UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Uncovering how bacteria sense and respond to chemically diverse corrinoids through cobalamin riboswitch gene regulation

Permalink
https://escholarship.org/uc/item/17j9j40q

Author
Kennedy, Kristopher John

Publication Date
2021

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Uncovering how bacteria sense and respond to chemically diverse corrinoids through cobalamin riboswitch gene regulation

By

Kristopher J. Kennedy

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 Arash Komeili
Professor Kathleen R. Ryan
Professor Michael A. Marletta

Fall 2021

Abstract

Uncovering how bacteria sense and respond to chemically diverse corrinoids through cobalamin riboswitch gene regulation

by

Kristopher J. Kennedy

Doctor of Philosophy in Microbiology

University of California, Berkeley

Professor Michiko E. Taga, Chair

Cells actively sense and integrate information about their internal and external environments to execute adaptive physiological responses. This makes it possible to survive fluctuating and oftentimes harsh environmental conditions. Controlling the expression of genes is part of this vital cellular process. Indeed, since the birth of molecular biology theory, a multitude of cellular mechanisms have been found to control virtually every aspect of the flow of genetic information from DNA to RNA to functional proteins. The revelation that RNAs are more than just passive messengers between DNA and proteins was a turning point in the field of molecular biology. We now recognize that many types of RNA actively fold into complex three-dimensional structures to carry out important cellular functions with sophistication, precision, and efficiency that rivals proteins. A prime example are riboswitches – structured noncoding RNAs that sense specific small molecule effectors by direct binding and function as *cis*-regulatory genetic switches. My dissertation research focuses on the forms and functions of the cobalamin riboswitches and how they are impacted by the exceptional chemical diversity of naturally occurring variants of cobalamin known as corrinoids.

I begin the first chapter with a broad overview of bacterial gene regulation with specific emphasis on mechanisms of transcriptional and translational control. This sets the stage for delving into the distinctive elements of riboswitch structure, mechanism, and function. I also describe how various riboswitch classes connect to the cellular processes which they control. Lastly, I describe in detail the cobalamin riboswitch class and the roles that corrinoids play in bacterial physiology.

The second chapter describes the bulk of my endeavors as a graduate student researcher in the Taga Lab examining the corrinoid specificity of cobalamin riboswitches. The large number of bacterial metabolic pathways that involve corrinoids and the numerous types of corrinoid cofactors and intermediates present a puzzle as to how cells can effectively use cobalamin riboswitches to control their corrinoid-related physiology. The approach I took leveraged two strengths of the Taga Lab's expertise: bacterial molecular genetics and biochemical production of commercially unavailable corrinoid molecules. I engineered an *in vivo* fluorescence reporter system to measure the responses of several cobalamin riboswitches to several corrinoids. From the patterns of corrinoid selectivity that I observed in my experiments, I developed a mechanistic hypothesis for

corrinoid specificity of cobalamin riboswitch-based gene regulation in bacteria. Furthermore, I propose a regulatory strategy that attempts to explain how corrinoid specificities of gene regulation and bacterial physiology are functionally connected.

In the third chapter, I explore the functional versatility of cobalamin riboswitches. As the second most prevalent class of riboswitches, I speculated that novel functional variations should arise from the diversity of cobalamin riboswitch sequences. To increase the likelihood of finding uncommon functional variants, I focused on atypical cobalamin riboswitch regulon architectures among bacterial species that specialize in corrinoid metabolism. This rationale enabled me to successfully identify a novel activator cobalamin riboswitch, dissect modular functionalities of tandemly linked cobalamin riboswitches, and develop new hypotheses about corrinoid-specific physiology.

Together, my research studies constitute a step towards reconciling the apparent oversimplicity of current cobalamin riboswitch models with the intrinsic complexity of the cellular processes they control.

Dedicated to my son Vladimir,

for reminding me each day of life's wonderful strangeness

Table of Contents

Acknowledgments

I thank my advisor Michi Taga for many years of mentorship and for creating ideal environments for me to develop as a scientist. I first met Michi in 2007 when I worked as an undergraduate researcher for one summer at MIT in Boston. She was a postdoctoral researcher at the time, and she introduced me to bacterial genetics while I assisted her with one of her side projects. Now looking back, I am certain that was the most formative experience of my scientific career. Yet somehow, at the end of that summer I expressed to Michi something along the lines of, 'You know, doing these genetic experiments is fun and all, but I still don't see what the big deal is about bacteria and vitamins.' I'm glad she didn't recall that conversation eight years later when I asked to join her lab. In addition to being an excellent scientist and teacher, Michi is the most patient and thoughtful supervisor I have encountered. I'll try my best to emulate her as I advance in my career.

I worked with many kind and wonderful people in the Taga Lab. Olga Sokolovskaya and Amanda Shelton were the senior graduate students when I joined the lab. Most of my knowledge about corrinoid biology I gained from conversations with them. Olga kindly gifted me several precious adenosylated corrinoids for my research. Later on, she also kindly gifted me a precious baby boy☺. I joined the Taga Lab with my classmate Gordon Pherribo and we've been close friends throughout grad school. Alexa Nicolas also became a close friend, particularly after she broke her leg and the Taga and Firestone Labs joined forces to bring her food while she recovered. Kenny Mok helped me in countless ways with my research projects. More importantly though, he has been my main source of unintentional workplace humor. Becky Procknow and Zoila Alvarez-Aponte are the newest batch of grad students in the group. They bring a lot of joy and enthusiasm to the lab. I am excited to read their future papers on riboswitches and soil bacteria. I worked directly with Becky and an undergraduate researcher Victoria Innocent on riboswitch projects. Working with them strengthened certain aspects of my own dissertation research, so I thank them for that experience. I also had many enlightening discussions with the postdoctoral researchers in the lab: Amrita Hazra, Florian Widner, Nicole Abreu, Sebastian Gude, and Zach Hallberg. Their expertise and perspectives on chemistry, physics, and biology had many important impacts on my work. The Taga Lab is a special place and I'll be sad when Michi finally kicks me out.

I thank my dissertation committee members Arash Komeili, Kathleen Ryan, Ming Hammond, and Michael Marletta, who provided helpful feedback at yearly meetings and other departmental events. I'm also grateful for the teaching experiences I had while serving as an instructor for Pat Zambryski's Secret Life of Plants course and Arash's undergraduate microbiology laboratory course. I also thank Rocio Sanchez, Lyn Rivera, Joanna Straley, and Dana Jantz for keeping the department and graduate program running smoothly. Rocio was particularly kind and encouraging as the graduate student affairs advisor.

One of the most rewarding parts of graduate school has been making new friends. I especially want to thank Gordon Pherribo, Alexa Nicolas, Patrick West, Kyle Meagher, Cassidy Crawford, Kirby Walls, Edi Wipf, Brianna Haining, Dylan McClung, Emma Kovak, Dax Ovid, and Liz Kellogg for their friendship over the years.

I also want to acknowledge a few of my past research mentors. Carol Pandey, my psychology professor at Los Angeles Pierce College, first introduced me to the idea of doing laboratory research. When my application for a summer research program was rejected because of poor grades, I thought that was the end of it. Luckily, she persuaded me to reapply and helped me find opportunities to strengthen my application for the following year (which was successful). Without her early guidance, my career in science might have ended before it even began.

As a transfer student at UCLA, I had several people looking out for me. Most important was Professor Elma Gonzalez who ran the Minority Access to Research Careers scholarship program. I always struggled with school but having Elma and the MARC program gave me a homebase at a huge campus and provided me with opportunities to do research and stay motivated.

In 2009, I enrolled in doctoral studies at Cornell University. Although I didn't complete my degree, I still learned the importance of scientific rigor from my advisor Ruth Collins. I also made great friends through the Collins Lab, especially Fabio and Vera Rinaldi. They helped me through some challenging moments in my life and I am very lucky to call them friends. I am also indebted to Maria Harrison for giving me the chance to return to scientific research after taking a break.

I thank Liam Holt for offering me a position in his lab at UC Berkeley when I moved back to California. Liam taught me how to get excited about unanticipated experimental results, which was a strange concept to me at the time. I also thank Will Ludington for providing lots of support when I applied for graduate programs once again.

I also want to thank all my teachers and classmates at Moon Dragon Dojo in Berkeley. Training in martial arts with Sensei Allyson Appen was an essential counterbalance to my focus on scientific research. Sensei Rosanne Boudreau and Sensei Laura Pirott are now two of my closest friends and supporters.

Finally, I thank all of my immediate and extended family who have cheered me on throughout my life. My parents Ceci and Patrick always loved and supported me, no matter how difficult I make it for them at times. My sisters Jessica and Kim are always available for a few laughs and a few words of encouragement over the phone when I've needed it most. Lastly, I thank my newest family members, Olga and our son Vladimir. I look forward to future adventures and scientific endeavors with them by my side.

