrate an organism that produces as a majority of its cobamides a base in the 6-substituted orientation, potentially providing new insights into how CobT binds its substrates. The relevance of this to *V. parvula* is, however, tempered by the fact that its native lower ligand bases, *p*-cresol, phenol, and benzimidazole, all contain an axis of symmetry (see Chapter 2). It is notable that *V. parvula* produces two CobT enzymes, one the canonical CobT studied here, and the second a member of the ArsAB family of CobT homologs first described in the phenolyl cobamide producing bacterium *S. ovata* (Chan and Escalante-Semerena, 2011; Crofts et al., 2013; Gronow et al., 2010; Hazra et al., 2013; Newmister et al., 2012; Stupperich et al., 1988).

Current evidence suggests that the presence of the *arsAB* genes is a requirement for the ability of an organism to activate phenolic lower ligand bases. It is not known why the *V. parvula* genome encodes both ArsAB and CobT. We have demonstrated that the *V. parvula* ArsAB enzyme can activate benzimidazoles (Hazra et al., 2013) and I demonstrated in Chapter 2 that expression of the *arsAB* genes allows *S. meliloti* to produce benzimidazolyl cobamides as well. Therefore the sole expression of the *arsAB* genes would seem to be sufficient for the production of [Bza]Cba and other benzimidazolyl cobamides by guided biosynthesis. Our new data provide possible indirect evidence that both CobT and ArsAB are active in *V. parvula* under our growth conditions. Heterologous expression of *V. parvula arsAB* in *S. meliloti* in the presence of 5-OMeBza **1** led to the production of mostly 5-substituted cobamides (see Chapter 2). Expression of *V. parvula cobT*, however, leads to the production of essentially only 6-substituted cobamides (Figure 3.11A). The simplest explanation for the ability of *V. parvula* to produce both isomers of asymmetric benzimidazoles by guided biosynthesis is that both *arsAB* and *cobT* are expressed under our growth conditions, and both contribute to the activation of 5-OMeBza **1**.

In *S. ovata*, the only other bacterial species documented as producing phenolyl cobamides and a relative of *V. parvula*, cobamides are required for acetogenesis. *V. parvula*, however, is not an acetogen but rather derives its energy from fermentation. The *V. parvula* genome is predicted to contain only two types of cobamide-dependent enzymes: methionine synthase and methylmalonyl-CoA mutase (for cobamide-dependent enzymes see Chapter 1) (Gronow et al., 2010). Methionine synthase catalyzes the final step in the biosynthesis of methionine and utilizes a methylated cobamide in a one-carbon transfer reaction. Methylmalonyl-CoA mutase uses an adenosylated cobamide for the interconversion of succinyl-CoA and methylmalonyl-CoA (Gruber et al., 2011). Bacteria of the genus *Veillonella* have been cited as producing high levels of cobamides (Friedmann and Cagen, 1970) and, in our laboratory, *V. parvula* produces greater amounts of cobamides than *S. enterica* or *S. meliloti* (data not shown). In *V. parvula,* methylmalonyl-CoA mutase is required as part of the methylmalonyl-CoA pathway for the fermentation of lactate to propionate, acetate, and CO_2. As the route for the fermentation of these substrates is energetically inefficient, it is possible that the high cobamide content of *V. parvula* reflects the need for a large cobamide-mediated metabolic flux (Denger and Schink, 1992; Ng and Hamilton, 1971; Seeliger et al., 2002). Both methionine synthase and methylmalonyl-CoA mutase bind cobamides in a "base-off/his-on" conformation in which a histidine residue of the enzyme displaces the lower ligand in coordinating with the central cobalt atom (Gruber et al., 2011). The major cobamides produced naturally by *V. parvula* contain phenolic lower ligands which are incapable of coordination which is consistent with the requirements of these two enzymes.

It has been observed that some cobamide-producing organisms synthesize a small but significant percentage of cobamides that are structural isomers of their dominant cobamide. It is not known whether these isomers are functionally equivalent despite comprising up to 20% of the total cobamide in natural producers of asymmetric lower ligand bases. Here we have demonstrated that specificity of substrate orientation is controlled by CobT enzyme homologs. This was especially evident in the CobT homologs we chose as examples which were each found to favor the synthesis of different ratios of 5- and 6-substituted products. We further demonstrated that *L. reuteri* seems to have the enzymatic ability to produce the uncharacterized [N9-Ade]Cba **21**, but

that for unknown reasons its cobamide does not retain this orientation. These findings point to an additional role for CobT in some circumstances in discriminating between lower ligand orientations during cobamide synthesis.

3.5   Materials and Methods

3.5.1   *Construction of CobT-overexpressing E. coli strains and purification of CobT homologs*

We have previously cloned *cobT* gene homologs from *Salmonella enterica*, *Lactobacillus reuteri, Veillonella parvula*, and *Sinorhizobium meliloti* into the pET28a plasmid to create IPTG-inducible, His-tagged expression vectors (Hazra et al., 2013). These were subsequently introduced into the *E. coli* strain BL21(DE3) for overexpression in LB medium with 25 µg/ml kanamycin and 250 µM isopropyl-β-D-thiogalactopyranoside (IPTG) (Hazra et al., 2013). CobT enzymes were harvested and purified by a nickel-NTA affinity chromatography as previously described (Hazra et al., 2013).

3.5.2   *In vitro CobT reactions with lower ligand substrates*

In vitro reactions with each CobT homolog were performed as before. Briefly, enzyme (10 µM), nicotinate mononucleotide (NaMN, 2 mM), lower ligand substrate (250 µM), and $MgCl_2$ (10 mM) were combined in 50 mM Tris-HCl at pH 7.5 for each reaction. The reaction was incubated at 37 ºC for 48 h then quenched at 100 ºC with 4% formic acid (Hazra et al., 2013). Some analyses required that the labile ribose-phosphate group be removed. This was accomplished by treatment with alkaline phosphatase as follows. Purified α-ribotide-phosphate compounds were prepared in 125 mM Tris at pH 7.9 with 10 mM $MgCl_2$ to which was added 40 units of calf intestinal phosphatase (New England Biolabs). The reactions were incubated on the bench top for 14 to 16 h then were heat inactivated at 100 ºC and filtered with a 10,000 MWCO filter (Pall). The substrates NaMN, 5-methoxybenzimidazole (5-OMeBza **1**) and adenine (Ade **15**) are commercially available and 5-hydroxybenzimidazole (5-OHBza **8**) was synthesized from 5-OMeBza **1** (Crofts et al., 2013).

3.5.3   *Purification of α-ribotide-phosphate and α-ribotide products*

Products of the CobT *in vitro* reactions were analyzed by high-performance liquid chromatography (HPLC) as previously described (Hazra et al., 2013). Briefly, the products of each reaction were filtered using a 10,000 MWCO filter (Pall) and injected onto an Agilent 4.6 x 150 mm XDB C-18 column with 5 µm pore size. The HPLC was run using 10 mM ammonium acetate at pH 6.5 and 100% methanol buffers on an Agilent 1200 series instrument equipped with an ultraviolet-visible (UV-vis) diode array detector. Analytes were monitored at 260 and 280 nm. Purification of individual peaks was carried out using the previously described method. Fractions were collected using an Agilent fraction collector unit. Pure fractions of each compound were pooled and stored overnight in a desiccation chamber to remove methanol following which the samples were freeze dried. The samples were resuspended in MilliQ filtered water and freeze-drying was repeated to remove volatile salts. Purified samples were stored at -80 ºC until further use.

3.5.4   Characterization of CobT products

UV-vis spectra were collected during HPLC purification using Agilent ChemStation software. For liquid chromatography tandem mass spectrometry analysis (LC-MS) an Agilent 6410 liquid chromatography-triple quadrupole mass spectrometer was used. Samples were run on the previously mentioned method using the XDB C-18 column as well as a 4.6 x 150 mm Zorbax SB-Aq column with 5 μm pore size with monitoring by MS2 scan. The conditions used have been previously described (Hazra et al., 2013). *In vitro* reactions containing CobT homologs from either *S. enterica* or *S. meliloti* were used to generate large quantities of benzimidazole-containing α-ribotides and α-ribotide-phosphates for further analysis. CobT homologs from *S. enterica* and *L. reuteri* were used to generate N7 and N9 α-ribotides and α-ribotide-phosphates respectively. For ^{1}H-NMR analysis, HPLC purified compounds were solubilized in deuterium oxide and freeze dried then brought to approximately 600 μl in deuterium oxide. The pH of each sample was verified to be between 6 and 6.5 prior to analysis. ^{1}H-NMR was performed using the California Institute for Quantitative Biosciences (QB3) 900 MHz Bruker instrument.

3.5.5   Strains of bacteria and culture conditions

Sporomusa ovata DSM 2662 was cultured at 30 ℃ anaerobically as previously described with 50 mM betaine as the carbon source (Mok and Taga, 2013). *Veillonella parvula* DSM 2008 was cultured anaerobically with an atmosphere of 80% N_2 and 20% CO_2 at 37 ℃ in the *S. ovata* medium with 10 g/L of sodium DL-lactate instead of betaine and with 5 μg/ml putrescine. *Lactobacillus reuteri* DSM 20016 was cultured anaerobically in a chemically defined medium at 37 ℃ under an atmosphere of 95% N_2 and 5% CO_2 (Santos et al., 2009). *Salmonella enterica* serovar Typhimurium strain LT2 was cultured aerobically with shaking at 37 ℃ in NCE media with 1,2-propoanediol as a carbon source and supplemented with 1 μM dicyanocobinamide (Gray and Escalante-Semerena, 2009b).

Sinorhizobium meliloti bluB::gus Gm[R] (*S. meliloti bluB*) (Campbell et al., 2006) was cultured aerobically at 30 ℃ in M9 minimal media as described in chapter 2 (Crofts et al., 2013; Maniatis et al., 1982). The *S. meliloti bluB*::gus Gm[R] *cobU*::*cobT$_{Lr}$* and *S. meliloti bluB*::gus Gm[R] *cobU*::*cobT$_{Se}$* strains were cultured similarly with the addition of 5 μM DMB or 1 μM cobalamin.

3.5.6   Construction of S. meliloti cobU genomic replacements

The strains *S. meliloti bluB*::gus Gm[R] *cobU*::*cobT$_{Lr}$* and *S. meliloti bluB*::gus Gm[R] *cobU*::*cobT$_{Se}$* were constructed as follows. The ORF of *S. enterica cobT* was amplified by PCR using the previously made vector pMP220+P$_{trp}$-*cobT$_{Se}$* as a template. The *L. reuteri* ORF had been previously codon optimized for expression in *S. meliloti* (Crofts et al., 2013). The ORF of *L. reuteri cobT* was amplified from the vector pMP220+P$_{trp}$-*cobT$_{Lr\ opt}$* containing the optimized ORF. Approximately 1 kb genomic fragments from immediately upstream and downstream of *cobU* were also amplified from an *S. meliloti* genomic DNA extract. The flanking 1 kb fragments and the *cobT* ORFs were assembled in the pK19 *mob sacB* vector (Schäfer et al., 1994) using the

Gibson cloning technique (Gibson et al., 2009). Each construct was introduced into the *S. meliloti bluB*::gus Gm[R] *cobU*::Spc[R] double mutant (Crofts et al., 2013) by triparental mating (Leigh et al., 1985). Screening for the replacement of the interrupted *cobU* gene by the *S. enterica* or *L. reuteri cobT* ORFs and loss of the suicide vector was achieved as previously described (Crofts et al., 2013; Leigh et al., 1985).

3.5.7 *Corrinoid extraction, HPLC analysis, and purification*

For corrinoid extractions, bacteria were cultured as above. *S. ovata* and *V. parvula* were cultured with either 500 µM 5-OMeBza **1** or 2 mM 5-OHBza **8**. *S. enterica* was cultured with 1 µM 5-OMeBza **1** and *S. meliloti* strains were cultured with 5 µM lower ligands unless otherwise noted. We found that production of [Ade]Cba was increased in *L. reuteri* when extra adenine (3 mg/L) was supplied in its medium. Corrinoids were extracted as in Chapter 2, but with an additional methanol wash step on the corrinoid extract pellet for higher yield.

All HPLC corrinoid analyses were performed using mobile phases of A, 0.1% formic acid in water and B, 0.1% formic acid in methanol. Corrinoid extracts were analyzed on an Agilent 1200 series HPLC with UV-Vis DAD using the previously described Zorbax SB-Aq column and method (Crofts et al., 2013).

[N7-Ade]Cba **18** was purified on a 9.4 x 250 mm Eclipse Plus C-18 column with 5 µm pore size. Purification was carried out using an in-line Agilent fraction collector using two sequential methods. First, the [Ade]Cba was eluted using a gradient HPLC method of 10 to 40% B over 17 minutes and 40 to 100% B over 0.1 minutes with a flow rate of 2 ml/min and temperature maintained at 45 ℃. Next, the cobamides were collected using the same column running an isocratic method at 30% B over 14 minutes with a flow rate of 1.5 ml/min and at 15 ℃.

Both isomers of [5(6)-OMeBza]Cba and [5(6)-OHBza]Cba were similarly purified using both gradient and isocratic HPLC methods. First, the cobamide isomers were collected as a single species on the 9.4 x 250 mm Eclipse Plus C-18 column using a gradient HPLC method consisting of 18% solvent B for 2.5 min, 18-45% solvent B over 18.5 min, and 45-100% solvent B over 0.5 min. The flow rate and temperature were 2 ml/min and 45 ℃ respectively. Second, the mixed isomers were injected onto the 4.6 x 150 mm Zorbax SB-Aq column and eluted using an isocratic method at a rate of 1 ml/min with the temperature maintained at 30℃ C. For [OMeBza]Cba isomers, 27% B was used in the elution. For the [OHBza]Cba isomers, 25% B was used in the elution.