Chapter 1 – An introduction to bacterial gene regulation with an emphasis on riboswitches

1.1 A brief overview of bacterial gene regulation

1.1.1 The central dogma of molecular biology

In the study of biology, we refine, reframe, and ultimately return to a prominent description of how genetic information flows within cells: the central dogma of molecular biology (CDMB) [1]. Today, the CDMB is a collective understanding of the fundamental way that cells carry out the molecular actions of life. The model is understood to mean that genetic information flows directionally from DNAs that are transcribed to RNAs that are translated to proteins. While there are important cases in which the information flows in different directions, the CDMB describes the rudiments. Built upon and intertwined with the CDMB are a multitude of ways in which cells control the flow of genetic information. These strategies are collectively referred to as 'gene regulation'. Simply put, gene regulation is the essential way that cells have evolved to express the right genes, at the right time, and in the right place. In this introductory chapter of my dissertation, I provide a broad overview of gene regulation and the mechanisms by which they occur in bacteria. Due to the enormous amount of knowledge that has accumulated on this topic, this overview does not aim to be complete. Rather, it should serve to highlight several of the major developments in our understanding of gene regulation and provide a backdrop for my particular research interests and approaches to studying the functions of a special bacterial gene regulatory element – the cobalamin riboswitch.

The history of microbial gene regulation began with the formative bacterial physiological studies of Jacob and Monod, and the ensuing molecular biological work of Gilbert and Müller-Hill [2-5]. Together, their insights provided the initial descriptions of the molecular basis by which cells sense a specific environmental factor (lactose) and respond in a precise physiological manner (induction of lactose catabolism) – or in other words, how cells express the right genes at the right time and place. Subsequent decades of research in microbial physiology, genetics, molecular biology, and other fields have revealed a vast range of sophisticated gene regulatory schemes across all domains of life [6]. The flow of genetic information in the CDMB occurs by two basic cellular processes: transcription of a DNA template to a messenger RNA (mRNA) by the RNA polymerase (RNAP) complex, and translation of an RNA to a protein by the ribosome complex. At every step of these processes, there exist regulatory mechanisms that act in concert with the CDMB machinery. It is useful to think of these mechanisms in terms of the actions of DNA, RNA, protein, and other molecular factors in the cell. In this brief overview, I will focus primarily on the RNA- and protein-based regulatory mechanisms that act at each stage of gene expression.

1.1.2 Regulation of transcriptional initiation

Efficient transcriptional initiation of a bacterial operon requires the assembly of an RNAP holoenzyme complex at a promoter site. The RNAP catalytic core complex is composed of two small α subunits and two subunits (known as β and β'), which together are sufficient to transcribe RNA from a template DNA sequence. The complete RNAP holoenzyme contains the catalytic core and an additional bound accessory protein called a sigma factor, which is required for efficient initiation of transcription. The most basic level of gene regulation occurs by controlling how

frequently transcription of an operon initiates. The binding affinity between a promoter DNA sequence and the RNAP components is a major determinant of transcription initiation frequency. Furthermore, in many bacteria there are different types of sigma factors within the cell that form complexes with the RNAP catalytic core. The differences in protein structure between alternative sigma factors allow them to recognize different promoter sequences. In general, a bacterial genome encodes several different sigma factors, but relies on one primary 'housekeeping' sigma factor to carry out transcription of most operons. The less frequently used sigma factors are called 'alternative sigma factors'. Alternative sigma factors will typically bind to a small subset of promoters with specific sequences. The subset of operons that are regulated by a given alternative sigma factor are often functionally related. For example, the *E. coli* genome encodes seven sigma factors, with σ^D being its main sigma factor for general transcription. The alternative sigma factor σ^N of *E. coli* is specifically active and recruited to promoters of operons containing genes involved in nitrogen metabolism, whereas *E. coli* σ^S is recruited to operons containing genes involved in stationary phase stress response [7-9]. The concept of how sigma factors in the RNAP holoenzyme bind to promoters to initiate transcription is relatively simple. However, the mechanisms by which sigma factors are regulated in response to cellular stress are complex and varied. Most commonly, under normal cellular conditions, an alternative sigma factor is not expressed or is in an inactive state due to binding to a cognate anti-sigma factor. Under stress conditions the anti-sigma factor directly senses or is indirectly stimulated by a relevant stress signal, which decreases its ability to bind and suppress its cognate sigma factor. Different species of bacteria contain different numbers and types of sigma factors that regulate a variety of physiological processes. At this broadest level of transcriptional gene regulation, specificity between sigma factors and promoters within an organism are likely reflective of large swaths of its evolutionary history.

More refined and complex control of transcriptional initiation can be found in the form of accessory factors known as transcription factors. Beyond the core assembly of the transcription initiation complex are transcription factors which bind near promoter sequences and regulate gene expression but are not strictly required for basal transcription initiation. Thus, transcription factors provide an additional layer of regulation to the transcriptional initiation activity of RNAP. A transcription factor can be categorized as an 'activator' or 'repressor' based on whether its binding to the promoter region causes transcription initiation frequency to increase or decrease, respectively. A classic example of a transcription factor is LacI, the repressor of the *lac* operon of *E. coli*. In the absence of lactose, LacI protein binds to a specific region of the *lac* promoter called the operator sequence. The operator sequence overlaps with part of the promoter. Thus, when LacI binds to the operator, RNAP holoenzyme is unable to bind and assemble at the *lac* promoter, thereby repressing expression of the operon. When lactose is present, it is converted in the cell to allolactose which allosterically binds LacI protein. A change in conformation in LacI when bound to allolactose decreases its binding affinity for the *lac* operator. Thus, RNAP can be recruited to the promoter to initiate transcription. The logic of this regulation is that the lactose catabolic genes encoded by the lac operon are only expressed when lactose is available in the cell as a carbon source. Otherwise, the cell conserves the costs of synthesizing lactose catabolic enzymes by repressing their transcription [10]. In a bacterial genome, there can be hundreds of functionally diverse transcription factors. Furthermore, a single promoter sequence often interacts with multiple transcription factors. Thus, the diversity of promoter architectures, sigma factors, and transcription

factors acting in concert provide a high degree of sophistication in the control of bacterial transcription initiation.

1.1.3 Co-transcriptional and post-transcriptional gene regulation

The process of ending the transcription of an mRNA by RNAP is yet another versatile control strategy. There are two main mechanisms of termination that occur widely across bacteria: rho-dependent and rho-independent termination. As their names suggest, the involvement of a protein factor called rho distinguishes these two termination mechanisms. Rho-dependent termination directly involves components of both transcription and translation as these processes are spatially and temporally coupled in bacteria. For that reason, I will explain mechanistic details of rho-dependent termination at the end of Section 1.4 where I describe several translational mechanisms of gene regulation.

The primary factor of rho-independent termination is the secondary structure of the mRNA as it is being transcribed. For this reason, rho-independent termination is also called intrinsic termination. The structure of the intrinsic terminator signal is rather simple: a short RNA hairpin structure immediately followed by a short uracil-rich tract of about 8 nucleotides. The stem of the terminator hairpin is typically GC-rich and five to seventeen base pairs in length, and the loop is three to ten nucleotides in length [11, 12]. Although the exact mechanistic details of how this structure impedes RNAP transcription are still being worked out, the simple structure of the terminator is relatively easy to predict and test experimentally. While an intrinsic terminator is sufficient to terminate transcription, its potency is often enhanced by intrinsic termination factor proteins NusA and NusG [13, 14]. These are highly conserved RNA binding proteins that specifically recognize, bind, and stabilize the intrinsic termination structure to increase the probability of transcription termination when encountered by RNAP. The simple structure and efficient mechanism of the intrinsic terminator makes it a versatile regulator that exists in several different contexts of transcriptional elongation. Most commonly, intrinsic terminators are at the 3' untranslated region (3'UTR) of an operon after the final open reading frame to define the end of the total expressed unit. However, intrinsic terminators can also be found among the open reading frames of polycistronic operons to suppress the basal transcription of downstream genes. Intrinsic terminators are also utilized within the specialized regulatory RNAs: attenuators and riboswitches.

Transcriptional attenuators are found in the 5' untranslated region (5'UTR) of many bacterial operons and can act as specific sensors of intracellular and environmental cues [15]. They are *cis*-regulators of the operon in which they are located. One of the most well-studied transcriptional attenuator regulates the *trp* operon of *Bacillus subtilis* [16, 17]. The 5'UTR of the *trp* transcript contains an RNA sequence that folds into one of two mutually exclusive secondary structures known as the OFF and ON structures. The OFF structure prematurely terminates transcription of downstream structural genes in the *trp* operon by stabilization of an intrinsic terminator hairpin upstream of the first open reading frame. The ON structure permits transcriptional readthrough of the terminator sequence by stabilization of an anti-terminator structure. The choice of folding into the ON or OFF structure is dependent on whether the RNA is bound to a tryptophan-sensing factor known as the *trp* RNA-binding Attenuation Protein (TRAP). TRAP senses high tryptophan concentrations in the cell by directly binding tryptophan molecules. The tryptophan-bound TRAP

structure specifically binds and stabilizes the OFF structure of the *trp* attenuator. Thus, at sufficiently high intracellular levels of tryptophan, expression of the *trp* genes is repressed, thereby conserving the cost of their expression. A variety of other attenuation mechanisms have been discovered, demonstrating their effectiveness as sensory systems to regulate bacterial gene expression [15, 18].