3.5.8 *Characterization of cobamides*

UV-vis spectra were collected during HPLC purification using Agilent ChemStation software. Following HPLC purification cobamides were further characterized on an Agilent 6410 liquid chromatograph-triple quadrupole mass spectrometer with multiple reaction monitoring (LC-MS/MS) and exact mass MS using an Agilent 6520 Q-TOF instrument at the UC Berkeley QB3 Mass Spectrometry Facility (TOF-MS). LC-MS/MS was performed using the 4.6 x 150 Zorbax SB-Aq column and the analytical LC method from Chapter 2. Multiple reaction monitoring was

used to identify cobamides based on their characteristic transitions and retention times. For [OMeBza]Cba 679.3 and 149.1 m/z were used, for [OHBza]Cba 672.3 and 135.1 m/z were used (Mok and Taga, 2013). For TOF-MS samples were prepared in 50% acetonitrile in water and injected onto an Agilent 1200 series LC in line with an Agilent 6520 Q-TOF instrument. The error of the instrument was less than 5 parts per million. For ^{1}H-NMR, samples were prepared and analyzed in the same manner as the α-ribotide molecules earlier.

Chapter 4

Ubiquity of benzimidazoles, biomarkers for cobamides, in environmental samples and laboratory media

4.1   Summary

Cobalamin (vitamin B$_{12}$), an essential micronutrient for humans, is a member of the cobamide family of cofactors. Cobamides are found in a variety of environmental microbial communities and growing evidence suggests that cobamide exchange is an important dynamic in these communities. Here I have utilized a new bioassay to survey a variety of laboratory-associated and environmental samples for benzimidazoles, biomarkers of a variety of cobamides. The bioassay relies on the calcofluor-bright phenotype of *Sinorhizobium meliloti bluB*, which is indicative of the presence of free benzimidazoles. We have used our bioassay as well as a new LCMS method to detect and quantify 5,6-dimethylbenzimidazole, the lower ligand of cobalamin, and other benzimidazoles in liquid media with sub-nanomolar detection. Our analyses found that benzimidazoles are present in *S. meliloti* culture supernatants, laboratory agars, yeast extract, bovine rumen fluid, termite fecal pellets, soil, creek water, and a Winogradsky column derived from mud from the San Francisco Bay. Benzimidazoles were found in nearly all of the samples tested. In particular, benzimidazoles were plentiful in samples of rumen fluid and Winogradsky column water, both environments known to have a large microbial content. Interestingly, I also detected significant levels of benzimidazoles in commercial yeast extract although there is no known mechanism for its presence there. Our results show that benzimidazoles may be common components of several environments. It is unclear if these are the results of microbial cobamide biosynthesis or other processes, but this does indicate benzimidazoles may be widely available to cobamide producing organisms.

4.2   Introduction

Cobamides are cobalt-containing modified tetrapyrroles that function as cofactors for a variety of metabolic processes in animals, protists, and prokaryotes (Roth et al., 1996). Vitamin B$_{12}$ (cobalamin, Figure 4.1A) is the best studied cobamide and is well known for its importance in human health. Cobalamin and other cobamides also function as cofactors for several ecologically important processes in prokaryotes. For example, cobamides are used as cofactors for methanogenesis, acetogenesis, the creation of toxic methyl mercury, and the detoxification of perchloroethylene (He et al., 2007; Parks et al., 2013; Ragsdale and Pierce, 2008; Roth et al., 1996). Cobamides are also found in host-associated and environmental microbial communities including the human gut (Allen and Stabler, 2008; Girard et al., 2009; Wakayama et al., 1984). Because cobamides are involved in so many important metabolic processes, there is an increasing need for high-throughput methods to survey the environment for cobamides.

Cobamides are modified tetrapyrroles that contain a central cobalt ion that coordinates two axial ligands. The β-ligand, also known as the upper ligand, is the catalytic site of the cofactor. Cobamides containing an adenosyl group as the upper ligand generally help catalyze structural

rearrangement reactions via a radical intermediate. Cobamides can contain a methyl group as the upper ligand to catalyze methyltransfer reactions (Banerjee and Ragsdale, 2003). The α-ligand, or lower ligand, is the main site of structural diversity in the cobamides. Cobalamin contains 5,6-dimethylbenzimidazole (DMB) as its lower ligand (Figure 4.1A). Other cobamides contain benzimidazoles (Figure 4.1B), purines (such as adenine), or phenolics (such as *p*-cresol) as lower ligands (Renz, 1999).
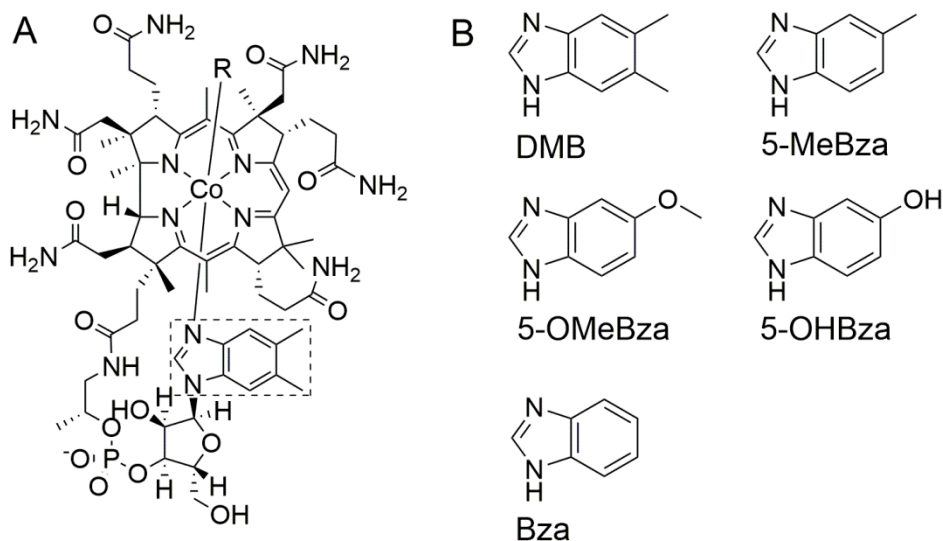

Figure 4.1 The calcofluor phenotype of *S. meliloti* is influenced by cobalamin-dependent metabolism. A) Structure of cobalamin. The lower ligand 5,6-dimethylbenzimidazole (DMB) is boxed. B) Structures of lower ligands used in this study. Abbreviations: 5-methylbenzimidazole (5-MeBza), 5-methoxybenzimidazole (5-OMeBza), 5-hydroxybenzimidazole (5-OHBza), benzimidazole (Bza).

There is growing evidence that cobamides and their biosynthetic precursors are shared among microbes in complex communities. For example, bioinformatic and experimental studies indicate that approximately half of bacteria that are capable of using cobamides do not produce cobamides *de novo*, and instead import cobamides produced by other organisms in their environment (Zhang et al., 2009). In addition, cobamide precursors such as cobinamide (Cbi), α-ribazole, and benzimidazoles are taken up and used by bacteria that lack the enzymes for their biosynthesis (Gray and Escalante-Semerena, 2010; Yi et al., 2012). In some cases complete cobamides are taken up and have their lower ligands exchanged in a process called remodeling (Gray and Escalante-Semerena, 2009a). Many bacteria such as *Escherichia coli* possess the *cobUTSC* genes necessary for lower ligand attachment but lack other genes for *de novo* cobamide synthesis, underscoring the importance of the availability of free benzimidazoles for the metabolism of some bacteria (Roth et al., 1996).

We recently reported that DMB and 5-methylbenzimidazole (5-MeBza), another cobamide lower ligand (Figure 4.1B), are present in microbial enrichment communities that perform cobamide-dependent dechlorination of trichloroethylene (Y. J. Men, E. C. Seth, S. Yi, T. S. Crofts, R. H. Allen, M. E. Taga, and L. Alvarez-Cohen, submitted). DMB has also been reported as a contaminant of common lab agar, pointing to its possible presence in multiple environments

(Anderson et al., 2008). Here, I expand on these observations by developing an extraction procedure, and bioassay for the purification, detection, and quantification of benzimidazoles, including DMB, from a variety of environmental sources. The bioassay is based on the "calcofluor-bright" (CFB) phenotype of *bluB* mutants of the bacterium *Sinorhizobium meliloti*, a phenotype that is rescued by the addition of DMB (Campbell et al., 2006), (Taga 2007). The CFB phenotype is defined as abnormally bright fluorescence on LB agar plates containing calcofluor, a dye (also used in laundry detergents) that fluoresces when bound to the bacterial exopolysaccharide succinoglycan (Leigh et al., 1985). The *bluB* gene encodes the enzyme responsible for the oxygen-dependent biosynthesis of DMB (Campbell et al., 2006; Gray and Escalante-Semerena, 2007; Taga et al., 2007). The results of this bioassay were compared to a newly developed liquid chromatography-tandem mass spectrometry (LCMS) analysis in order to quantify specific benzimidazoles in each sample. The results of this study show that benzimidazoles are present at picomolar levels in a variety of environmental samples, as well as in common laboratory media components.

4.3   Results

4.3.1   *The calcofluor bright phenotype reflects DNA stress in S. meliloti.*

To investigate the extent to which the calcofluor phenotype of the *S. meliloti bluB* mutant is influenced by the DMB concentration, a range of concentrations of DMB were tested for their effect on the calcofluor phenotype by a filter disc diffusion assay. This showed that the size of the calcofluor-dim zone surrounding the DMB-soaked disc was dependent on the DMB concentration, suggesting that this phenotype could be used for quantification of DMB (Figure 4.2A).

The observation that cobalamin deficiency in *S. meliloti* results in a CFB phenotype suggests that one or more enzymes that use cobalamin as a cofactor influence the level or structure of succinoglycan. Of the three known cobalamin-dependent enzymes in *S. meliloti*, two, methionine synthase and methylmalonyl CoA mutase, were previously shown not to significantly affect the calcofluor phenotype of *S. meliloti* and thus cannot explain the CFB phenotype of the *bluB* mutant (Campbell et al., 2006). Therefore, we hypothesized that the CFB phenotype is a result of loss of activity of the cobalamin-dependent (Class II) ribonucleotide reductase (RNR) encoded by *nrdJ* (Cowles and Evans, 1968; Lawrence and Stubbe, 1998). To test whether the CFB phenotype of the *bluB* mutant is due to insufficient RNR activity of the cobalamin-dependent RNR NrdJ, we expressed an *E. coli* cobalamin-independent (Class I) RNR encoded by *nrdA* and *nrdB* (Carlson et al., 1984) on a plasmid in a *bluB* mutant background and found that the strain produces a dim phenotype on calcofluor plates (data not shown). To further examine the link between the CFB phenotype and reduction in RNR activity, we used the *Sm*NrdAB[+] strain ($\Delta nrdJ$ with the plasmid expressing *E. coli nrdAB*) (Taga and Walker, 2010). The effect of hydroxyurea, a specific inhibitor of Class I RNRs (Elford, 1968), was tested by growing this strain on an agar plate with a hydroxyurea-soaked filter disc. The addition of hydroxyurea leads to the formation of a zone of inhibition, and a CFB phenotype is evident in the colonies surrounding the zone of inhibition (Figure 4.2BC), further indicating an association between the CFB phenotype and a reduction in RNR activity.
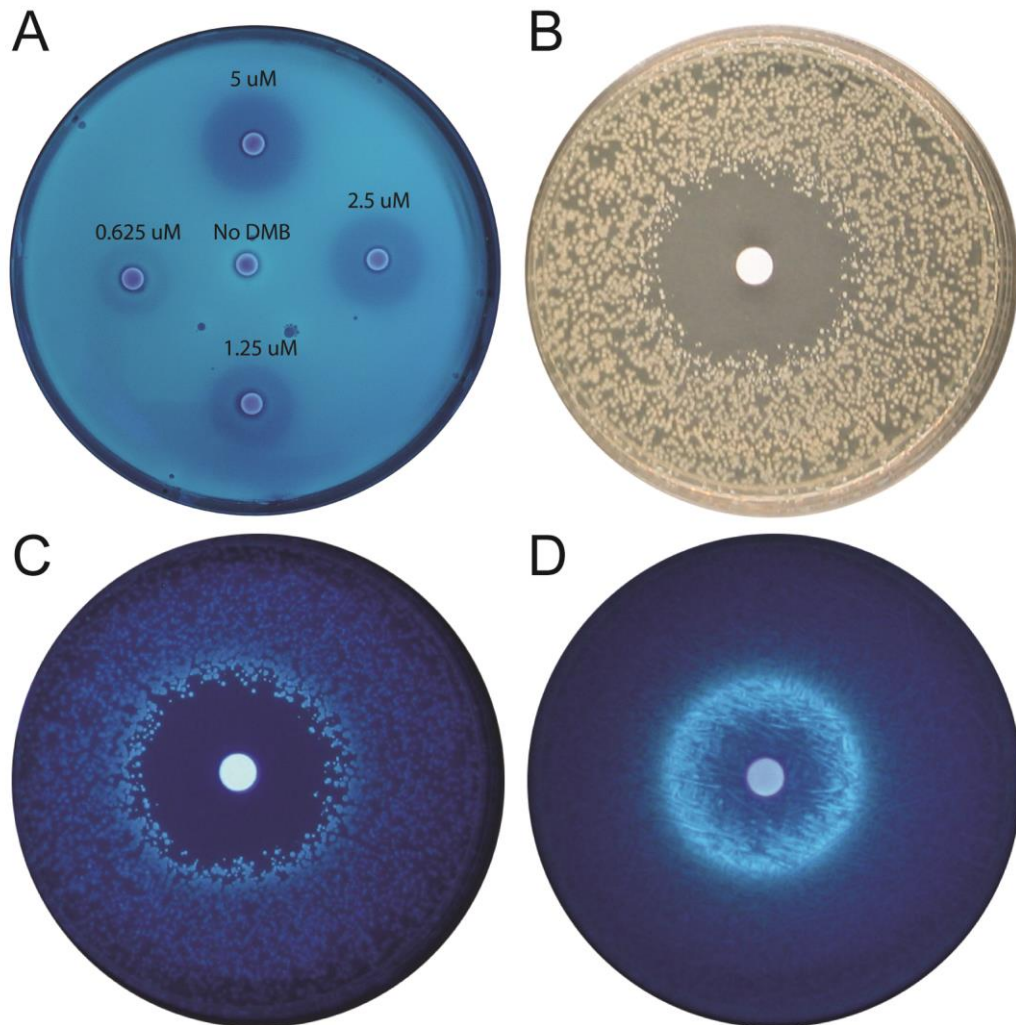
Figure 4.2 The calcofluor bright phenotype of *S. meliloti* due to DNA stress. A) Calcofluor fluorescence response of *S. meliloti bluB* to supplementation with exogenous DMB on sterile filter discs. Clockwise from the top: 10 μl of 5 μM; 2.5 μM; 1.25 μM; or 0.625 μM DMB; or (center) water was applied to each disc. B) Lawn of *Sm*NrdAB+ with hydroxyurea applied to a filter disc in the center of the plate photographed under white light. C) Same as in B) photographed under UV light. D) A lawn of wild type *S. meliloti* strain Rm1021 on a calcofluor plate with nalidixic acid applied to the filter disc.

One explanation for the induction of the CFB phenotype by inactivation of RNR is that it is a response to DNA stress, since reduction in RNR activity results in an inadequate supply of deoxyribonucleotides (dNTPs) and leads to replication fork arrest (Odsbu et al., 2009). To explore this possibility, we exposed wild-type *S. meliloti* to nalidixic acid, a DNA gyrase inhibitor that induces DNA damage (Pommier et al., 2010). Nalidixic acid (Figure 4.2D) induced a CFB phenotype adjacent to the zone of inhibition in the filter disc assay, similar to the phenotype of *Sm*NrdAB[+] exposed to hydroxyurea, suggesting that the CFB phenotype of the *bluB* mutant could be a response to DNA damage.

4.3.2   *Calcofluor fluorescence phenotype of* S. meliloti bluB *in liquid media*

Previous experiments have focused on the binary relationship between the CFB phenotype of *S. meliloti bluB* and the presence or absence of DMB and/or cobalamin when grown on solid media (Campbell et al., 2006; Taga et al., 2007). Because I observed that the CFB phenotype shows a dose-dependent relationship with the concentration of DMB (Figure 4.2A), I became interested in harnessing the CFB phenotype as a method to detect DMB and possibly other cobamide lower ligand bases. I therefore investigated whether the calcofluor phenotype would function in liquid media, which would be more amenable to quantification and high throughput. To assay the calcofluor phenotype in liquid cultures, wild-type *S. meliloti* and *S. meliloti bluB* were grown in the presence and absence of 0.5 µM DMB or 1 µM cobalamin. Following addition of calcofluor, a multiwell plate reader was used to quantify the fluorescence levels. This experiment recapitulated the calcofluor phenotype on solid media: the *bluB* mutant displayed 20-fold higher fluorescence than wild-type *S. meliloti* (Figure 4.3A).The *S. meliloti bluB* cultures grown with either DMB or cobalamin also displayed a dim fluorescent phenotype similar to wild-type *S. meliloti* (Figure 4.3A). Increasing the amount of time that the bacteria grew in the presence or absence of DMB prior to the addition of calcofluor led to a greater separation of the dim and bright phenotypes (Figure 4.3B). After growth, the optimal amount of time to incubate the cells with calcofluor is 5 h. Longer development of the cultures with calcofluor led to an increase in the fluorescence of both the bright and dim strains (Figure 4.3C).

My next goal was to determine if the CFB phenotype showed a dose-dependent relationship with DMB concentration. Cultures of *S. meliloti bluB* were grown with DMB at concentrations ranging from 244 pM to 1 µM. After reaching stationary phase (approximately 65 h) the cultures were incubated with calcofluor for 5 h before fluorescent measurements were taken. The results in Figure 4.3D show a relationship between fluorescence, and to a lesser extent the final O.D.$_{600}$ of the cultures, and the concentration of DMB added. The linear range of these dose-response curves can therefore be used to quantify DMB concentrations in unknown samples. After repeated experiments, I found that the concentration of DMB that results in half maximal fluorescence (EC$_{50}$) can range from approximately 1 to 20 nM.
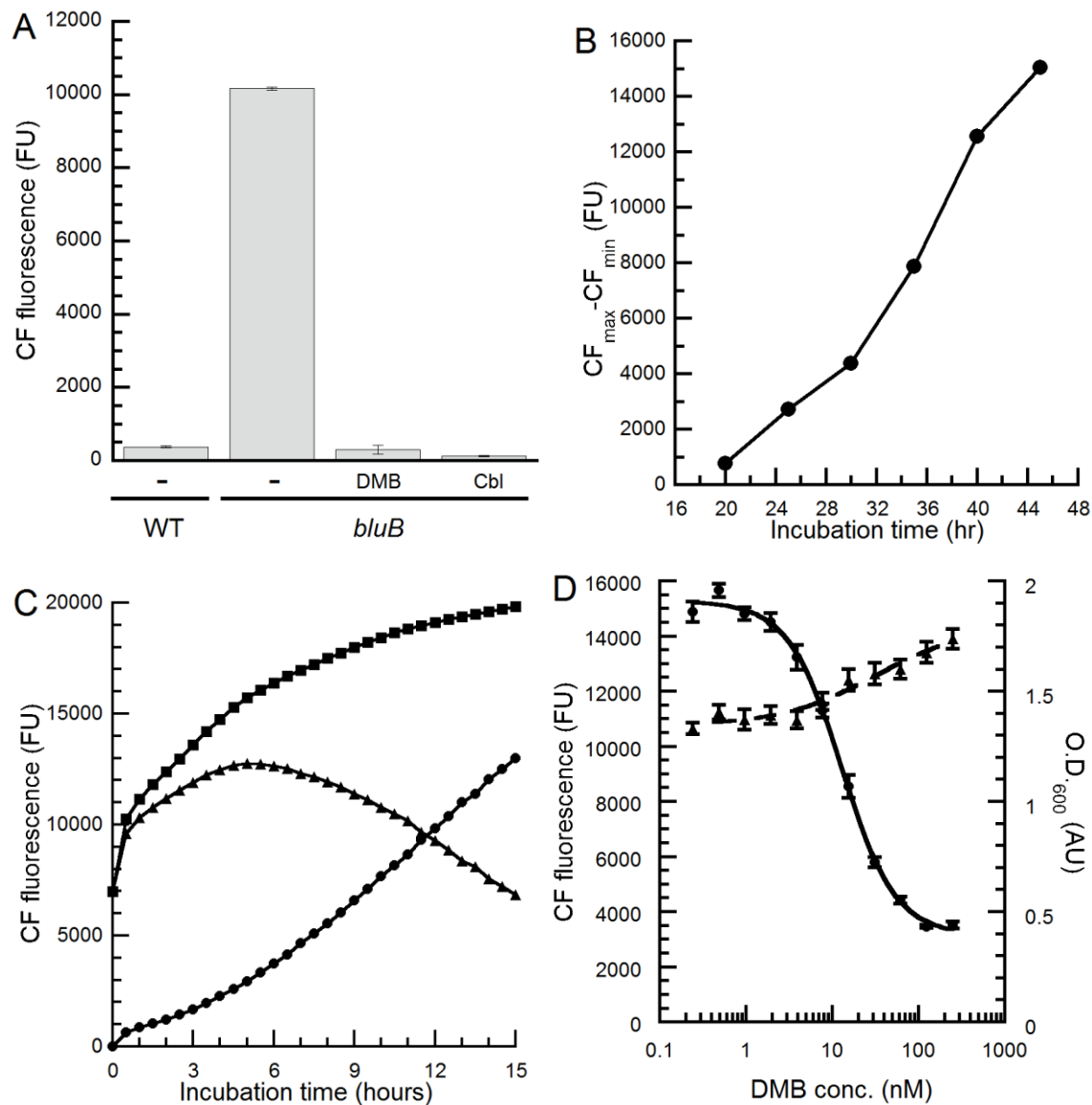
Figure 4.3 The *S. meliloti* calcofluor bioassay to quantify benzimidazoles. A) Calcofluor (CF) fluorescent phenotype of liquid cultures of wild type *S. meliloti* strain Rm1021 (WT), an *S. meliloti bluB* mutant (*bluB*), or *S. meliloti bluB* grown in the absence (-) or presence of 0.5 μM DMB (DMB) or 1 μM cobalamin (Cbl). Error bars represent the standard error of three independent experiments. B) Difference between bright (CFmax) and dim (CFmin) phenotypes as a function of *S. meliloti bluB* incubation period with (CFmin) or without (CFmax) DMB. C) Development of calcofluor fluorescence over time. After growth of *S. meliloti bluB* with and without DMB for approximately 48 h, calcofluor was added and fluorescence was monitored. Squares represent the bright phenotype, circles the dim phenotype, and triangles the difference between the two. 5 h was identified as the optimal incubation time. D) The dose-response relationship between concentration of DMB and either the calcofluor fluorescence (left axis, circles) or the culture density (O.D.600) (right axis, triangles). Error bars represent the standard error of six independent experiments.

4.3.3 The CFB phenotype of S. meliloti bluB *is also rescued in a dose-dependent manner by other cobamide-associated benzimidazoles*

As described in Chapter 2, I previously found that *S. meliloti bluB* can incorporate a variety of benzimidazoles in addition to DMB as cobamide lower ligands, and that many of these cobamides can support the growth of *S. meliloti* (Crofts et al., 2013; Hazra et al., 2013). I therefore tested whether the bioassay could be used to detect benzimidazoles other than DMB. To determine to what degree the CFB phenotype of *S. meliloti bluB* can be rescued by other benzimidazoles, I used the bioassay to examine the calcofluor response to a variety of benzimidazoles (Figure 4.4). As expected based on the results shown in Chapter 2, all of the benzimidazoles tested were also able to rescue the CFB phenotype (Figure 4.4), though their EC_{50} values were approximately 5- to 45 -fold higher than for DMB. The addition of cobalamin also rescued the CFB phenotype of *S. meliloti bluB* in a dose-dependent fashion, though 14-fold higher concentrations were required (Figure 4.4F). Therefore, this calcofluor-based bioassay could be useful for detecting a variety of benzimidazoles and possibly cobamides.

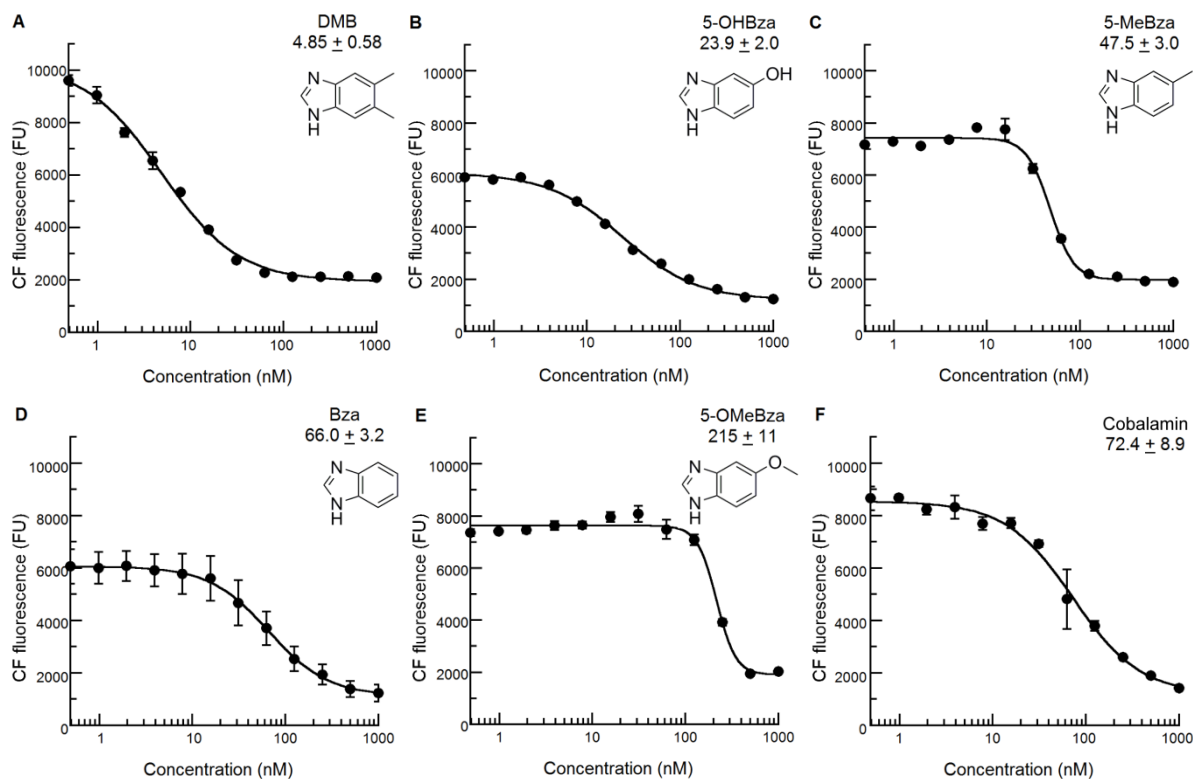**Figure 4.4 Calcofluor response of *S. meliloti bluB* to other benzimidazoles and cobalamin.** The ability of a variety of benzimidazoles and cobalamin to rescue the calcofluor bright phenotype of *S. meliloti bluB* was assayed. The structure and EC_{50} of each compound are displayed in the upper right corner of each graph along with the curve-fit error. Error bars represent the standard error of three independent experiments.