Riboswitches are RNAs that function in a very similar fashion as attenuators, except without the involvement of any cue-sensing protein factor like TRAP. Instead, the riboswitch RNA directly senses the relevant molecular cue to influence downstream regulatory structures including intrinsic terminators and translational control structures. The various classes, mechanisms, and functions of riboswitch RNAs will be discussed extensively in section 2 of this chapter and later on in chapters 2 and 3.

1.1.4 Translational control of gene expression

Similar to transcription, the fundamental components of translation provide a basal level of gene regulation, while accessory gene regulatory elements provide additional layers of control at different stages of translation. Translation of a gene begins with an initiation phase where ribosomal components are recruited and assembled at a ribosome binding site (RBS) of an mRNA transcript. Initiation is a highly coordinated process that results in an initiation complex containing the 70S ribosome (composed of the large 50S and small 30S subunits) loaded with a specialized transfer RNA (tRNA) for translating the first methionine codon. Three initiation factor proteins (IF1, IF2, and IF3) are also required to coordinate the loading of the initial tRNA [19]. The first step in the translation initiation phase is the recruitment of the 30S subunit to the RBS by complementary base pairing. Unlike the RNAP subunits, the ribosomal subunits are each composed of both RNA (rRNAs) and small ribosomal proteins, with the rRNAs providing the main tertiary structural folds and the catalytic mechanism of peptide bond formation. The RBS contains a short sequence of about six nucleotides known as the Shine Dalgarno (SD) sequence, which is complementary to a conserved nucleotide sequence of the 16S rRNA component of the 30S subunit known as the anti-Shine Dalgarno sequence (anti-SD) [20, 21]. In general, deviations away from a consensus SD sequence will decrease the basal translational initiation rate of a gene [22]. Thus, variations in SD sequences provide a mechanism for regulating the basal translation initiation rate on a gene-by-gene basis [23]. After the initiation complex is assembled at the RBS, translation shifts to the elongation phase in which the ribosome reads the codon sequence to synthesize the polypeptide chain. The translational elongation rate of a polypeptide sequence is dependent on the complex interplay between codon usage, amino acid concentrations, tRNA concentrations, post-transcriptional modifications of tRNAs, and other factors [24]. Although these factors are numerous and play essential roles in ultimately producing the final products of the CDMB, the translational elongation phase is not controlled in any direct manner by riboswitch-based gene regulation. Thus, I will move on to describe two more major forms of bacterial gene regulation.

Small bacterial RNAs (sRNAs) are a prevalent and versatile type of gene regulatory factor that can regulate bacterial translation, as well as the stability of mRNAs [25]. In fact, sRNAs can also

modulate the activities of protein targets by direct binding. Unlike *cis*-acting transcriptional attenuators and riboswitches, which directly regulate downstream genes in their operon, sRNAs regulate genes in *trans*, often targeting multiple operons distributed throughout a genome. In many cases, sRNAs target mRNA transcripts by complementary base pairing within the RBS and start codon sequences. This effectively sequesters the RBS region by preventing ribosomal binding to inhibit translation initiation of the target gene. Expression is additionally repressed because the sRNA-mRNA complexes are also degradation signals for RNases. In many cases, specific chaperone proteins bind sRNAs to help direct their interaction with target mRNAs and to protect the sRNAs from degradation by RNases before reaching a target mRNA. The most well-studied sRNA chaperone is Hfq [26]. This RNA binding protein is ubiquitous among bacteria. Through the numerous sRNAs that it chaperones, Hfq has been shown to coordinate complex physiological processes including general growth, stress responses, virulence, quorum sensing, and many others [27-31].

The last gene regulatory mechanism to be discussed in this section is rho-dependent transcriptional termination, which terminates transcription of mRNAs that are not being actively translated by ribosomes. Interestingly, this mechanism relies on the coupling of transcription and translation. In bacteria, transcription and translation are spatiotemporally coupled because ribosomes begin translating mRNA transcripts while in the process of being fully transcribed by RNAP. This differs from eukaryotes where transcription and translation take place in distinct cellular compartments and thus translation of a gene product begins after an mRNA is fully transcribed. The highly conserved bacterial termination factor rho is a homo-hexameric protein complex that binds to mRNAs as they are being actively transcribed by RNAP. The presence of ribosomes bound to an mRNA precludes rho from effectively binding rho utilization (rut) sites commonly distributed throughout mRNA transcripts. Once rho binds to a nascent mRNA at a rut site, it begins translocating along the transcript in the 3' direction towards the RNAP complex. Transcriptional termination is triggered if rho reaches the mRNA exit site of the transcribing RNAP. The exact mechanistic details of how rho causes the termination are still being actively investigated. However, several transcriptomic studies show that rho-dependent termination is a potent and widely occurring form of gene regulation in bacteria. Any process that impedes translation of an mRNA transcript by preventing or disrupting ribosomal binding to that transcript may additionally cause transcriptional termination. Like intrinsic termination, the process of rho-dependent termination is embedded within the mechanisms of other gene regulatory elements including sRNAs and riboswitches.

The past several decades of molecular biology research have revealed the intricate and highly coordinated nature of bacterial transcription and translation. At every step of these processes, it appears that bacteria have evolved gene regulatory mechanisms to precisely control the flow of genetic information. As I move on to Section 2, I will refer back to key mechanisms highlighted above, which are either part of or are intimately tied to riboswitch-based regulation.

1.2 Riboswitch-based gene regulation

1.2.1 Introduction to riboswitch structure and function

Since their initial discovery two decades ago, riboswitches have emerged as efficient and widespread gene regulatory elements that sense and respond to a wide range of molecular cues. Much like transcriptional attenuators, riboswitches are cis-regulatory structured RNAs found within the 5'UTR of bacterial mRNAs. They also modulate the expression of the downstream structural genes of their mRNA transcript. The key distinction between riboswitches and attenuators is that riboswitch RNAs sense small molecule effectors by direct binding, whereas attenuators rely on a protein chaperone to indirectly sense its effector. Riboswitch RNAs are composed of two functional domains: the effector-binding domain called the aptamer, and the regulatory domain called the expression platform. The aptamer domain folds into a specific tertiary structure enabling it to bind a specific small molecule effector. The expression platform modulates intrinsic transcriptional termination or translation initiation of downstream gene targets in response to changes in aptamer structure upon binding the effector (Figure 1.1). Riboswitch classes are named by their cognate effector molecules which are commonly primary metabolites such as organic cofactors, nucleosides, amino acids, and essential metals (Figure 1.2). Currently, there are at least 22 major confirmed classes of riboswitches, some of which have multiple subtypes with distinct effector specificities and/or phylogenetic origins. In the following sections, I will describe various mechanisms and configurations that riboswitches use to sense their cognate effectors and regulate gene expression. Then, I will highlight specific examples of riboswitches and the roles they play in bacterial physiology.

1.2.2 Kinetic and thermodynamic control of transcriptional and translational riboswitches

Most riboswitches are predicted to regulate downstream target genes of their mRNA transcript either by a transcriptional termination or translational control mechanism. The regulatory outcomes of both mechanisms appear the same - genes are either expressed or repressed like an ON/OFF switch. However, the details of each mechanism have important implications for the conditions under which they can operate and the types of organisms in which they can function. In translational riboswitches, the translation initiation rate of a target gene is modulated by controlling the accessibility of the RBS to the ribosome by way of a hairpin structure containing the RBS sequence (Figure 1.1A). When the hairpin is formed, the RBS is sequestered, and the ribosome is unable to bind. Thus, translation initiation is inhibited. In contrast, transcriptional riboswitches control the rate of premature termination of the transcript by controlling the formation of an intrinsic terminator hairpin in the expression platform (Figure 1.1B).

The mechanism of transcriptional riboswitches is often described as a kinetically driven process whereas the translational riboswitch mechanism is described as thermodynamically driven. Time is the main factor that distinguishes the kinetic and thermodynamic features of these mechanisms. In general, it appears that transcriptional riboswitches are more likely to be kinetically driven because there are only a few seconds from when the RNAP begins transcribing the expression platform to when it reaches the transcriptional terminator sequence. Within this short window of time, the effector must bind the aptamer and then the aptamer and expression platform structures

must fold into their effector-bound conformations. The speed, accuracy, and robustness of these RNA folding steps determine the proportion of prematurely terminated transcripts to full-length transcripts, and ultimately the expression level of the target genes [32, 33].

In contrast, translational riboswitches are more likely to be thermodynamically driven because there is a much longer time period in which effector-aptamer binding and RNA folding processes are relevant, with the translational switching mechanism likely to continue long after the RNAP has finished synthesizing the transcript. With the longer timeframe for binding and RNA folding to take place, the system transitions from a state of fast initial kinetics to a state tending towards equilibrium. Thus, the equilibrium rates of effector binding and unbinding, and the energetic favorability of the ON and OFF RNA folds are major determinants of the rate of translational initiation. However, initial co-transcriptional folding of translational riboswitch structures still plays a major role in producing correctly folded RNAs [34, 35].