4.3.4   Detection of benzimidazoles in environmental samples

To test whether the calcofluor bioassay can quantify DMB present in complex mixtures, I first conducted the assay with serial dilutions of spent media of various *S. meliloti* cultures that contain a range of DMB concentrations. The strains tested were wild type *S. meliloti* Rm1021, the *bluB* mutant, and *bluB* mutant strains containing plasmids expressing *bluB* at different levels (Campbell et al., 2006; Yu et al., 2012). The plasmid pMS03-*bluB* expresses *bluB* at a high level in *S. meliloti* such that when grown on solid media adjacent to a *bluB* mutant, it can induce a calcofluor-dim phenotype in neighboring colonies (Yu et al., 2012). The plasmid pJL1031 does not have this ability, suggesting that it expresses BluB at lower levels (K. Kennedy, personal communication). The concentration of DMB present in these samples was estimated based on the average fluorescence of dilutions of triplicate samples in comparison to triplicate samples of DMB standards. Only wells with fluorescence that fell within the range of the DMB standard curve were used. Based on these calculations, 6- to 10-fold higher concentrations of DMB were detected in the pMS03-*bluB* spent media compared to the strain containing pJL1031 and Rm1021, and no DMB was detected in spent media of the *bluB* mutant (Table 4.1). This indicates that the bioassay can be used to measure DMB in biological samples.

Table 4.1 Benzimidazoles levels in complex samples as determined by the *S. meliloti bluB* bioassay and LCMS

Sample	DMB equivalents detected by bioassay	Benzimidazoles detected by LCMS					
		DMB	MeBza	OMeBza	Bza	OHBza	Expected bioassay equivalents[*]
Cell free supernatants							
Wild-type (nM)	31.8 ± 6.24						
bluB (nM)	ND[+]						
pMSO3 (nM)	307 ± 54.8						
pJl1031 (nM)	50.3 ± 15.0						
Laboratory agars							
Technical (pmol/g)	2.78 ± 1.16	0.242	0.286	0.231	2.48	6.67	1.81
Plant cell culture tested (pmol/g)	0.950 ± 0.458	ND	0.196	ND	1.01	ND	0.0944

Noble (pmol/g)	2.48 ± 1.66	0.307	0.370	0.347	0.974	0.942	0.615
Agarose (pmol/g)	1.84 ± 1.52	0.845	0.300	0.339	0.653	ND	0.931
Yeast extract (pmol/g)	26.0 ± 1.21	6.89	7.42	1.84	4.99	1.10	8.28
Cellulytic microbial communities							
Rumen 1 (pM)	262 ± 76.1	490	ND	ND	1500	ND	600
Rumen 2 (pM)	367 ± 33.9	562	ND	ND	1250	ND	654
Termite (pmol/g)	0.851 ± 0.255	0.181	0.259	ND	0.442	ND	0.240
Environmental							
Eucalyptus grove soil (pmol/g)	0.694 ± 0.291	ND	ND	ND	0.483	ND	0.0355
Creek bank soil (pmol/g)	0.531 ± 0.361	ND	ND	ND	0.197	ND	0.0145
Creek water (pM)	7.36 ± 3.35	ND	ND	ND	4.86	ND	0.357
Winogradsky column water (pM)	104 ± 35.5	23.9	24.5	ND	22.6	ND	28.1

[*] Expected bioassay equivalents, see the Materials and Methods section for detailed description of calculation
[+] ND (not detected)

In order to assay environmental samples that may contain benzimidazoles at concentrations below the detection limit of the bioassay, I developed a two-step process to concentrate and purify benzimidazoles from complex mixtures (see Materials and Methods section for details). The yield of each benzimidazole varied, as the level of benzimidazole standards retained during purification ranged from 0 to 70% (Figure 4.5). The low retention of 5-OHBza is mitigated in part by the greater sensitivity of *S. meliloti bluB* towards this substrate compared to other benzimidazoles (Figure 4.4B).

It has previously been reported that DMB is a contaminant of standard laboratory agar and that noble agar has less DMB contamination (Anderson et al., 2008). I have also observed this indirectly based on the amount of cobalamin produced by *S. meliloti bluB* when grown on minimal media plates prepared with different agars (data not shown). To determine the

concentration of DMB in laboratory agars, I tested extracts from three grades of agar in addition to agarose. The bioassay detected 1-3 pmol/g benzimidazoles (expressed as expected bioassay equivalents) in each grade of agar, though levels did not differ significantly from each other (Table 4.1). Based on these values, I estimate that a generic agar plate, containing approximately 25 ml of 1.5% agar, would contain less than 2 pmol of DMB, or a concentration of ~75 pM. This is consistent with our earlier findings, in that this concentration is too low to rescue the CFB phenotype (Figure 4.3D). In addition to agars, I also examined commercial yeast extract. Surprisingly, I saw approximately 10-fold higher levels of DMB in the yeast extract sample (26 pmol/g) than the agar samples. Based on this value, a typical LB agar plate would contain ~3 pmol of DMB from yeast extract (LB medium would contain ~130 pM). Again this is below the level required to rescue the CFB phenotype but may explain the ability of the *S. meliloti bluB* mutant to form colonies when DMB or cobalamin is not provided.
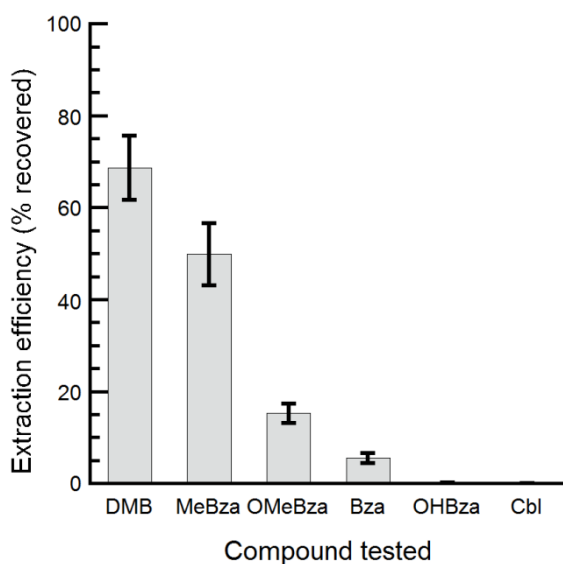
Figure 4.5 Efficiency of extraction of benzimidazoles and cobalamin. Aqueous solutions of standards of the indicated compounds were subjected to liquid extraction with ethyl acetate and C18 solid phase extraction. The recovery of each compound was quantified by high performance liquid chromatography (HPLC). Error bars represent the standard error of three replicates.

The gastrointestinal communities of animals are sources of diverse cobamides. For example, bovine rumen, cow and pig manure, human feces, and wood-feeding insects have been shown to have concentrations of corrinoids up to 2 μg/g (Allen and Stabler, 2008; Brown et al., 1955; Girard et al., 2009; Wakayama et al., 1984). Because corrinoid remodeling, an activity that requires the availability of free lower ligands, has been shown to occur in some microbial communities, I hypothesized that environments with high concentrations of corrinoids would also contain significant amounts of free lower ligand bases (Allen and Stabler, 2008; Yi et al., 2012). Rumen fluid and fecal pellets from a laboratory colony of termites were analyzed by the calcofluor bioassay to determine the benzimidazole content. I found ~300 pM benzimidazoles in the two rumen fluid samples (Table 4.1). In contrast, the termite fecal pellets appear to have

lower concentrations of benzimidazoles on a per gram basis (approximately 0.9 pmol/g) (Table 4.1).

Finally, I analyzed samples taken from locations around the UC Berkeley campus. Samples were collected and extracted on the same day and came from the top soil of the campus eucalyptus grove, gravely soil adjacent to Strawberry creek, and water from Strawberry creek. The two soil samples contained levels of benzimidazoles similar to the termite gut (0.5-0.7 pmol/g) as did the creek water (7 pM) (Table 4.1). I also analyzed the benzimidazoles present in the top layer of a laboratory Winogradsky column established over 20 years ago by Prof. Sydney Kustu with mud from the San Francisco bay. Benzimidazole levels were comparable to those found in rumen fluid (100 pM) (Table 4.1).

To confirm our bioassay results, and to analyze the presence of individual benzimidazoles, we analyzed the same samples by LCMS. The method used was developed in collaboration with Yujie Men as part of our lab's ongoing collaboration with the Alvarez-Cohen lab. A manuscript describing the LCMS method and its application for the analysis of benzimidazoles in a trichloroethene-degrading microbial enrichment culture has been submitted for publication.

The bioassay and LCMS methods were largely similar, though with some significant differences (Table 4.1). The values calculated from LCMS analysis in Table 4.1 are systematically lower than those found using the bioassay. The reason for this could be that for several samples certain benzimidazoles were detected by LCMS, but were found in concentrations below the accurate range of the instrument. As an example, eucalyptus grove soil was found by bioassay to contain approximately 0.7 pmol/g of DMB equivalents. In contrast the LCMS data indicate the presence of only 0.035 pmol/g. The discrepancy is due in part to the exclusion of values for DMB, 5-MeBza, 5-OMeBza, and 5-OHBza as these fell outside the limit of quantification of the LCMS method. For details on the calculation of expected bioassay equivalents see the Materials and Methods section.

4.4   Discussion

Due to its medical significance, cobalamin has often been studied with an eye towards human health. This is despite the fact that it is cofactor synthesized solely by prokaryotes for their own use (Roth et al., 1996). Cobamides contribute to microbial processes involving methyl transfer reactions, carbon skeleton rearrangements, and reductive dechlorination (Banerjee and Ragsdale, 2003; Gruber et al., 2011). While not necessarily directly impacting human health, these environmental processes nevertheless are environmentally and industrially important. It is in this context that I have demonstrated a surprising ubiquity of free benzimidazoles available in the environment.

Our bioassay takes advantage of the CFB phenotype of the *S. meliloti bluB* mutant which can be modulated by DMB and other cobamide-associated benzimidazoles. Our experiments suggest that the CFB phenotype is a reflection of altered succinoglycan production, which itself is the result of DNA stress due to low ribonucleotide reductase activity in the absence of cobalamin

(Figure 4.2BCD). Increased exopolysaccharide production is a known side effect of *S. meliloti* growth under stressful conditions (Rinaudi et al., 2006; Soto et al., 2006).

The tendency of some *S. meliloti* mutants, or wild-type strains grown under certain conditions, to develop a CFB phenotype is well established and is an experimentally useful phenotype (Campbell et al., 2006; Finan et al., 1985; Leigh et al., 1985). The requirement of solid media and its qualitative nature limit the applications of this phenotype. By adapting the CFB phenotype to liquid media I have made it amenable to high throughput analyses. With the use of a plate reader, the assay can also be made quantitative.

When I applied our bioassay to a variety of samples to determine its utility, I was able to detect the presence of DMB or other benzimidazoles (all of which were grouped together as "DMB equivalents") in almost all of the samples (Table 4.1). LCMS analysis of our bioassay samples, using a method developed Men *et al* (Y. J. Men, E. C. Seth, S. Yi, T. S. Crofts, R. H. Allen, M. E. Taga, and L. Alvarez-Cohen, submitted), complemented the bioassay results (Table 4.1). The bioassay providing a high throughput but coarse assay and LCMS provided the ability to differentiate and quantify each type of benzimidazole. The two methods did diverge significantly in the cases of some samples. Notably, the bioassay detected higher levels of benzimidazoles in environmental samples than the LCMS. Two issues might explain this difference. First, although Table 4.1 records many of these samples as having "not detectable" levels benzimidazoles, many of these were in fact detected but fell below the range at which accurate concentrations could be measured. In contrast, the bioassay in effect pools the levels of all benzimidazoles, potentially allowing for higher or lower sensitivity depending on the biological activity of the benzimidazoles present. Second, the LCMS method is only able to detect and quantify benzimidazoles that it has been programmed to search for. In this way new lower ligands might not be detected by LCMS but may still be able to rescue to CFB phenotype of the *bluB* mutant.

I previously observed that an *S. meliloti cobU* (the *cobT* homolog) mutant, unlike the *bluB* mutant, is unable to grow on LB without cobalamin supplementation (Crofts et al., 2013). At the time, this phenotype was unexpected because *bluB* and *cobT* function in the same pathway and therefore should have the same mutant phenotypes. It now appears likely that trace amounts of DMB present in components of growth media are sufficient to support growth of the *bluB* mutant, whereas a functioning lower ligand attachment pathway, which includes *cobU*, is required for scavenging of DMB. Both the yeast extract in LB and agar itself may provide a source of lower ligands. This study also proved the interesting observation that benzimidazoles are present in commercial yeast extract, despite the absence of the *bluB* gene or any cobamide biosynthetic or dependent enzymes in yeast. We do not know the origin of these benzimidazoles, though it is possible that they may have been formed abiotically from flavins during the production of the material (Maggio-Hall et al., 2003).

The other samples containing significantly higher levels of benzimidazoles were extracts from rumen fluid and a Winogradsky column (Table 4.1). These environments are known to contain high microbial content and, in the case of rumen fluid, to contain cobamides themselves making the detection of benzimidazoles reasonable (Girard et al., 2009). The prevalence of benzimidazoles in rumen fluid and the TCE-dechlorinating enrichment community is also

interesting biochemically. The anaerobic route of benzimidazole biosynthesis is unknown, and the appearance of benzimidazoles in these environments indicates their production in anaerobic environments.

The Winogradsky column water and rumen fluid contained over 100 nM DMB equivalents. As is shown in Figure 4.3D, these concentrations would be sufficient to fully rescue the CFB phenotype of *S. meliloti bluB*, and may be sufficiently concentrated to be utilized by other environmental organisms as well. Along these lines, our laboratory has previously shown that lower ligand bases provided at a concentration of 37 nM were sufficient to allow cobamide-dependent growth of *D. mccartyi* via its corrinoid salvaging and remodeling pathways when a corrinoid is also added (Yi et al., 2012). Our laboratory also previously examined the growth inhibition of the acetogenic bacterium *Sporomusa ovata* in the presence of benzimidazoles. That study reported that a benzimidazole concentration over 300 nM was required for guided biosynthesis leading to growth inhibition, a level slightly higher than the physiological concentrations we detected (Mok and Taga, 2013). It therefore appears that the benzimidazole levels I detected fall into an intermediate zone, where levels are high enough to support growth of organisms requiring an exogenous lower ligand base but not high enough to outcompete the native lower ligand bases of cobamide producing bacteria.

Often bioassay screens represent the first step when searching for small molecules of interest in complex samples. The high-throughput and low cost nature of bioassays makes them effective screens for initial detection of compounds of interest. The detection of benzimidazoles in diverse biological samples opens up the possibility that cobamide remodeling and the sharing of benzimidazoles are more widespread than previously appreciated.

4.5   Materials and Methods

4.5.1   *Bacterial culturing*

All *S. meliloti* strains are derivatives of the wild type strain Rm1021. *S. meliloti bluB::gus Gm^R* (Campbell et al., 2006) cultures were grown with aeration at 30 °C on LB agar plates, in liquid LB supplemented with 2.5 mM $CaCl_2$, 2.5 mM $MgSO_4$ (LBMC), or in M9 minimal medium (M9) with 0.2% sucrose, 10 µg/L biotin, 1 mg/mL methionine, and 10 µM $CoCl_2$ (Maniatis et al., 1982). When necessary, antibiotics were added at the following concentrations (µg/ml): gentamicin, 25; tetracycline, 10; streptomycin, 500; and spectinomycin, 100.

4.5.2   *S. meliloti calcofluor phenotype on agar plates*

To assay the calcofluor phenotype on solid media, 100 µl of a bacterial culture was spread on LB agar containing 0.02% calcofluor and 10 mM HEPES pH 7.4. Sterile filter discs containing 10 µl of each compound being tested were placed on the surface. The plates were incubated for 2-3 d at 30 °C and photographed under ultraviolet (UV) light.

4.5.3   *S. meliloti bioassay to detect benzimidazoles*

Three independent cultures of *S. meliloti bluB* were grown in LBMC for 40-48 h. The cultures were diluted to an O.D.$_{600}$ of 0.04 in M9 and mixed with known concentrations of DMB or dilutions of a sample to be tested in a total volume of 200 µl per well in 96-well microtiter plates. All solutions were sterilized by passing through a 0.22 µm filter prior to use. The edges of the plates were sealed with parafilm and the plates were placed on a moist paper towel in a sealed plastic bag to limit evaporation. The plates were incubated in a shaking incubator at 30 ℃ for 55-70 h. Calcofluor was then added to a final concentration of 80 µg/ml, bringing the total volume to 300 µl per well, and the plate was incubated at room temperature in the dark for 5 h. Bioassay results are reported as DMB equivalents because the bioassay cannot differentiate unknown benzimidazoles. The fluorescence phenotype was measured using a BioTek Synergy2 plate reader with excitation at 360 nm and emission measured at 460 nm. Sigmoidal curves were solved using the KaleidaGraph program with the equation m1+(m2-m1)/(1+(x/m3)^m4), where m1, m2, m3, m4, and x correspond to the minimum fluorescence value, the maximum fluorescence value, the EC$_{50}$, the slope at EC$_{50}$, and the dilution factor of the sample, respectively. The error was calculated by the curve-fit program.

4.5.4   Bioassay sample preparation

5-hydroxybenzimidazole was synthesized from 5-methoxybenzimidazole as previously described (Crofts et al., 2013; Renz et al., 1993).

To obtain cell-free supernatants for analysis, cultures of Rm1021 and *bluB* mutant strains containing the plasmids pMS03-*bluB* (Yu et al., 2012) or pJL1031 (*bluB* cloned under its native promoter in pFAJ1700) (Campbell et al., 2006) were grown in triplicante in M9. All samples were diluted to an O.D.$_{600}$ of 0.85 and cell-free supernatants were collected by centrifugation of each culture at 9,000 g for 5 min followed by filtration through a 10,000 MWCO filter (PALL). Samples were stored at -20 ℃ prior to use.

7.5 g of technical grade agar (Difco 281210), plant cell culture tested agar (Sigma A7921), noble agar (US Biological C10021012), agarose (Fisher BP160-500), and yeast extract (Becton Dickinson 211929) was suspended in 300 ml of phosphate buffered saline adjusted to pH 8.0 (1X PBS). Rumen fluid from two cows fed on a high forage diet (UC Davis) was collected in September, 2009, and stored at -20 °C. 35 ml aliquots of clarified rumen fluid were brought to 75 ml in 1X PBS adjusted to pH 8.0. 10 g of termite fecal pellets obtained from a laboratory colony of *Zootermopsis* sp. (a gift from Erica Seth) was suspended in 75 ml of 1X PBS and the pH was adjusted to 8.0. Particulates were removed from samples by passage through a syringe containing a cotton plug.

Topsoil from the UC Berkeley campus eucalyptus grove and soil adjacent to Strawberry Creek on the UC Berkeley campus were collected in September, 2013. 10 g of each sample was suspended in 75 ml of 1X PBS and brought to pH 8.0. Water from Strawberry Creek was collected from the same location. The Winogradsky column, inoculated over 20 years ago with mud from the San Francisco bay, was a gift from Sydney Kustu. 1 L of creek water and 75 ml of the upper portion of the Winogradsky column were mixed with NaCl, KCl, Na$_2$HPO$_4$, and

K$_2$HPO$_4$ to adjust the salt and buffer content to that of 1X PBS, and the pH was subsequently adjusted to 8.0.

4.5.5 *Extraction of benzimidazoles*

The following procedure was used to extract benzimidazoles from the above preparations, except the *S. meliloti* supernatants which were used without extraction. Samples were mixed with an equal volume of ethyl acetate and shaken vigorously in a separatory funnel. The organic layer was removed, and the extraction was repeated with the aqueous layer by adding an equal volume of ethyl acetate again. The organic phase from each extraction was pooled, and solvent was removed by rotary evaporation at 50 °C. Samples were resuspended in 1X PBS and applied to a 360 mg C18 Sep-Pak (Waters) previously activated with 3 ml of methanol and equilibrated with 6 ml deionized H$_2$O. The cartridge was washed with 6 ml of 30% methanol and eluted with 2 ml of 70% methanol. The eluate was collected and dried under reduced pressure at 45 °C. The final product was resuspended in deionized H$_2$O, filtered through a 10,000 MWCO spin filter (Pall) and stored at -20 °C prior to analysis.

4.5.6 *HPLC and LCMS verification of extraction efficiency and DMB concentrations in bioassay samples*

To determine the extraction efficiency of each benzimidazole shown in Figure 4.1B, 250 nmol of each of the above components was resuspended in 15 ml of 1X PBS pH 8.0 and extracted as detailed above. Samples were analyzed by using an Agilent 1200 series high performance liquid chromatography (HPLC) equipped with a diode array detector. Samples were run on an Agilent SB-Aq 4.5 x 150 mm column with 5 µm pores at 1.5 ml/min at 40 °C with a gradient of 2 to 100% buffer B over 10 min, where buffer A was 10 mM ammonium acetate pH 6.5 and buffer B was 100% methanol. Quantification was achieved by comparing peak areas at 280 nm with authentic standards. For LCMS detection and quantification of benzimidazoles extracted from the environment, we used a recently reported method (Y. J. Men, E. C. Seth, S. Yi, T. S. Crofts, R. H. Allen, M. E. Taga, and L. Alvarez-Cohen, submitted). Mass spectrometry was carried out on an Agilent 6410 liquid chromatograph-triple quadrupole mass spectrometer with multiple reaction monitoring. Samples were loaded on a 1.8 µm pore size, 3.0 x 50 mm Agilent C18 Eclipse Plus column and eluted using a gradient of 0.1% formic acid in water against 0.1% formic acid in methanol. The following precursor and product ions were monitored during quantification of benzimidazoles: 119.1 and 65 for Bza, 133.1 and 77 for 5-MeBza, 147.1 and 131 for DMB, 149.2 and 79 for 5-OMeBza, 135.1 and 53 for 5-OHBza. The concentrations of each lower ligand were converted into "expected bioassay equivalents" using the following calculations. Each benzimidazole detectable by the LCMS method has been tested for its ability to rescue the *bluB* CFB phenotype (Figure 4.4). The relative EC$_{50}$ values of each benzimidazole can be compared to the EC$_{50}$ of DMB. The ratio of these two values provides a factor that can be used to convert the concentration of any benzimidazole into the concentration of DMB which would rescue the CFB phenotype to an equivalent extent. For example, based on their EC$_{50}$ values, 5-MeBza is approximately 10-fold less effective at rescuing the CFB phenotype than DMB. If LCMS analysis determines that a sample contains 10 nM 5-MeBza the expected bioassay equivalent would be 1 nM DMB, or simply 1 nM.

Chapter 5

Summary, conclusions, and future directions

5.1   Dissertation results in context

At the start of my graduate studies the CobT enzyme from *Salmonella enterica* was the most thoroughly characterized CobT homolog. *In vitro* and X-ray crystallographic experiments provided evidence that the *S. enterica* enzyme has low substrate specificity and will bind or activate a wide variety of lower ligand bases (Cheong et al., 2001; Trzebiatowski and Escalante-Semerena, 1997). Several bacteria have also been shown to be amenable to guided biosynthesis (Allen and Stabler, 2008; Mok and Taga, 2013; Perlman and Barrett, 1958). Taken together, these two lines of evidence led to the hypothesis that substrate promiscuity is a characteristic of the CobT enzyme family and that cobamide identity is determined by lower ligand availability.

In this dissertation I have provided evidence for a more nuanced hypothesis. In Chapter 2 and a companion paper that I co-authored (Hazra et al., 2013), I have provided evidence that under physiological conditions CobT substrate specificity is an important factor in determining which cobamides are produced. These findings support the new hypothesis that lower ligand availability is necessary for the production of diverse cobamides, but substrate specificity of CobT limits the cobamides that can be produced. In Chapter 3, I noted that *S. meliloti* and *V. parvula* each produced a pair of cobamides when grown with asymmetric lower ligand bases. Our hypothesis that these additional products represented structural isomers was verified by multiple analytical methods. These findings demonstrate that CobT substrate specificity applies to substrate orientation as well as identity. Finally, I demonstrated in Chapter 4 that benzimidazoles are prevalent in the environment and are accessible to organisms capable of lower ligand attachment.

The work in this dissertation can be summarized with the following statements: 1) CobT enzyme substrate specificity and lower ligand availability together determine cobamide identity, 2) The orientation of asymmetric lower ligands is further specified by CobT, and 3) Benzimidazoles may be plentiful in the environment and are available for incorporation into cobamides. This is represented graphically in Figure 5.1.

These statements support a model in which there are four general types of cobamide-producing prokaryotes (Table 5.1). The first group could be represented by *S. meliloti*. These prokaryotes both synthesize a lower ligand, such as DMB, and only incorporate a subset of lower ligand bases. For wild type *S. meliloti*, the CobU enzyme prevents both the incorporation of endogenous adenine and low concentrations of exogenous benzimidazoles (Chapter 2) (Hazra et al., 2013). The second group of organisms can be represented by *D. mccartyi*, which utilizes lower ligands from the environment but, through CobT, preferentially attaches only those it can use best (Chapter 2) (Yi et al., 2012). The third group consists of organisms that show little lower ligand attachment specificity but still use an endogenous lower ligand. Good representatives of this group are *S. ovata* and *V. parvula* which are capable of attaching a range of lower ligands but produce cobamides using endogenous phenol and *p*-cresol (Chapters 2, 3) (Mok and Taga,

2013). High external concentrations of benzimidazoles can outcompete the native lower ligands, but environmental concentrations do not appear to be high enough (Chapter 4). The final group is made up of cobamide-producing organisms that use an exogenous lower ligand but also show little specificity in attachment. *S. enterica* can be considered part of this group as it can utilize endogenous purines as lower ligands but it will preferentially use exogenous DMB (Anderson et al., 2008).
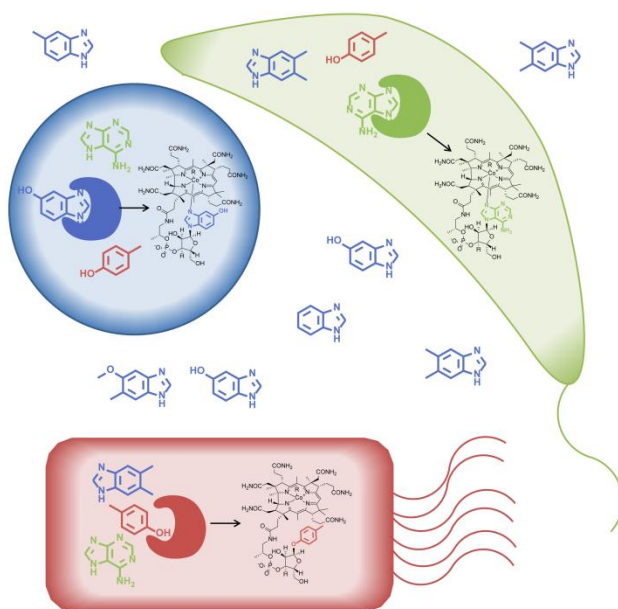

Figure 5.1 Availability of endogenous and exogenous lower ligands and CobT specificity dictate cobamide diversity. Cobamide producing organisms have access to exogenous lower ligands and, in some cases, endogenous lower ligands. Substrate concentration and/or CobT specificity determine what lower ligands are incorporated. In some organisms additional enzymes may play a role in governing specificity. Adapted from Crofts et al., 2013.

The level of promiscuity of an organism's cobamide-dependent enzyme(s) should determine what category a given organism belongs in. For example, the *S. meliloti* ribonucleotide reductase NrdJ does not utilize phenolyl or purinyl cobamides and *S. meliloti* CobU is incapable of activating phenolics or purines (Chapter 2). On the other hand, *S. enterica* is capable of utilizing a variety of different cobamides and it appropriately falls into the category of least control over lower ligand attachment (Chan and Escalante-Semerena, 2011; Trzebiatowski and Escalante-Semerena, 1997). The causality of this argument could, however, be reversed. Substrate promiscuity in lower ligand attachment may have driven the evolution of promiscuous cobamide-dependent enzymes. *S. ovata* appears to be an exception to this hypothesis, as it will incorporate a variety of benzimidazoles even though benzimidazolyl cobamides do not support its metabolism (Mok and Taga, 2013). I hypothesized in Chapter 2 that benzimidazoles may not be present in high concentrations in the environment of *S. ovata* and in Chapter 4 I supported this hypothesis by demonstrating that in many environments benzimidazoles are below inhibitory concentrations for *S. ovata*.