There is one more layer of complexity for translational riboswitches that must be considered. As a result of inhibiting translation initiation, there is likely a significant increase in rho-dependent termination and RNase-mediated decay of the regulated RNA transcript [36]. While this has been observed in few specific riboswitches, the extent to which this generally occurs is still unclear [37]. If gene repression by translational riboswitches is commonly followed by rho-dependent termination and/or mRNA decay, then this would provide a robust means for translational riboswitches to effectively control the expression of all the structural genes within an operon, not just the very first one. Interestingly, a recent report suggests that for *B. subtilis*, the transcription and translation process is not as tightly coupled as previously assumed [38]. Thus, rho-dependent termination may not be a fundamental feature of riboswitch-based gene expression in some bacteria, particularly the Firmicutes phylum.

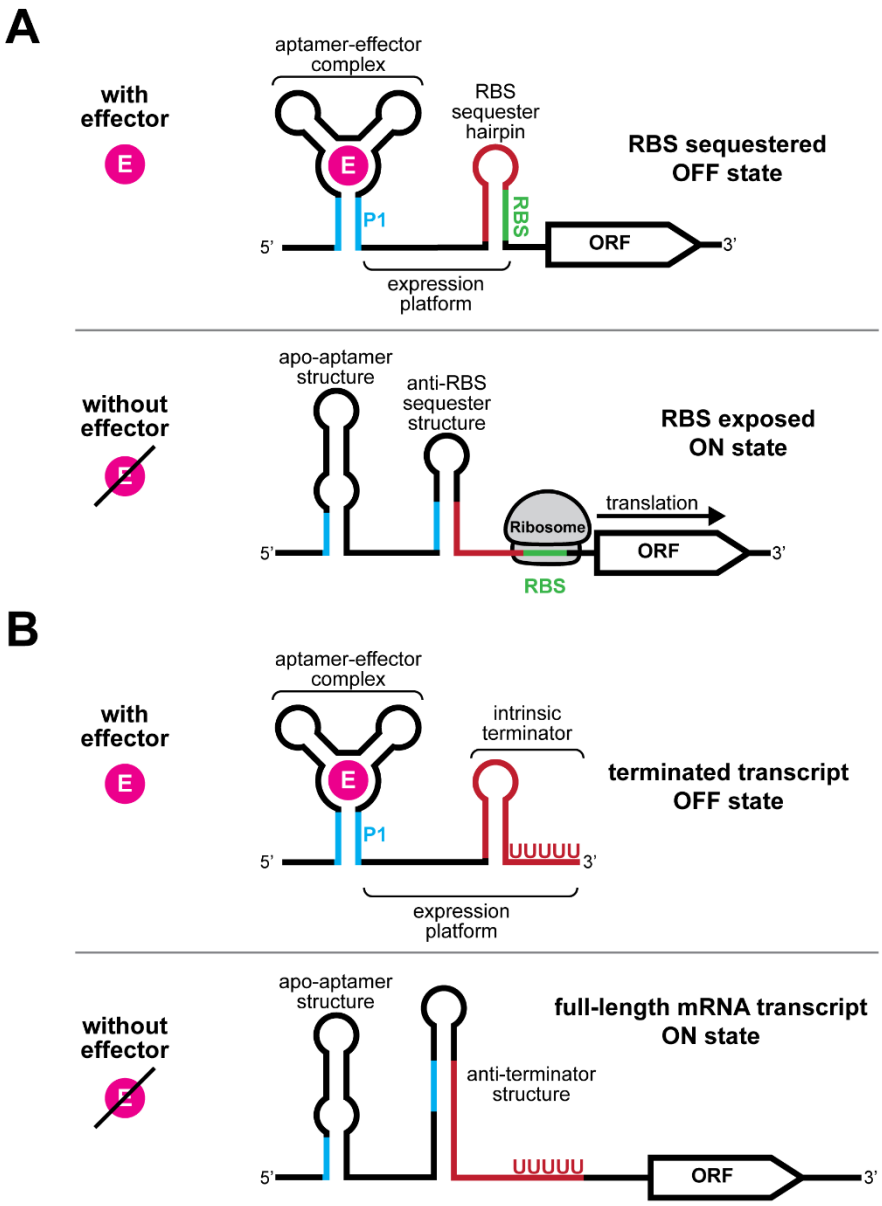


Figure 1.1 – Common riboswitch mechanisms. (A) Translation initiation control. (B) Transcriptional termination control. In both mechanisms, effector-dependent changes in aptamer structure, particularly the P1 stem, influence expression platform regulatory structures. RBS = ribosome binding site, ORF = open reading frame.

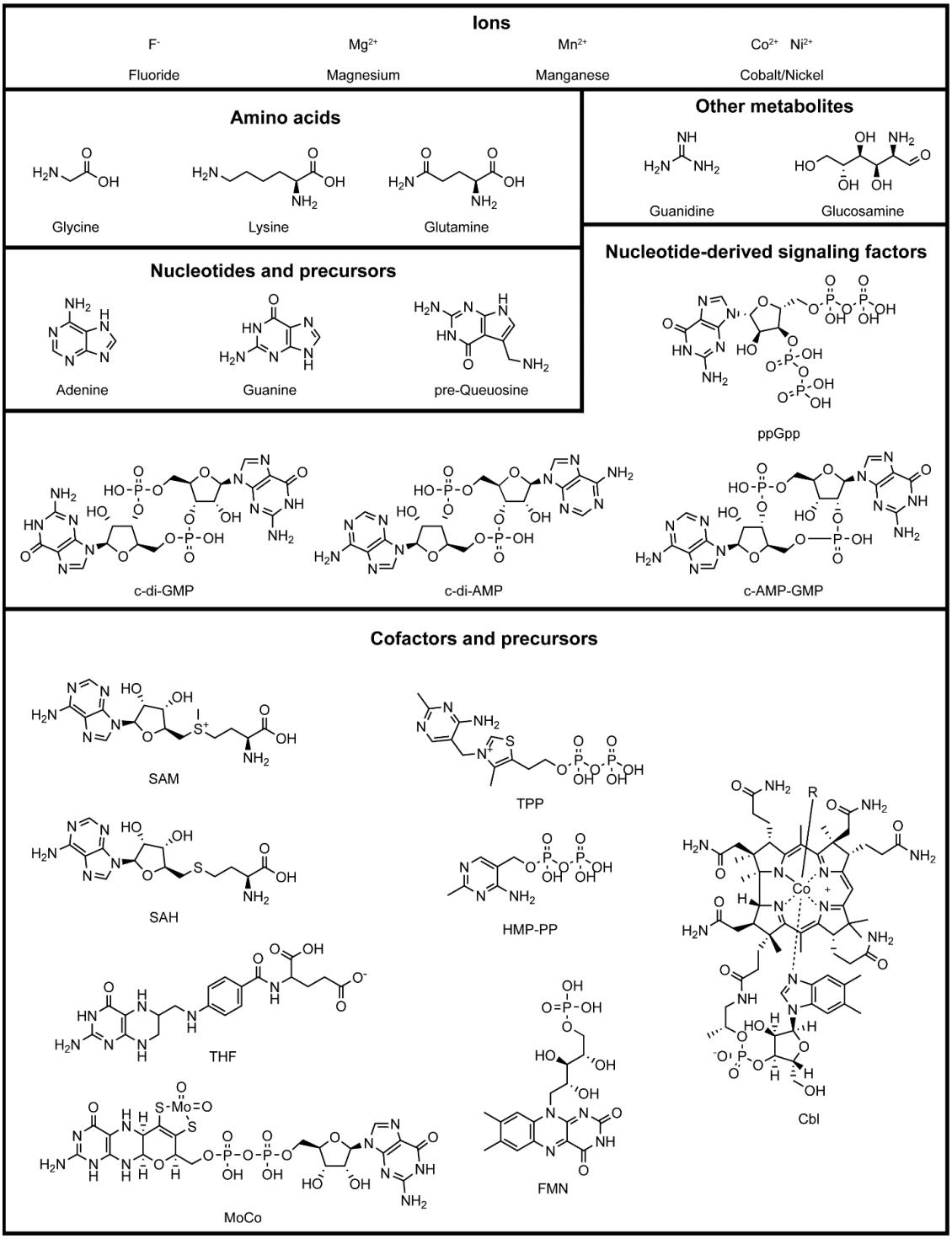


Figure 1.2 – Naturally occurring small molecule riboswitch effectors.

1.2.3 Less prevalent riboswitch mechanisms and features

Although the vast majority of riboswitches are predicted to control transcriptional termination or translation, there are notable cases of riboswitches that use other regulatory mechanisms [39]. These include modulation of self-cleaving ribozyme activity in *glmS* riboswitches, mRNA splicing of certain thiamin pyrophosphate (TPP) riboswitches, and indirect control of gene targets by sRNAs by special cases of cobalamin riboswitches [40-46]. While these mechanisms and features are less common among riboswitches, they demonstrate the versatility of riboswitches as a form of gene regulation.