Table 5.1 Four types of cobamide-producing organisms

		Lower ligand attachment	
		High Specificity	Low Specificity
Source of lower ligand bases	Endogenous Bases	*S. meliloti*	*S. ovata*, *V. parvula*
	Exogenous Bases	*D. mccartyi*	*S. enterica*

5.2 Unanswered questions in cobamide research

Several new questions have been generated by the work described in the preceding chapters. I have shown that cobamide-producing organisms regulate what lower ligand bases they incorporate, but the evolutionary advantage of this is unknown. I provided some hypotheses for the existence of cobamide diversity in Chapter 1, among them better cofactor catalysis and as a means to prevent competing organisms from scavenging cobamides. Another issue of importance to the broader field is the apparent preference for cobalamin among prokaryotes. While cobalamin is the most important cobamide for humans, it is not necessarily the most important cobamide for life. Despite this, only two organisms, *S. ovata* and *S. multivorans* (see Chapter 1), have been identified that function less well with cobalamin than other cobamides (Keller et al., 2013; Mok and Taga, 2013; Stupperich et al., 1990). One possible reason so many organisms appear to favor cobalamin is that culturing methods have produced a bias. Recalcitrant organisms are often cultured in media supplemented with a vitamin solution, and invariably the cobamide provided has been cobalamin, the only commercially available cobamide. It is therefore worth considering the possibility that some bacteria thought to be unculturable may instead require a cobamide other than cobalamin for growth. This too is a line of thought that may be worth pursuing.

I have also raised a series of questions within more limited areas. First, in Chapter 2, it is not known why *Lactobacillus reuteri* was the only organism tested that would not incorporate benzimidazole lower ligands. When expressed in *S. meliloti*, the *L. reuteri cobT* allowed attachment of benzimidazoles and *in vitro* the *L. reuteri* CobT preferred DMB over any other lower ligand (Hazra et al., 2013). It is apparent then that lower ligand specificity in *L. reuteri* is not controlled simply by lower ligand availability or by CobT specificity. We have hypothesized that perhaps CobS, the enzyme that functions after CobT in the lower ligand attachment pathway, may be playing a role in specificity. Substrate specificity with respect to the lower ligand has not been characterized in CobS, making this a potentially fruitful direction for future studies.

The second question also concerns *L. reuteri*. In Chapter 3 the *L. reuteri* CobT was found to catalyze the formation of the unusual ribotide α-N9-adenine-ribotide-phosphate. However, we were unable to purify the corresponding N9-adeninylcobamide from either *L. reuteri* cultures or cultures of *S. meliloti* expressing the *L. reuteri cobT*. Some possibilities for the apparent mismatch between *in vitro* and *in vivo* experiments are explored in Chapter 3 but it is apparent that *L. reuteri* remains an interesting model cobamide-producing bacterium.

The third question concerns the presence of benzimidazoles in the environment and laboratory materials. One result from Chapter 4 was the discovery of benzimidazoles in commercial yeast extract. Yeast are not known to produce corrinoids and have no known corrinoid-dependent enzymes, making the origins of these benzimidazoles puzzling (Roth et al., 1996). Similarly, while the presence of DMB in laboratory agar has previously been hypothesized, the origin is again unknown (Anderson et al., 2008). This puzzle is of interest to us and we are planning on investigating it further. Benzimidazoles were also detected in an extract from a Winogradsky column. The microbial community in this sample is undefined, though both aerobic and anaerobic metabolisms likely are in place. The presence or absence of corrinoids in this community has not been established but may be of interest given the diversity of benzimidazoles present.

As an increasing number of important environmental communities such as the human gut or soil continue to be investigated, the role of cobamides besides cobalamin will likely become more evident. Understanding how diverse cobamides are formed is one of the first steps to understanding their functions.

References

Allen, R.H., and Stabler, S.P. (2008). Identification and quantitation of cobalamin and cobalamin analogues in human feces. Am. J. Clin. Nutr. *87*, 1324–1335.

Anderson, P.J., Lango, J., Carkeet, C., Britten, A., Kräutler, B., Hammock, B.D., and Roth, J.R. (2008). One pathway can incorporate either adenine or dimethylbenzimidazole as an alpha-axial ligand of B12 cofactors in *Salmonella enterica*. J. Bacteriol. *190*, 1160–1171.

Anonymous (1955). Biochemistry of Vitamin B12. Nature *175*, 875–877.

Banerjee, R., and Ragsdale, S.W. (2003). The many faces of vitamin B12: catalysis by cobalamin-dependent enzymes. Annu. Rev. Biochem. *72*, 209–247.

Barchielli, R., Boretti, G., Julita, P., Migliacci, A., and Minghetti, A. (1957). A natural guanine-containing analogue of vitamin B12. Biochim. Biophys. Acta *25*, 452.

Barker, H.A., Weissbach, H., and Smyth, R.D. (1958). A coenzyme containing pseudovitamin B(12). Proc. Natl. Acad. Sci. U. S. A. *44*, 1093–1097.

Barker, H.A., Smyth, R.D., Weissbach, H., Munch-Peterson, A., Toohey, J.I., LADD, J.N., Volcani, B.E., and Wilson, R.M. (1960a). Assay, purification, and properties of the adenylocobamide coenzyme. J. Biol. Chem. *235*, 181–190.

Barker, H.A., Smyth, R.D., Weissbach, H., Toohey, J.I., Ladd, J.N., and Volcani, B.E. (1960b). Isolation and properties of crystalline cobamide coenzymes containing benzimidazole or 5, 6-dimethylbenzimidazole. J. Biol. Chem. *235*, 480–488.

Battersby, A.R. (2000). Tetrapyrroles: the pigments of life. Nat. Prod. Rep. *17*, 507–526.

Biedendieck, R., Malten, M., Barg, H., Bunk, B., Martens, J.-H., Deery, E., Leech, H., Warren, M.J., and Jahn, D. (2010). Metabolic engineering of cobalamin (vitamin B12) production in *Bacillus megaterium*. Microb. Biotechnol. *3*, 24–37.

Blanche, F., Cameron, B., Crouzet, J., Debussche, L., Thibaut, D., Vuilhorgne, M., Leeper, F.J., and Battersby, A.R. (1995). Vitamin B12: How the Problem of Its Biosynthesis Was Solved. Angew. Chemie Int. Ed. English *34*, 383–411.

Brey, R.N., Banner, C.D., and Wolf, J.B. (1986). Cloning of multiple genes involved with cobalamin (Vitamin B12) biosynthesis in *Bacillus megaterium*. J. Bacteriol. *167*, 623–630.

Brown, K.L. (2005). Chemistry and enzymology of vitamin B12. Chem. Rev. *105*, 2075–2149.

Brown, F.B., Cain, J.C., Gant, D.E., Parker, L.F., and Smith, E.L. (1955). The vitamin B12 group; presence of 2-methyl purines in factors A and H and isolation of new factors. Biochem. J. *59*, 82–86.

Brushaber, K.R., O'Toole, G.A., and Escalante-Semerena, J.C. (1998). CobD, a novel enzyme with L-threonine-O-3-phosphate decarboxylase activity, is responsible for the synthesis of (R)-1-amino-2-propanol O-2-phosphate, a proposed new intermediate in cobalamin biosynthesis in *Salmonella typhimurium* LT2. J. Biol. Chem. *273*, 2684–2691.

Butler, P., Ebert, M.-O., Lyskowski, A., Gruber, K., Kratky, C., and Kräutler, B. (2006). Vitamin B12: a methyl group without a job? Angew. Chem. Int. Ed. Engl. *45*, 989–993.

Cameron, B., Blanche, F., Rouyez, M.C., Bisch, D., Famechon, A., Couder, M., Cauchois, L., Thibaut, D., Debussche, L., and 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. J. Bacteriol. *173*, 6066–6073.

Campbell, G.R.O., Taga, M.E., Mistry, K., Lloret, J., Anderson, P.J., Roth, J.R., and Walker, G.C. (2006). *Sinorhizobium meliloti bluB* is necessary for production of 5,6-dimethylbenzimidazole, the lower ligand of B12. Proc. Natl. Acad. Sci. U. S. A. *103*, 4634–4639.

Carlson, J., Fuchs, J. a, and Messing, J. (1984). Primary structure of the *Escherichia coli* ribonucleoside diphosphate reductase operon. Proc. Natl. Acad. Sci. U. S. A. *81*, 4294–4297.

Chan, C.H., and Escalante-Semerena, J.C. (2011). ArsAB, a novel enzyme from *Sporomusa ovata* activates phenolic bases for adenosylcobamide biosynthesis. Mol. Microbiol. *81*, 952–967.

Chan, C.H., Newmister, S. a, Talyor, K., Claas, K.R., Rayment, I., and Escalante-Semerena, J.C. (2013). Dissecting cobamide diversity through structural and functional analyses of the base-activating CobT enzyme of Salmonella enterica. Biochim. Biophys. Acta *1840*, 464–475.

Chen, W.P., and Kuo, T.T. (1993). A simple and rapid method for the preparation of gram-negative bacterial genomic DNA. Nucleic Acids Res. *21*, 2260.

Chen, C., Natale, D. a, Finn, R.D., Huang, H., Zhang, J., Wu, C.H., and Mazumder, R. (2011). Representative proteomes: a stable, scalable and unbiased proteome set for sequence analysis and functional annotation. PLoS One *6*, e18910.

Cheong, C.G., Escalante-Semerena, J.C., and Rayment, I. (1999). The three-dimensional structures of nicotinate mononucleotide:5,6- dimethylbenzimidazole phosphoribosyltransferase (CobT) from *Salmonella typhimurium* complexed with 5,6-dimethybenzimidazole and its reaction products determined to 1.9 Å resolution. Biochemistry *38*, 16125–16135.

Cheong, C.G., Escalante-Semerena, J.C., and Rayment, I. (2001). Structural investigation of the biosynthesis of alternative lower ligands for cobamides by nicotinate mononucleotide: 5,6-dimethylbenzimidazole phosphoribosyltransferase from *Salmonella enterica*. J. Biol. Chem. *276*, 37612–37620.

Cheong, C.G., Escalante-Semerena, J.C., and Rayment, I. (2002). Capture of a labile substrate by expulsion of water molecules from the active site of nicotinate mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase (CobT) from *Salmonella enterica*. J. Biol. Chem. *277*, 41120–41127.

Collins, H.F., Biedendieck, R., Leech, H.K., Gray, M., Escalante-Semerena, J.C., McLean, K.J., Munro, A.W., Rigby, S.E.J., Warren, M.J., and Lawrence, A.D. (2013). *Bacillus megaterium* has both a functional BluB protein required for DMB synthesis and a related flavoprotein that forms a stable radical species. PLoS One *8*, e55708.

Cowles, J.R., and Evans, H.J. (1968). Some properties of the ribonucleotide reductase from *Rhizobium meliloti*. Arch. Biochem. Biophys. *127*, 770–778.

Crofts, T.S., Seth, E.C., Hazra, A.B., and Taga, M.E. (2013). Cobamide structure depends on both lower ligand availability and CobT substrate specificity. Chem. Biol. *6*, 1265–1274.

Crowfoot-Hodgkin, D. (1965). Die Röntgenstrukturanalyse komplizierter Moleküle. Nobel-Vortrag am 12. Dezember 1964. Angew. Chemie *77*, 954–962.

Denger, K., and Schink, B. (1992). Energy conservation by succinate decarboxylation in *Veillonella parvula*. J. Gen. Microbiol. *138*, 967–971.

Dion, H., Calkins, D., and Pfiffner, J. (1954). 2-methyladenine, an hydrolysis product of pseudovitamin B12d. J. Am. Chem. Soc. *76*, 948–949.

Dion, H.W., Calkins, D.G., and Pfiffner, J.J. (1952). Hydrolysis products of pseudovitamin B12. J. Am. Chem. Soc. *74*, 1108–1108.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797.

Eisenreich, W., and Bacher, A. (1991). Biosynthesis of 5-hydroxybenzimidazolylcobamid (factor III) in *Methanobacterium thermoautotrophicum*. J. Biol. Chem. *266*, 23840–23849.

Elford, H.L. (1968). Effect of hydroxyurea on ribonucleotide reductase. Biochem. Biophys. Res. Commun. *33*, 129–135.

Fan, C., and Bobik, T. a (2008). The PduX enzyme of *Salmonella enterica* is an L-threonine kinase used for coenzyme B12 synthesis. J. Biol. Chem. *283*, 11322–11329.

Fan, C., Fromm, H.J., and Bobik, T. a (2009). Kinetic and functional analysis of L-threonine kinase, the PduX enzyme of *Salmonella enterica*. J. Biol. Chem. *284*, 20240–20248.

Fellay, R., Frey, J., and Krisch, H. (1987). Interposon mutagenesis of soil and water bacteria: a family of DNA fragments designed for in vitro insertional mutagenesis of gram-negative bacteria. Gene *52*, 147–154.

Finan, T.M., Hartweig, E., LeMieux, K., Bergman, K., Walker, G.C., and Signer, E.R. (1984). General transduction in *Rhizobium meliloti*. J. Bacteriol. *159*, 120–124.

Finan, T.M., Hirsch, a M., Leigh, J.A., Johansen, E., Kuldau, G. a, Deegan, S., Walker, G.C., and Signer, E.R. (1985). Symbiotic mutants of *Rhizobium meliloti* that uncouple plant from bacterial differentiation. Cell *40*, 869–877.