Another uncommon, but noteworthy feature of riboswitches is tandem configuration where multiple riboswitches are linked in series to regulate a single operon. Bioinformatic predictions suggest that tandem riboswitches are configured as doublets of the same or different classes of riboswitch [47]. In fact, no tandem arrangements of greater than two riboswitches have been identified in comparative genomic studies. However, one case of a naturally occurring tandem triplet riboswitch from a specific environmental strain isolate of *Bacillus thuringiensis* has been reported [48]. This triple tandem riboswitch is composed of three transcriptional riboswitches of the same class which sense the effector molecule cyclic-di-guanosine monophosphate (c-di-GMP). The tandem configuration of this riboswitch was shown to increase the sensitivity to c-di-GMP by lowering the concentration of c-di-GMP required to trigger gene regulatory switching. The modular architecture and functional consequences of tandem riboswitch arrangement will be discussed in further detail in Chapter 3.

1.2.4 Nucleoside and nucleotide sensing riboswitches

Perhaps the simplest metabolites to envision as riboswitch effectors are nucleosides and nucleotides. After all, a riboswitch RNA is composed of nucleotides that interact with each other through base-pairing and stacking interactions to produce secondary and tertiary structures. It seems that a structured RNA should be well-suited for detecting free nucleobases by direct binding. Indeed, several classes of riboswitches detect nucleosides, nucleoside precursors, as well as nucleotide-derived signaling molecules [49].

The purine riboswitch class detects adenine and guanine nucleosides to regulate genes of purine biosynthesis, interconversion, and transport [50]. The purine riboswitch aptamer secondary structure is composed of three base-paired helices (P1, P2, and P3) that form around a three-way junction serving as the effector binding site. Conserved residues within the three-way junction form a binding site for the purine effector, and a single conserved position in the binding site varies as either uracil or cytosine to mediate selective effector detection by Watson-Crick base pairing to adenine or guanine, respectively. The effector selectivity of a purine riboswitch can be accurately predicted solely based upon the presence of cytosine or uracil at this conserved position [51, 52].

Interestingly, both repressor and activator purine riboswitches exist. In both cases, purine binding to the aptamer stabilizes formation of the P1 stem. In activator purine riboswitches, P1 stem stabilization promotes formation of an anti-terminator or anti-RBS hairpin structure to disrupt an intrinsic terminator hairpin or RBS hairpin, thereby upregulating gene expression. Repressor purine riboswitches do the opposite, with P1 stabilization disrupting the anti-terminator or anti-

RBS structures and promoting the terminator or RBS hairpin [49, 53, 54]. Effector-dependent stabilization of a conserved P1 stem to direct the structural formation of the expression platform is a common mechanism found across many riboswitches, not just the purine class. Purine riboswitches that regulate purine biosynthesis operons tend to repress gene expression in response to purines as a feedback inhibition mechanism to maintain homeostatic purine production. In contrast, purine riboswitches upregulate purine efflux transporters in response to overaccumulation of purines in the cytoplasm [53, 55].

The pre-queuosine$_1$ (PreQ1) riboswitch class detects PreQ1, a precursor to the nucleoside queuosine. Queuosine is produced *de novo* as a post-transcriptional modification of 7-deazaguanine (itself a post-transcriptional modification of guanine) in the wobble position of certain tRNA anticodons [56, 57]. This hyper-modification of the wobble position to queuosine has been shown to increase translational efficiency of tyrosine, aspartic acid, asparagine, and histidine in bacteria, fungi, plants, and animals. Intriguingly, only bacteria are known to produce queuosine *de novo*, while eukaryotes must salvage queuine (the nucleobase derivative of queuosine) from their environment. This had led to speculation that queuine may be an essential micronutrient in humans and other animals [58, 59]. Bacteria detect intracellular PreQ1 with riboswitches in order to control the expression of queuosine biosynthetic genes and transporters [60, 61]. Much like the purine riboswitches, PreQ1 riboswitches direct the formation of regulatory structures of the expression platform by effector binding-dependent stabilization of a conserved P1 stem [61]. There are three major subclasses of PreQ1 riboswitches (named PreQ1-I, PreQ1-II, and PreQ1-III) with distinct folds, suggesting that they may have evolved independently [62-64]. There are also independently evolved isoenzymes in the biosynthetic pathway of queuosine [57, 65-67]. Yet curiously, disruption of queuosine biosynthesis has not been strongly associated with any growth major defects in bacteria [68]. Although the genuine physiological significance of queuosine remains unclear, the multiple paths that have evolved to control its production underscore its importance.

Several nucleotide-derived bacterial signaling molecules such as cyclic dinucleotides, guanine tetraphosphate (ppGpp), and 5-aminoimidazole-4-carboxamide riboside 5'-triphosphate (ZTP) are also detected by riboswitches [69-72]. The riboswitches that sense cyclic di-guanosine monophosphate (c-di-GMP) and cyclic di-adenosine monophosphate (c-di-AMP) are phylogenetically distinct classes. Two riboswitch classes, c-di-GMP-I and c-di-GMP-II, specifically recognize c-di-GMP, whereas a single c-di-AMP riboswitch class detects c-di-AMP. Unlike the purine riboswitches, each of the cyclic dinucleotide riboswitches relies on a combination of canonical and noncanonical base pairing and base stacking interactions between aptamers and effectors to achieve selective binding [73-75]. Interestingly, a subclass of c-di-GMP-I riboswitches found in the *Geobacter* genus have altered selectivity that permits either promiscuous binding to both c-di-GMP and cyclic AMP-GMP, or selective binding for cyclic AMP-GMP. The changes in selectivity are primarily achieved by nucleotide variation of a single conserved residue in the aptamer that base pairs with the effector [76, 77].

1.2.5 Amino acid sensing riboswitches

Three amino acids (lysine, glutamine, and glycine) have been identified as riboswitch effectors. The initial discoveries of the lysine and glycine riboswitches were unexpected. Given that amino acids are much smaller molecules than the previously known nucleoside and organic cofactor effectors, it was surprising that structured RNAs could discriminate lysine and glycine from the eighteen other standard amino acids [78]. The lysine riboswitches have typical characteristics of riboswitches in that they regulate genes involved in the synthesis and uptake of lysine and they operate by transcriptional termination and translational control mechanisms [60, 79-81]. The glutamine riboswitches are similarly hypothesized to regulate biosynthesis and transport of glutamine. However, precise regulatory mechanisms of glutamine riboswitches have not been fully elucidated [82, 83].

Glycine riboswitches are unusual in that they commonly contain a double tandem aptamer domain fused to a single expression platform, although some glycine riboswitches contain a more typical architecture of a single aptamer fused to an expression platform. Initially, it was suggested that the double aptamer configuration served a cooperative binding mechanism, perhaps to increase glycine sensitivity [84]. Subsequent *in vitro* biochemical and *in vivo* studies showed that the binding properties of glycine to singlet glycine aptamers is similar to doublet aptamers, so it is unclear what functionality, if any, the tandem double aptamer configuration serves [85-88]. Interestingly, a recent phylogenetic analysis suggests that the ancestral glycine riboswitch was a double aptamer type and that single aptamer glycine riboswitches are evolutionary descendants [89]. Another other notable feature of glycine riboswitches is that they commonly operate as activators of genes required for glycine degradation and efflux. Mutation of glycine riboswitches in *B. subtilis* led to severe growth inhibition in high glycine conditions, suggesting that excess glycine is toxic to cells. Thus, rather than controlling biosynthesis of glycine via feedback inhibition, it appears that glycine riboswitches activate genes to counter toxic over-accumulation of glycine. [85, 90].

1.2.6 Metal and fluoride sensing riboswitches

Divalent metals are essential structural cofactors and coenzymes in bacteria. However, excess intracellular concentrations of these metals can be highly toxic to cells. Therefore, intracellular metal concentrations are tightly regulated in bacterial cells. Riboswitches are one of several regulatory systems that bacteria employ to sense and respond to fluctuating metal concentrations [91]. The currently known classes of metal sensing riboswitches are magnesium (Mg^{2+}) riboswitches, manganese (Mn^{2+}) riboswitches, and Nickel/Cobalt (NiCo) riboswitches [92]. Mg^{2+} riboswitches bind divalent metals promiscuously with relatively low affinity. Intracellular Mg^{2+} concentrations are normally in the high µM to low mM range, whereas other divalent metals are maintained at low µM or below. Thus, Mg^{2+} is considered to be the physiologically relevant effector of the Mg^{2+} riboswitches. Notably, the Mg^{2+} riboswitch aptamer contains binding sites for four Mg^{2+} ions. Binding of Mg^{2+} to riboswitches typically represses Mg^{2+} uptake genes to prevent overaccumulation of Mg^{2+} from the extracellular environment [93].