Ford, J.E., Holdsworth, E.S., and Kon, S.K. (1955). The biosynthesis of vitamin B12-like compounds. Biochem. J. *59*, 86–93.

Friedmann, H.C. (1965). Partial purification and properties of a single displacement trans-N-glycosidase. J. Biol. Chem. *240*, 413–418.

Friedmann, H.C., and Cagen, L.M. (1970). Microbial biosynthesis of B12-like compounds. Annu. Rev. Microbiol. *24*, 159–208.

Friedmann, H.C., and Fyfe, J.A. (1969). Pseudovitamin B 12 biosynthesis. Enzymatic formation of a new adenylic acid, 7-alpha-D-ribofuranosyladenine 5'-phosphate. J. Biol. Chem. *244*, 1667–1671.

Friedmann, H.C., and Harris, D.L. (1965). The formation of α-glycosidic 5'-nucleotides by a single displacement trans-N-glycosidase. J. Biol. Chem. *240*, 406–412.

Friedrich, W., and Bernhauer, K. (1957). Zur Chemie und Biochemie der „Cobalamine", VII. 2-Methylmercapto-Adenin-Cobalamin-Analogon, ein Neuer B12-Faktor des Faulschlammes. Chem. Ber. *90*, 1966–1973.

Friedrich, W., and Bernhauer, K. (1958). Zur Chemie und Biochemie der „Cobalamine", VIII. Über die 5- und 6-Methyl-benzimidazol-cobalamin-Analoga. Chem. Ber. *91*, 1665–1670.

Fyfe, J.A., and Friedmann, H.C. (1969). Vitamin B 12 biosynthesis. Enzyme studies on the formation of the alpha-glycosidic nucleotide precursor. J. Biol. Chem. *244*, 1659–1666.

Gibson, D.G., Young, L., Chuang, R., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.

Girard, C.L., Santschi, D.E., Stabler, S.P., and Allen, R.H. (2009). Apparent ruminal synthesis and intestinal disappearance of vitamin B12 and its analogs in dairy cows. J. Dairy Sci. *92*, 4524–4529.

Grabau, C., and Roth, J.R. (1992). A *Salmonella typhimurium* cobalamin-deficient mutant blocked in 1-amino-2-propanol synthesis. J. Bacteriol. *174*, 2138–2144.

Gray, M.J., and Escalante-Semerena, J.C. (2007). Single-enzyme conversion of FMNH2 to 5,6-dimethylbenzimidazole, the lower ligand of B12. Proc. Natl. Acad. Sci. U. S. A. *104*, 2921–2926.

Gray, M.J., and Escalante-Semerena, J.C. (2009a). The cobinamide amidohydrolase (cobyric acid-forming) CbiZ enzyme: a critical activity of the cobamide remodelling system of *Rhodobacter sphaeroides*. Mol. Microbiol. *74*, 1198–1210.

Gray, M.J., and Escalante-Semerena, J.C. (2009b). In vivo analysis of cobinamide salvaging in *Rhodobacter sphaeroides* strain 2.4.1. J. Bacteriol. *191*, 3842–3851.

Gray, M.J., and Escalante-Semerena, J.C. (2010). A new pathway for the synthesis of α-ribazole-phosphate in *Listeria innocua*. Mol. Microbiol. *77*, 1429–1438.

Gray, M.J., Tavares, N.K., and Escalante-Semerena, J.C. (2008). The genome of *Rhodobacter sphaeroides* strain 2.4.1 encodes functional cobinamide salvaging systems of archaeal and bacterial origins. Mol. Microbiol. *70*, 824–836.

Gronow, S., Welnitz, S., Lapidus, A., Nolan, M., Ivanova, N., Glavina Del Rio, T., Copeland, A., Chen, F., Tice, H., Pitluck, S., et al. (2010). Complete genome sequence of *Veillonella parvula* type strain (Te3). Stand. Genomic Sci. *2*, 57–65.

Gruber, K., Puffer, B., and Kräutler, B. (2011). Vitamin B12-derivatives-enzyme cofactors and ligands of proteins and nucleic acids. Chem. Soc. Rev. *40*, 4346–4363.

Guimarães, D.H., Weber, A., Klaiber, I., Vogler, B., and Renz, P. (1994). Guanylcobamide and hypoxanthylcobamide-Corrinoids formed by *Desulfovibrio vulgaris*. Arch. Microbiol. *162*, 272–276.

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. *41*, 95–98.

Harada, F., Gross, H.J., Kimura, F., Chang, S.H., Nishimura, S., and RajBhandary, U.L. (1968). 2-Methylthio N6-(Δ2-isopentenyl) adenosine: A component of tyrosine transfer RNA. Biochem. Biophys. Res. Commun. *33*, 299–306.

Hazra, A.B., Tran, J.L. a, Crofts, T.S., and Taga, M.E. (2013). Analysis of substrate specificity in CobT homologs reveals widespread preference for DMB, the lower axial ligand of Vitamin B12. Chem. Biol. *20*, 1275–1285.

He, J., Holmes, V.F., Lee, P.K.H., and Alvarez-Cohen, L. (2007). Influence of vitamin B12 and cocultures on the growth of *Dehalococcoides* isolates in defined medium. Appl. Environ. Microbiol. *73*, 2847–2853.

Hodgkin, D.C., Kamper, J., Mackay, M., Pickworth, J., Trueblood, K.N., and White, J.G. (1956). Structure of Vitamin B12. Nature *178*, 64–66.

Hoffmann, B., Oberhuber, M., Stupperich, E., Bothe, H., Buckel, W., Konrat, R., Kräutler, B., and Kra, B. (2000). Native corrinoids from *Clostridium cochlearium* are adeninylcobamides: spectroscopic analysis and identification of pseudovitamin B(12) and factor A. J. Bacteriol. *182*, 4773–4782.

Hörig, J.A., Renz, P., and Heckmann, G. (1978). [5-15N]Riboflavin as precursor in the biosynthesis of the 5,6-dimethylbenzimidazole moiety of vitamin B12. A study by 1H and 15N magnetic resonance spectroscopy. J. Biol. Chem. *253*, 7410–7414.

Irion, E., and Ljungdahl, L. (1965). Isolation of factor 3m coenzyme and cobyric acid coenzyme plus other B12 factors from *Clostridium thermoaceticum*. Biochemistry *4*, 2780–2790.

Kamikubo, T., and Matsuura, T. (1969). Incorporation of purine bases into Vitamin B12 molecule. Agric. Biol. Chem. *33*, 1207–1209.

Keck, B., and Renz, P. (2000). *Salmonella typhimurium* forms adenylcobamide and 2-methyladenylcobamide, but no detectable cobalamin during strictly anaerobic growth. Arch. Microbiol. *173*, 76–77.

Keck, B., Munder, M., and Renz, P. (1998). Biosynthesis of cobalamin in *Salmonella typhimurium*: transformation of riboflavin into the 5,6-dimethylbenzimidazole moiety. Arch. Microbiol. *171*, 66–68.

Keller, S., Ruetz, M., Kunze, C., Kräutler, B., Diekert, G., and Schubert, T. (2013). Exogenous 5,6-dimethylbenzimidazole caused production of a non-functional tetrachloroethene reductive dehalogenase in *Sulfurospirillum multivorans*. Environ. Microbiol. n/a–n/a.

Kohler, H.E. (1988). Isolation of cobamides from Methanothrix soehngenii: 5-methylbenzimidazole as the α-ligand of the predominant cobamide. Arch. Microbiol. *150*, 219–223.

Krasna, A.I., Rosenblum, C., and Sprinson, D.B. (1957). The conversion of L-threonine to the Dg-1-amino-2-propanol of vitamin B12. J. Biol. Chem. *225*, 745–750.

Kräutler, B., Moll, J., and Thauer, R.K. (1987). The corrinoid from *Methanobacterium thermoautotrophicum* (Marburg strain). Spectroscopic structure analysis and identification as Co beta-cyano-5'-hydroxybenzimidazolyl-cobamide (factor III). Eur. J. Biochem. *162*, 275–278.

Kräutler, B., Kohler, H.P., and Stupperich, E. (1988). 5'-Methylbenzimidazolyl-cobamides are the corrinoids from some sulfate-reducing and sulfur-metabolizing bacteria. Eur. J. Biochem. *176*, 461–469.

Kräutler, B., Fieber, W., Ostermann, S., Fasching, M., Ongania, K., Gruber, K., Kratky, C., Mikl, C., Siebert, A., and Diekert, G. (2003). The cofactor of tetrachloroethene reductive dehalogenase of *Dehalospirillum multivorans* is Norpseudo-B12, a new type of a natural corrinoid. Helv. Chim. Acta *86*, 3698–3716.

Kräutler, B. (1987). Thermodynamictrans-effects of the nucleotide base in the B12 coenzymes. Helv. Chim. Acta *70*, 1268–1278.

Ladd, J.N., Hogenkamp, H.P., and Barker, H.A. (1961). Structure of cobamide coenzymes: influence of pH on absorption spectra and ionophoretic mobilities. J. Biol. Chem. *236*, 2114–2118.

Lamm, L., Hörig, J.A., Renz, P., and Heckmann, G. (1980). Biosynthesis of vitamin B12. Experiments with the anaerobe *Eubacterium limosum* and some labelled substrates. Eur. J. Biochem. *109*, 115–118.

Lamm, L., Heckmann, G., and Renz, P. (1982). Biosynthesis of vitamin B12 in anaerobic bacteria. Mode of incorporation of glycine into the 5,6-dimethylbenzimidazole moiety in *Eubacterium limosum*. Eur. J. Biochem. *122*, 569–571.

Lawrence, C.C., and Stubbe, J. (1998). The function of adenosylcobalamin in the mechanism of ribonucleoside triphosphate reductase from *Lactobacillus leichmannii*. Curr. Opin. Chem. Biol. *2*, 650–655.

Leigh, J.A., Signer, E.R., and Walker, G.C. (1985). Exopolysaccharide-deficient mutants of *Rhizobium meliloti* that form ineffective nodules. Proc. Natl. Acad. Sci. U. S. A. *82*, 6231–6235.

Lezius, A.G., and Barker, H.A. (1965). Corrinoid compounds of *Methanobacillus omelianskii*. I. Fractionation of the corrinoid compounds and identification of factor 3 and factor 3 coenzyme. Biochemistry *4*, 510–518.

Littlefield, J.W., and Dunn, D.B. (1958). The occurrence and distribution of thymine and three methylated-adenine bases in ribonucleic acids from several sources. Biochem. J. *70*, 642–651.

Ljungdahl, L., Irion, E., and Wood, H.G. (1965). Total synthesis of acetate from CO2. I. Co-methylcobyric acid and CO-(methyl)-5-methoxybenzimidazolylcobamide as intermediates with *Clostridium thermoaceticum*. Biochemistry *4*, 2771–2780.

Lopes, J.N., and Cruz, F.S. (1976). Chemically defined media for growing anaerobic bacteria of the genus *Veillonella*. Antonie Van Leeuwenhoek *42*, 411–420.

Maggio-Hall, L.A., and Escalante-Semerena, J.C. (1999). In vitro synthesis of the nucleotide loop of cobalamin by *Salmonella typhimurium* enzymes. Proc. Natl. Acad. Sci. U. S. A. *96*, 11798–11803.

Maggio-Hall, L.A., Dorrestein, P.C., Escalante-Semerena, J.C., and Begley, T.P. (2003). Formation of the dimethylbenzimidazole ligand of coenzyme B(12) under physiological conditions by a facile oxidative cascade. Org. Lett. *5*, 2211–2213.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). Molecular cloning : a laboratory manual (Plainview, NY: Cold Spring Harbor Lab. Press).

Martens, J.H., Barg, H., Warren, M.J., and Jahn, D. (2002). Microbial production of vitamin B12. Appl. Microbiol. Biotechnol. *58*, 275–285.

Miller, M. a., Pfeiffer, W., and Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In 2010 Gateway Computing Environments Workshop (GCE), (IEEE), pp. 1–8.

Minot, G.R., and Murphy, W.P. (1926). Treatment of pernicious anemia by a special diet. JAMA J. Am. Med. Assoc. *87*, 470–476.

Mok, K.C., and Taga, M.E. (2013). Growth inhibition of *Sporomusa ovata* by incorporation of benzimidazole bases into cobamides. J. Bacteriol. *195*, 1902–1911.

Moore, S.J., Lawrence, A.D., Biedendieck, R., Deery, E., Frank, S., Howard, M.J., Rigby, S.E.J., and Warren, M.J. (2013). Elucidation of the anaerobic pathway for the corrin component of cobalamin (vitamin B12). Proc. Natl. Acad. Sci. U. S. A. 6–11.

Munder, M., Vogt, J.R., Vogler, B., and Renz, P. (1992). Biosynthesis of vitamin B12 in anaerobic bacteria. Experiments with *Eubacterium limosum* on the incorporation of D-[1-13C]erythrose and [13C]formate into the 5,6-dimethylbenzimidazole moiety. Eur. J. Biochem. *204*, 679–683.

Newmister, S.A., Chan, C.H., Escalante-Semerena, J.C., and Rayment, I. (2012). Structural Insights into the Function of the Nicotinate Mononucleotide:phenol/p-cresol Phosphoribosyltransferase (ArsAB) Enzyme from *Sporomusa ovata*. Biochemistry *51*, 8571–8582.

Ng, S.K., and Hamilton, I.R. (1971). Lactate metabolism by *Veillonella parvula*. J. Bacteriol. *105*, 999–1005.

Odsbu, I., Morigen, and Skarstad, K. (2009). A reduction in ribonucleotide reductase activity slows down the chromosome replication fork but does not change its localization. PLoS One *4*, e7617.

Parks, J.M., Johs, A., Podar, M., Bridou, R., Hurt, R. a, Smith, S.D., Tomanicek, S.J., Qian, Y., Brown, S.D., Brandt, C.C., et al. (2013). The genetic basis for bacterial mercury methylation. Science *339*, 1332–1335.

Perlman, D. (1959). Microbial synthesis of cobamides. Adv. Appl. Microbiol. *1*, 87–122.

Perlman, D. (1971). Processes for the biosynthesis of cobamides. Methods Enzymol. *934*, 75–82.

Perlman, D., and Barrett, J. (1958). Biosynthesis of cobalamin analogues by *Propionibacterium arabinosum*. Can. J. Microbiol. *4*, 9–15.

Pol, A., van der Drift, C., and Vogels, G.D. (1982). Corrinoids from *Methanosarcina barkeri*: Structure of the α-ligand. Biochem. Biophys. Res. Commun. *108*, 731–737.

Pommier, Y., Leo, E., Zhang, H., and Marchand, C. (2010). DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. Chem. Biol. *17*, 421–433.

Puigbò, P., Guzmán, E., Romeu, A., and Garcia-Vallvé, S. (2007). OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. Nucleic Acids Res. *35*, W126–31.

Puigbò, P., Romeu, A., and Garcia-Vallvé, S. (2008). HEG-DB: a database of predicted highly expressed genes in prokaryotic complete genomes under translational selection. Nucleic Acids Res. *36*, D524–7.

Ragsdale, S.W., and Pierce, E. (2008). Acetogenesis and the Wood-Ljungdahl pathway of CO(2) fixation. Biochim. Biophys. Acta *1784*, 1873–1898.

Renz, P. (1999). Biosynthesis of the 5,6-dimethylbenzimidazole moiety of cobalamin and of the other bases found in natural corrinoids. In Chemistry and Biochemistry of B12, R. Banerjee, ed. (New York: John Wiley and Sons, Inc), pp. 557–575.

Renz, P., Blickle, S., and Friedrich, W. (1987). Two new vitamin B-12 factors from sewage sludge containing 2-methylsulfinyladenine or 2-methylsulfonyladenine as base component. Eur. J. Biochem. *163*, 175–179.

Renz, P., Endres, B., Kurz, B., and Marquart, J. (1993). Biosynthesis of vitamin B12 in anaerobic bacteria. Transformation of 5-hydroxybenzimidazole and 5-hydroxy-6-methylbenzimidazole into 5,6-dimethylbenzimidazole in *Eubacterium limosum*. Eur. J. Biochem. *217*, 1117–1121.

Rickes, E.L., Brink, N.G., Koniuszy, F.R., Wood, T.R., and Folkers, K. (1948a). Crystalline Vitamin B12. Science *107*, 396–397.

Rickes, E.L., Brink, N.G., Koniuszy, F.R., Wood, T.R., and Folkers, K. (1948b). Comparative data on Vitamin B12 from liver and from a new source, *Streptomyces griseus*. Science *108*, 634–635.

Rinaudi, L., Fujishige, N. a, Hirsch, A.M., Banchio, E., Zorreguieta, A., and Giordano, W. (2006). Effects of nutritional and environmental conditions on *Sinorhizobium meliloti* biofilm formation. Res. Microbiol. *157*, 867–875.

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

Roessner, C. a, Huang, K.-X., Warren, M.J., Raux, E., and Scott, a I. (2002). Isolation and characterization of 14 additional genes specifying the anaerobic biosynthesis of cobalamin (vitamin B12) in *Propionibacterium freudenreichii* (*P. shermanii*). Microbiology *148*, 1845–1853.

Rossi, M., Glusker, J.P., Randaccio, L., Summers, M.F., Toscano, P.J., and Marzilli, L.G. (1985). The structure of a B12 coenzyme: methylcobalamin studies by x-ray and NMR methods. J. Am. Chem. Soc. *107*, 1729–1738.

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

Roth, J.R., Lawrence, J.G., and Bobik, T.A. (1996). Cobalamin (coenzyme B12): synthesis and biological significance. Annu. Rev. Microbiol. *50*, 137–181.

Saneyoshi, M., Oashi, Z., Harada, F., and Nishimura, S. (1972). Isolation and characterization of 2-methyladenosine from *Escherichia coli* tRNA Glu 2 , tRNA Asp 1 , tRNA His 1 and tRNA Arg . Biochim. Biophys. Acta *262*, 1–10.

Santos, F., Vera, J.L., Lamosa, P., de Valdez, G.F., de Vos, W.M., Santos, H., Sesma, F., and Hugenholtz, J. (2007). Pseudovitamin B(12) is the corrinoid produced by *Lactobacillus reuteri* CRL1098 under anaerobic conditions. FEBS Lett. *581*, 4865–4870.

Santos, F., Teusink, B., Molenaar, D., van Heck, M., Wels, M., Sieuwerts, S., de Vos, W.M., and Hugenholtz, J. (2009). Effect of amino acid availability on vitamin B12 production in *Lactobacillus reuteri*. Appl. Environ. Microbiol. *75*, 3930–3936.

Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., and Pühler, A. (1994). Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and

pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene *145*, 69–73.

Scherer, P., Höllriegel, V., Krug, C., Bokel, M., and Renz, P. (1984). On the biosynthesis of 5-hydroxybenzimidazolylcobamide (vitamin B12-factor III) in *Methanosarcina barkeri*. Arch. Microbiol. *138*, 354–359.

Seeliger, S., Janssen, P.H., and Schink, B. (2002). Energetics and kinetics of lactate fermentation to acetate and propionate via methylmalonyl-CoA or acrylyl-CoA. FEMS Microbiol. Lett. *211*, 65–70.

Seshadri, R., Adrian, L., Fouts, D.E., Eisen, J. a, Phillippy, A.M., Methe, B.A., Ward, N.L., Nelson, W.C., Deboy, R.T., Khouri, H.M., et al. (2005). Genome sequence of the PCE-dechlorinating bacterium *Dehalococcoides ethenogenes*. Science *307*, 105–108.

Smith, E.L., and Parker, L.F.J. (1948). Purification of anti-pernicious anaemia factor. Biochem. J. *43*, viii.

Soto, M.J., Sanjuán, J., and Olivares, J. (2006). Rhizobia and plant-pathogenic bacteria: common infection weapons. Microbiology *152*, 3167–3174.

Stabler, P. (1999). B12 and nutrition. In Chemistry and Biochemistry of B12, R. Banerjee, ed. (New York: John Wiley and Sons, Inc), pp. 343–365.

Stadtman, T.C. (1960). Synthesis of adenine-B12 coenzyme by *Clostridium sticklandii*: relationship to one-carbon metabolism. J. Bacteriol. *79*, 904–905.

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics *22*, 2688–2690.

Stamatakis, A., Hoover, P., and Rougemont, J. (2008). A rapid bootstrap algorithm for the RAxML Web servers. Syst. Biol. *57*, 758–771.

Stubbe, J. (1994). Binding site revealed of nature's most beautiful cofactor. Science *266*, 1663–1664.

Stupperich, E., and Eisinger, H.J. (1989). Biosynthesis of para-cresolyl cobamide in *poromusa ovata*. Arch. Microbiol. *151*, 372–377.

Stupperich, E., and Kräutler, B. (1988). Pseudo vitamin B12 or 5-hydroxybenzimidazolyl-cobamide are the corrinoids found in methanogenic bacteria. Arch. Microbiol. *149*, 268–271.

Stupperich, E., Steiner, I., and Eisinger, H.J. (1987). Substitution of Co alpha-(5-hydroxybenzimidazolyl)cobamide (factor III) by vitamin B12 in *Methanobacterium thermoautotrophicum*. J. Bacteriol. *169*, 3076–3081.

Stupperich, E., Eisinger, H.J., and Kräutler, B. (1988). Diversity of corrinoids in acetogenic bacteria. P-cresolylcobamide from *Sporomusa ovata*, 5-methoxy-6-methylbenzimidazolylcobamide from *Clostridium formicoaceticum* and vitamin B12 from *Acetobacterium woodii*. Eur. J. Biochem. *172*, 459–464.

Stupperich, E., Eisinger, H.J., and Kräutler, B. (1989). Identification of phenolyl cobamide from the homoacetogenic bacterium *Sporomusa ovata*. Eur. J. Biochem. *186*, 657–661.

Stupperich, E., Eisinger, H.J., and Schurr, S. (1990). Corrinoids in anaerobic bacteria. FEMS Microbiol. Rev. *87*, 355–359.

Stupperich, E., Eisinger, H.J., Kerssebaum, R., and Nexø, E. (1993). Fluorinated vitamin b(12) analogs are cofactors of corrinoid-dependent enzymes: a f-labeled nuclear magnetic resonance probe for identifying corrinoid-protein interactions. Appl. Environ. Microbiol. *59*, 599–603.

Summers, M.F., Marzilli, L.G., and Bax, A. (1986). Complete proton and carbon-13 assignments of coenzyme B12 through the use of new two-dimensional NMR experiments. J. Am. Chem. Soc. *108*, 4285–4294.

Taga, M.E., and Walker, G.C. (2010). *Sinorhizobium meliloti* requires a cobalamin-dependent ribonucleotide reductase for symbiosis with its plant host. Mol. Plant. Microbe. Interact. *23*, 1643–1654.

Taga, M.E., Larsen, N. a, Howard-Jones, A.R., Walsh, C.T., and Walker, G.C. (2007). BluB cannibalizes flavin to form the lower ligand of vitamin B12. Nature *446*, 449–453.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. *28*, 2731–2739.

Tanioka, Y., Yabuta, Y., Yamaji, R., Shigeoka, S., Nakano, Y., Watanabe, F., and Inui, H. (2009). Occurrence of pseudovitamin B12 and its possible function as the cofactor of cobalamin-dependent methionine synthase in a cyanobacterium *Synechocystis* sp. PCC6803. J. Nutr. Sci. Vitaminol. (Tokyo). *55*, 518–521.

Tollinger, M., Konrat, R., and Kräutler, B. (1999). The Structure of Methylcob(III)alamin in Aqueous Solution - A Water Molecule as Structuring Element of the Nucleotide Loop. Helv. Chim. Acta *82*, 1596–1609.

Trzebiatowski, J.R., and Escalante-Semerena, J.C. (1997). Purification and characterization of CobT, the nicotinate-mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase enzyme from *Salmonella typhimurium* LT2. J. Biol. Chem. *272*, 17662–17667.

Vogt, J.R., and Renz, P. (1988). Biosynthesis of vitamin B-12 in anaerobic bacteria. Experiments with *Eubacterium limosum* on the origin of the amide groups of the corrin ring and of N-3 of the 5,6-dimethylbenzimidazole part. Eur. J. Biochem. *171*, 655–659.

Wakayama, E.J., Dillwith, J.W., Howard, R.W., and Blomquist, G.J. (1984). Vitamin B12 levels in selected insects. Insect Biochem. *14*, 175–179.

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

Weiner, M.P., Costa, G.L., Schoettlin, W., Cline, J., Mathur, E., and Bauer, J.C. (1994). Site-directed mutagenesis of double-stranded DNA by the polymerase chain reaction. Gene *151*, 119–123.

Weissbach, H., Ladd, J.N., Volcani, B.E., Smyth, R.D., and Barker, H.A. (1960). Structure of the adenylcobamide coenzyme: degradation by cyanide, acid, and light. J. Biol. Chem. *235*, 1462–1473.

Wolf, J.B., and Brey, R.N. (1986). Isolation and genetic characterizations of *Bacillus megaterium* cobalamin biosynthesis-deficient mutants. J. Bacteriol. *166*, 51–58.

Woodson, J.D., and Escalante-Semerena, J.C. (2004). CbiZ, an amidohydrolase enzyme required for salvaging the coenzyme B12 precursor cobinamide in archaea. Proc. Natl. Acad. Sci. U. S. A. *101*, 3591–3596.

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

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

Wurm, R., Weyhenmeyer, R., and Renz, P. (1975). On the biosynthesis of 5-methoxybenzimidazole. Precursor-function of 5-hydroxybenzimidazole, benzimidazole and riboflavin. Eur. J. Biochem. *56*, 427–432.

Wurm, R., Renz, P., and Heckmann, G. (1980). The biosynthesis of the vitamin B12-analog 5-methoxybenzimidazolecobamide in *Clostridium thermoaceticum*. FEMS Microbiol. Lett. *7*, 11–13.

Yan, J., Ritalahti, K.M., Wagner, D.D., and Löffler, F.E. (2012). Unexpected specificity of interspecies cobamide transfer from *Geobacter* spp. to organohalide-respiring *Dehalococcoides mccartyi* strains. Appl. Environ. Microbiol. *78*, 6630–6636.

Yan, J., Im, J., Yang, Y., and Löffler, F.E. (2013). Guided cobalamin biosynthesis supports *Dehalococcoides mccartyi* reductive dechlorination activity. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *368*, 20120320.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene *33*, 103–119.

Yi, S., Seth, E.C., Men, Y., Stabler, S.P., Allen, R.H., Alvarez-Cohen, L., and Taga, M.E. (2012). Versatility in Corrinoid Salvaging and Remodeling Pathways Supports Corrinoid-Dependent Metabolism in *Dehalococcoides mccartyi*. Appl. Environ. Microbiol. *78*, 7745–7752.

Yu, T.-Y., Mok, K.C., Kennedy, K.J., Valton, J., Anderson, K.S., Walker, G.C., and Taga, M.E. (2012). Active site residues critical for flavin binding and 5,6-dimethylbenzimidazole biosynthesis in the flavin destructase enzyme BluB. Protein Sci. *21*, 839–849.

Zayas, C.L., and 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. J. Bacteriol. *189*, 2210–2218.

Zayas, C.L., Claas, K., and Escalante-Semerena, J.C. (2007). The CbiB protein of *Salmonella enterica* is an integral membrane protein involved in the last step of the de novo corrin ring biosynthetic pathway. J. Bacteriol. *189*, 7697–7708.

Zhang, Y., Rodionov, D. a, Gelfand, M.S., and Gladyshev, V.N. (2009). Comparative genomic analyses of nickel, cobalt and vitamin B12 utilization. BMC Genomics *10*, 78.