Mn^{2+} riboswitches commonly regulate Mn^{2+} uptake and efflux genes to maintain homeostatic intracellular Mn^{2+} concentrations. Mn^{2+} toxicity typically results from Mn^{2+} competing with iron binding to iron-dependent enzymes. Tolerance of Mn^{2+} varies greatly between bacterial species

depending on which iron-binding proteins are required for cell survival in a given environment [94]. Like the Mg^{2+} riboswitches, Mn^{2+} riboswitch aptamers contain multiple Mn^{2+} ion binding sites. Structural probing of the Mn^{2+} riboswitch aptamer and analysis of the crystal structure suggests that the effector binding sites of the Mn^{2+} aptamer allow selective binding of Mn^{2+} and exclusion of Mg^{2+}. However, the precise mechanism of selectivity remains unclear [95, 96]. The NiCo riboswitch operates in a similar manner to the Mg^{2+} and Mn^{2+} riboswitches, with the aptamer containing multiple effector binding sites, and having metal transport proteins as regulatory targets. The NiCo riboswitches help bacterial cells maintain intracellular concentrations of Ni^{2+} and Co^{2+} in the sub μM to low μM range [97].

The last riboswitch to mention in this section is the fluoride (F^-) riboswitch. Fluorine is not a metal, but rather a highly reactive halogen that is potently toxic to bacteria. There are no known beneficial roles of F^- in bacterial physiology. The F^- riboswitch senses intracellular F^- to activate expression of efflux pumps that expel toxic fluoride compounds from the cell. Although named a F^- riboswitch, the effector of these riboswitches is actually magnesium fluoride, which itself is a toxic compound that readily forms in the cytoplasm [98, 99].

1.2.7 Organic cofactor sensing riboswitches

Several riboswitch classes detect organic cofactors and related metabolites. These cofactors include cobalamin (Cbl), flavin mononucleotide (FMN), *S*-adenosylmethionine (SAM), *S*-adenosylhomocysteine (SAH), molybdenum cofactor (MoCo), tetrahydrofolate (THF), and thiamin pyrophosphate (TPP) [60, 100]. As seen in the riboswitches already discussed, organic cofactor riboswitches regulate genes involved in the biosynthesis and uptake of their cognate effector molecules in order to maintain homeostasis [101]. These riboswitches also operate by the typical transcriptional termination and translational control mechanisms [102]. Here, I will highlight just a few of the organic cofactor riboswitches that have notable structural and functional features that are distinct from the riboswitches previously discussed. The cobalamin riboswitch (Cbl-riboswitch) will be discussed in detail in Section 3.

SAM is an essential enzyme co-substrate required for numerous challenging methyl transfer and methylation reactions in all organisms. SAM donates methyl groups or forms radicals to catalyze biochemical reactions including the modification of DNA, RNA, proteins, and lipids, as well as the synthesis of many secondary metabolites. In many cases, methyltransferase reactions demethylate SAM to produce SAH, which cannot serve as a methyl donor. Thus, SAH must be catabolized to homocysteine and adenosine and fed back into SAM biosynthesis to regenerate a methyl donor [103]. Bacteria use SAM-specific, SAH-specific, and promiscuous SAM/SAH riboswitches to sense the intracellular concentrations of SAM and its demethylated byproduct SAH [104-107]. In concordance with the fundamental role of SAM in cell physiology, there are four known classes of SAM-specific riboswitches that are thought to have evolved independently in bacteria. X-ray crystal structures of the aptamer domains of SAM- and SAH-specific riboswitches have revealed mechanisms of effector specificity. SAM-specific binding is achieved by interactions between residues in the effector binding site and the sulfonium cation of SAM [107]. In contrast, the SAH riboswitch achieves specificity by steric effects, in which the effector binding site allows binding of the less bulky and uncharged aminocarboxypropyl side chain of SAH but

excludes the bulkier and charged SAM side chain [104, 105]. Effector promiscuity of the SAM/SAH riboswitch results from an overall lack of interaction between the RNA and the aminocarboxypropyl side chain [105].

Flavin coenzymes are a group of chemically related metabolites including FMN, flavin adenine dinucleotide (FAD), and riboflavin. The core chemical feature of flavins is the isoalloxazine tricyclic ring structure. Numerous essential biochemical reactions are catalyzed by FMN, FAD, and several naturally occurring modified forms of FMN and FAD [108]. Although many flavins exist, FMN appears to be the only flavin coenzyme that is directly detected by riboswitches to maintain homeostatic flavin concentrations in bacterial cells [109, 110]. The secondary metabolite roseoflavin is an interesting example of an antibiotic that targets riboswitches as a primary mode of action [111, 112]. Roseoflavin was discovered as a natural product of *Streptomyces davawensis* that inhibits growth of other bacterial species. *B. subtilis* takes up roseoflavin and chemically modifies it to roseoflavin mononucleotide (RoFMN) and roseoflavin adenine dinucleotide (RoFAD). RoFMN and RoFAD bind FMN- and FAD-dependent enzymes, respectively, but are inactive cofactors [113, 114]. Surprisingly, it was observed that *S. davawensis* does not resist its own roseoflavin by preventing production of inactive cofactor forms . Instead, it was found that the *S. davawensis* FMN riboswitch that regulates flavin biosynthesis is insensitive to RoFMN, whereas the *B. subtilis* FMN riboswitches are sensitive to RoFMN [115]. Indeed, *B. subtilis* mutants selected for growth in the presence of roseoflavin had mutations in the FMN riboswitch, and overproduced active flavins to overcome roseoflavin toxicity. The discovery of the mode of action of roseoflavin has significant implications in drug development, as it demonstrated that riboswitches can be effectively targeted by antibiotics [116, 117].

The TPP riboswitch is the most prevalent class of riboswitch among bacteria. It is also the only riboswitch commonly found in eukaryotes including fungi and plants, though it is not present in animals [118]. TPP-dependent enzymes are required for several essential biochemical reactions of carbohydrate and amino acid metabolism. The TPP riboswitches of bacteria operate by transcriptional termination and translation control, but in eukaryotes TPP riboswitches modulate alternative splicing to regulate target gene expression [119]. In the fungus *Neurospora crassa*, TPP riboswitch activity results in alternatively spliced transcripts containing distinct translational start sites and different open reading frames [41]. In *Arabidopsis thaliana*, the alternative spliceforms produced by this mechanism differ in stability, thus modulating gene expression by controlling the quantities of transcript [42].

1.3 Cbl-riboswitch regulation of corrinoid-related metabolism

1.3.1 Structures and functions of Cbl-riboswitches and their effectors

Cbl-riboswitches are the second most abundant class of riboswitches among bacteria [64]. They commonly regulate genes involved in or related to the biosynthesis, transport, and usage of Cbl, an enzyme cofactor. Cbl-riboswitches are found in 5'UTRs of operons that they regulate. They have the typical riboswitch architecture of an aptamer domain fused to an expression platform and operate by both transcriptional termination and translational control mechanisms [101, 102, 120].

In these ways, Cbl-riboswitches are quite similar to other riboswitch classes. However, there are two exceptional features of Cbl-riboswitches that will drive most of the questions of my research.

The first exceptional feature of Cbl-riboswitches is the mechanism for transducing the effector-binding state of the aptamer to the downstream regulatory structures of the expression platform. In most other riboswitch classes, the formation of the P1 stem is dependent upon aptamer-effector binding, which allows the P1 stem to directly influence the switching of downstream expression platform structures (Figure 1.1). The transduction mechanism in Cbl-riboswitches is less direct. Aptamer-Cbl binding stabilizes a long-range kissing loop interaction between the conserved loops L5 of the aptamer and L13 of the expression platform (Figure 1.3) [121, 122]. The kissing loop stabilizes formation of stem P13, which directs the downstream formation of an intrinsic terminator hairpin or RBS sequester hairpin. All reported Cbl-riboswitches appear to use this kissing loop mechanism to operate as repressors of their target genes . The kissing loop structure and mechanism becomes a highly important part of my work described Chapter 2.

The other exceptional feature of the Cbl-riboswitch is that there exist several naturally occurring chemical variants of the cognate effector Cbl (Figure 1.4A). It is unclear whether these Cbl variants called corrinoids are also effectors of Cbl-riboswitches. The corrinoid biosynthetic pathway is complex consisting of roughly 30 enzymatic steps including separate parallel aerobic and anerobic parts of the pathway [123-125]. In addition, the corrinoid biosynthetic pathway produces multiple end products, which are enzyme cofactors called ‘complete corrinoids’ (and also referred to as ‘cobamides’) [126, 127]. Cbl is just one type of complete corrinoid. Complete corrinoids chemically vary in two parts of the chemical structure: the upper and lower axial ligands of the central cobalt ion (Figure 1.4B). There are three naturally occurring upper ligands (methyl, hydroxy, and 5’-deoxyadenosyl) and at least 12 naturally occurring lower ligands (Figure 1.4C) [128-138].

Interestingly, bacteria are known to produce, import and use specific cobamides for their metabolic needs [139-154]. Yet, many corrinoid-dependent species are unable to produce corrinoids *de novo* (auxotrophy), suggesting that they rely on corrinoids produced by other bacteria in the environment [155, 156]. The Taga Lab and others hypothesize that corrinoid auxotrophy and specificity are major drivers of interspecies interactions within microbial communities [127, 157-162]. Therefore, it has been useful to study corrinoid specificity of proteins such as enzymes and transporters in this context, but little is known about the corrinoid specificity of gene regulation by Cbl-riboswitches [163]. This is the central problem addressed in my research in Chapters 2 and 3.

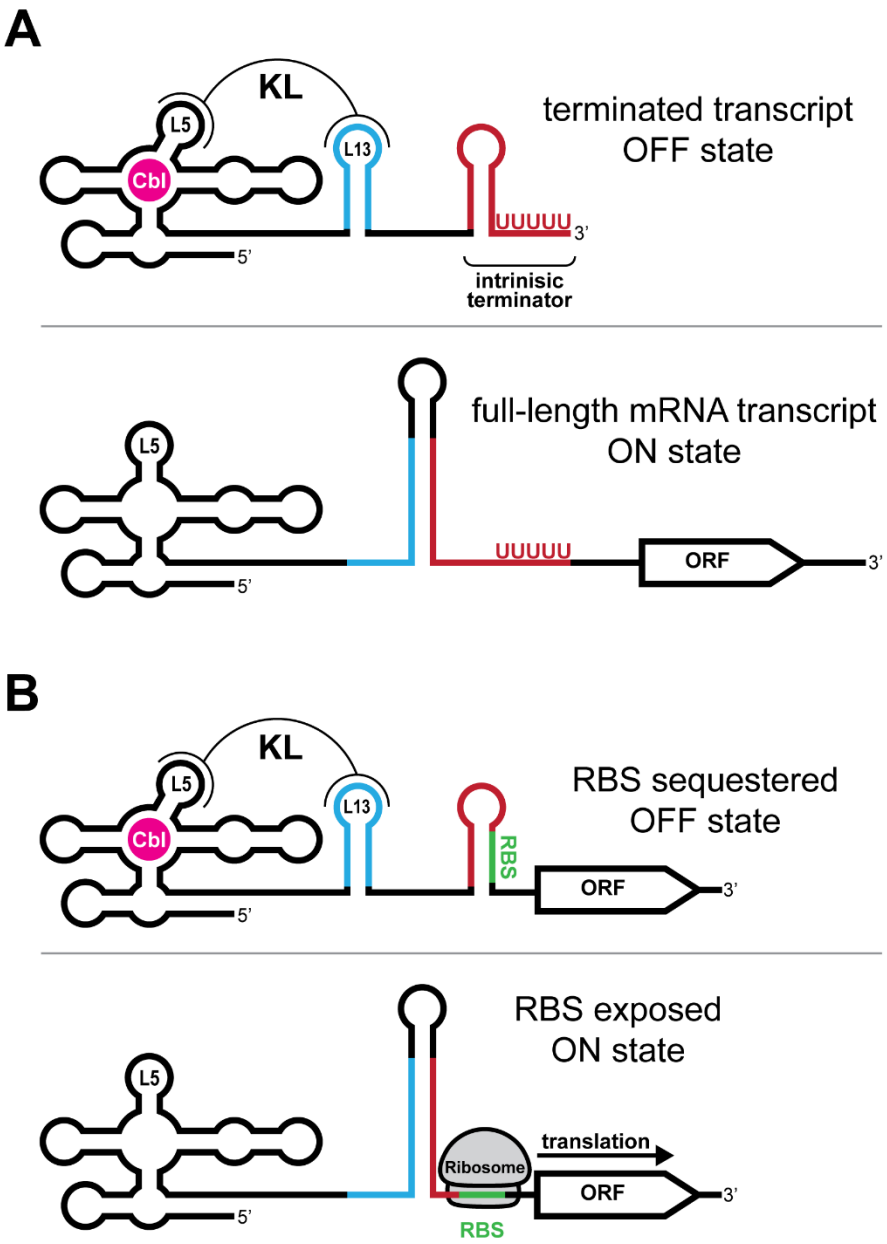


Figure 1.3 – A kissing loop interaction drives Cbl-riboswitch regulatory mechanisms. The aptamer-effector complex stabilizes a kissing loop to control (A) transcriptional termination and (B) translational initiation. L5 = loop 5, L13 = loop 13, KL = kissing loop, RBS = ribosome binding site, ORF = open reading frame.

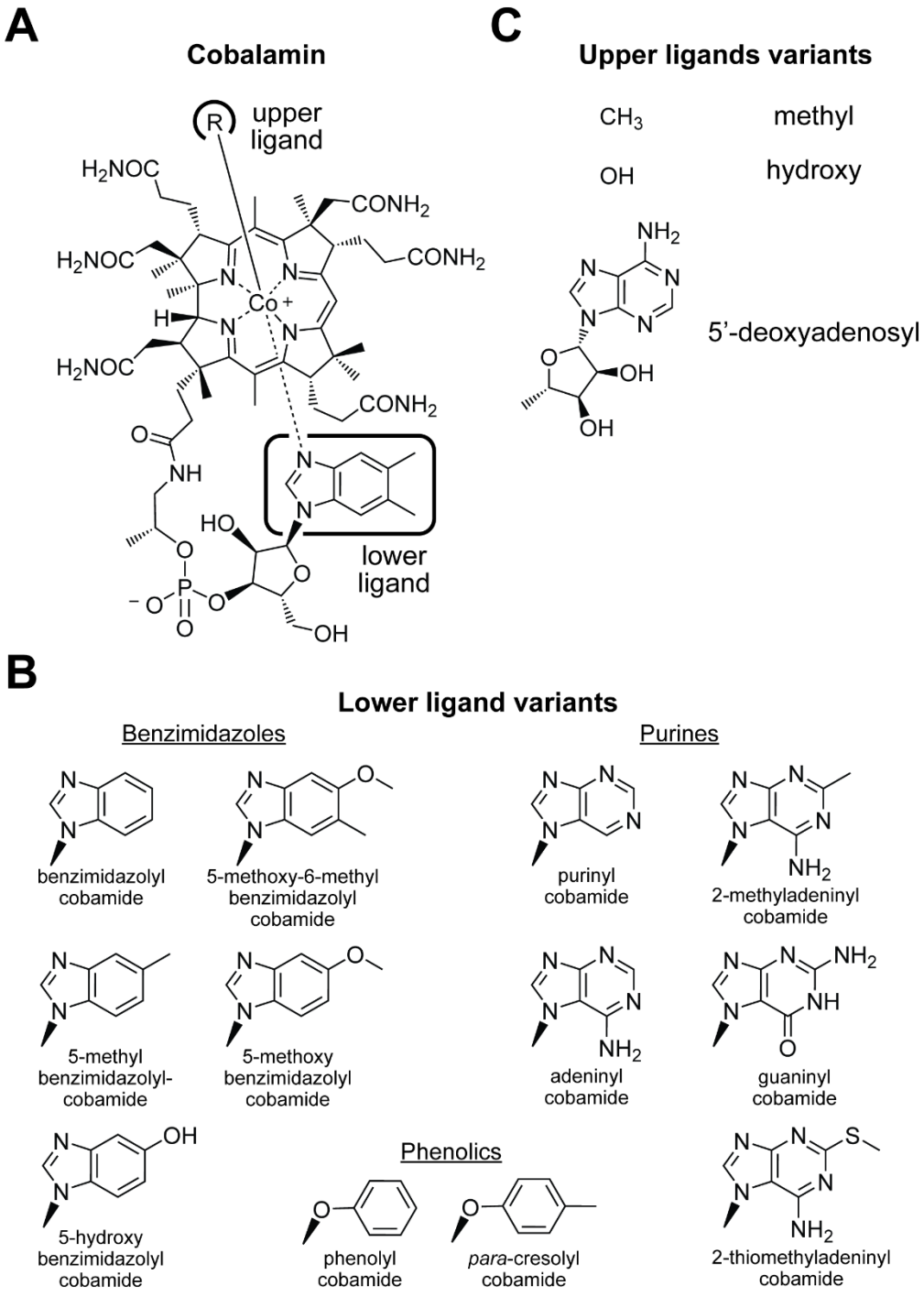


Figure 1.4 – Chemical diversity of corrinoid cofactors. (A) Chemical structure of cobalamin (Cbl). R-group indicates variable upper ligand. The lower ligand of Cbl is 5,6-dimethylbenzimidazole. (B) Alternative lower ligands grouped by chemical classes – benzimidazoles, purines, and phenolics. (C) Natural variants of the catalytic upper ligand moiety of corrinoids.

1.3.2 Physiological processes regulated by Cbl-riboswitches

Corrinoid-dependent enzymes are involved in numerous bacterial metabolisms. These include common processes such as methionine synthesis (via corrinoid-dependent methionine synthase), synthesis of deoxyribonucleotides (via corrinoid-dependent ribonucleotide reductase), certain post-transcriptional modifications of tRNAs (via corrinoid-dependent epoxyqueuosine reductase), and the catabolism of certain branched chain amino acids, short chain fatty acids, and polyhydroxyalkanoates (via methylmalonyl-CoA mutase) [57, 65, 164-167]. Corrinoids are also involved in less prevalent processes such as synthesis of certain antibiotics, fermentation of various carbon and nitrogen sources (glutamate, lysine, ornithine, ethanolamine, and propanediol), mercury methylation, and respiration of organohalides [168-183]. The versatility of corrinoid-dependent reactions stems in part from the chemical diversity of the corrinoid upper ligand (Figure 1.4C). The methyl upper ligand of methylcobalamin mediates methyl transfer reactions, whereas the 5'-deoxyadenosyl upper ligand of adenosylcobalamin catalyzes radical-mediated carbon-carbon bond rearrangements [184]. The breadth and diversity of metabolisms involving corrinoids likely reflects the catalytic utility of cobalt.

Cbl-riboswitches commonly regulate expression of corrinoid biosynthetic genes, corrinoid uptake genes, and genes encoding corrinoid-independent enzymes. The regulation of corrinoid biosynthesis and uptake by Cbl-riboswitches is considered a form of feedback control, as seen in other riboswitch classes [102]. The biosynthesis and uptake of corrinoids leads to accumulation of intracellular corrinoids, which in turn triggers the Cbl-riboswitch to repress expression of the biosynthetic and uptake genes. In the Cbl producer species *Bacillus megaterium* and *Propionibacterium* strain UF1, mutations in Cbl-riboswitches regulating Cbl biosynthesis operons result in overproduction of Cbl [185-187]. Thus, this regulatory strategy presumably minimizes the costs associated corrinoid biosynthesis and uptake, and perhaps maintains homeostatic levels of intracellular corrinoids.

In contrast to corrinoid biosynthesis and uptake, the regulation of corrinoid-independent enzymes by Cbl-riboswitches is not explained by simple feedback control. The rationale of this regulation requires an explanation of a special characteristic of certain corrinoid-dependent enzymes. Namely, bacteria often possess both corrinoid-dependent and corrinoid-independent isoenzymes and can use either enzyme to carry out the same biochemical reaction. There is some evidence suggesting that corrinoid-dependent catalysis is more robust or efficient than the corrinoid-independent alternative, so it is beneficial to use the corrinoid-dependent isoenzyme when possible [188, 189]. Thus, it appears that the Cbl-riboswitch regulation of the corrinoid-independent isoenzyme allows the cell to choose which isoenzyme is expressed and used depending on the intracellular concentration of corrinoids. There exist corrinoid-dependent and -independent isoenzymes of methionine synthase, ribonucleotide reductase, and epoxyqueuosine reductase. Corrinoid-independent methionine synthases and ribonucleotide reductases are common regulatory targets of Cbl-riboswitches, although one potential example of a Cbl-riboswitch regulating a corrinoid-independent epoxyqueuosine reductase has been observed in *Stenotrophomonas maltophilia* [47]. The impact of corrinoid chemical diversity on this regulatory scheme is unknown and I explore it in my work in Chapter 2.

1.3.3 The utility of Cbl-riboswitches for unveiling new biology

Studying corrinoid specificity of Cbl-riboswitches should lead to better understanding of how bacteria sense and respond to chemically diverse corrinoids. However, another interesting benefit of understanding the intricacies of corrinoid-dependent metabolism and gene regulation is that this knowledge can help elucidate new facets of corrinoid biology [190]. The Taga Lab used this kind of approach to discover new parts of the corrinoid biosynthetic pathway. In Hazra et al. 2015, members of the Taga Lab identified the genes required for the anaerobic biosynthesis of 5,6-dimethylbenzimidazole (DMB), the lower ligand base of Cbl (Figure 1.4A) [141]. Previously, it was known that the enzyme BluB converts FMN to DMB through a single but complex enzymatic reaction [191-193]. Critically, BluB catalysis requires oxygen, yet some strictly anaerobic bacterial species are also capable of producing DMB *de novo*. Several isotopic labeling analyses had shown that anaerobic biosynthesis of DMB in *Eubacterium limosum* most likely branches from the purine biosynthetic pathway [194-201]. Thus, Hazra et al. 2015 set out to find the hypothesized genes of anaerobic DMB synthesis in *E. limosum*. Their key strategy was to examine Cbl-riboswitch regulated operons containing genes of unknown function in the genome of *E. limosum* reasoning that this pathway would likely be regulated. The authors pinpointed strong candidate genes and subsequently tested and confirmed several enzymatic steps of the pathway in *E. limosum* and other bacterial species. Identifying these genes may have proved much more challenging without the Cbl-riboswitch analysis to provide the initial clues.

The intrinsic connections between gene regulators and their targets are further utilized in my own research approaches. In Chapter 2, I explore how the corrinoid selectivity of Cbl-riboswitches relates to the corrinoid selectivity of their regulatory targets. In Chapter 3, I use the relationship between riboswitch and target gene function to discover Cbl-riboswitches with novel functionalities.

1.4 Chapter 1 conclusion

The study of molecular biology and gene regulation in bacteria has fundamentally shaped our understanding of how cells function. Naturally as we gain new insights, new questions arise, especially as the field of microbiology seeks to understand the behaviors of microbes in natural environments and in the context of complex microbial communities. In the chapters that follow, I continue the rich line of inquiry - 'How do cells express the right genes at the right time and place?' - by investigating the connections between Cbl-riboswitches, corrinoids, and the metabolisms that they control.

Chapter 2 – Patterns and mechanisms of cobalamin riboswitch responses to chemically diverse corrinoid cofactors

2.1 Abstract

Microbes rely on cobalamin and other corrinoids for many key metabolic functions. Yet, access to corrinoids in microbial environments is often limited by costly biosynthesis performed by a small number of *de novo* corrinoid producer species. Thus, corrinoid-related genes are often highly regulated by cobalamin riboswitches, non-coding RNAs that were originally found to bind and respond to cobalamin. While over a dozen corrinoids are naturally produced by microbes, it is unknown how cobalamin riboswitches generally respond to corrinoids other than cobalamin. To address this problem, we developed a live cell reporter system in *Bacillus subtilis* to systematically measure gene regulatory activity of cobalamin riboswitches. Characterization of 38 reporter strains revealed cobalamin riboswitches that range from semi-selective to promiscuous in their response to corrinoids. Through a series of analyses involving chimeric riboswitches, spectroscopy of corrinoid tail analogs, and riboswitch structural models, we propose a mechanism in which cobalamin riboswitches are sensitive to the conformational differences between corrinoid tail variants, rather than differences in the chemical structures of corrinoid tail. Lastly, we test and find support for a gene regulatory strategy in which corrinoid selectivity of the cobalamin riboswitch complements the corrinoid-specific requirements of the cell. Overall, these findings provide new conceptual and mechanistic insights into how corrinoid chemical diversity impacts bacterial gene regulation, further underscoring the crucial role of corrinoid physiology in microbial systems.

2.2 Introduction

Controlling gene expression is an essential task that cells accomplish in a variety of ways. Non-coding RNAs are one such means of gene regulation, acting in parallel or in concert with historically better-studied protein-based mechanisms [202]. In bacteria and archaea, riboswitches are a widespread type of gene regulatory RNA with the distinct ability to sense particular intracellular metabolites by direct binding in order to control gene expression [203]. These RNAs are typically located in the 5'-untranslated region of mRNA, functioning as cis-acting regulators of downstream genes in the operon. Riboswitches contain two functional domains: an aptamer and an expression platform. The aptamer domain adopts a three-dimensional structure that can bind its cognate effector molecule. The expression platform domain is a regulatory switch that interprets the ligand-binding state of the upstream aptamer to promote or disrupt the transcription or translation of downstream genes [204]. The diversity of riboswitch effectors and regulatory mechanisms has revealed fundamental insights into how bacteria sense and respond to dynamic environments and has also driven new approaches for precise control and manipulation of microbes for human purposes [116, 205-207].

The first riboswitch discovered was the *Escherichia coli btuB* cobalamin (Cbl) riboswitch, which directly binds Cbl (a cofactor form of vitamin B_{12}) as its cognate effector (Figure 2.1A) [101]. Cbl-riboswitches typically regulate genes involved in the synthesis, transport, and usage of Cbl, an enzyme cofactor required to catalyze challenging chemical transformations such as methyl transfer reactions and radical-mediated carbon-carbon bond rearrangements [184, 208]. Cbl-dependent catalysis is present in common metabolic pathways including methionine synthesis, deoxyribonucleotide synthesis, tRNA modification, and the degradation of certain amino acids, fatty acids, and biopolymers [171, 173, 174, 176, 177, 183, 209-216]. Cbl is also required for rarer metabolic processes involved in antibiotic synthesis, mercury methylation, catabolism of steroids, and many others [168, 179-181, 217-226]. Comparative genomic studies indicate that Cbl-dependent metabolism is common among microbes, with Cbl-riboswitches being a prevalent mode of gene regulation [64, 156, 227]. However, an underappreciated aspect among Cbl-riboswitch studies is that Cbl is just one of several naturally occurring chemical analogs called corrinoids [127], which potentially act as Cbl-riboswitch effectors.

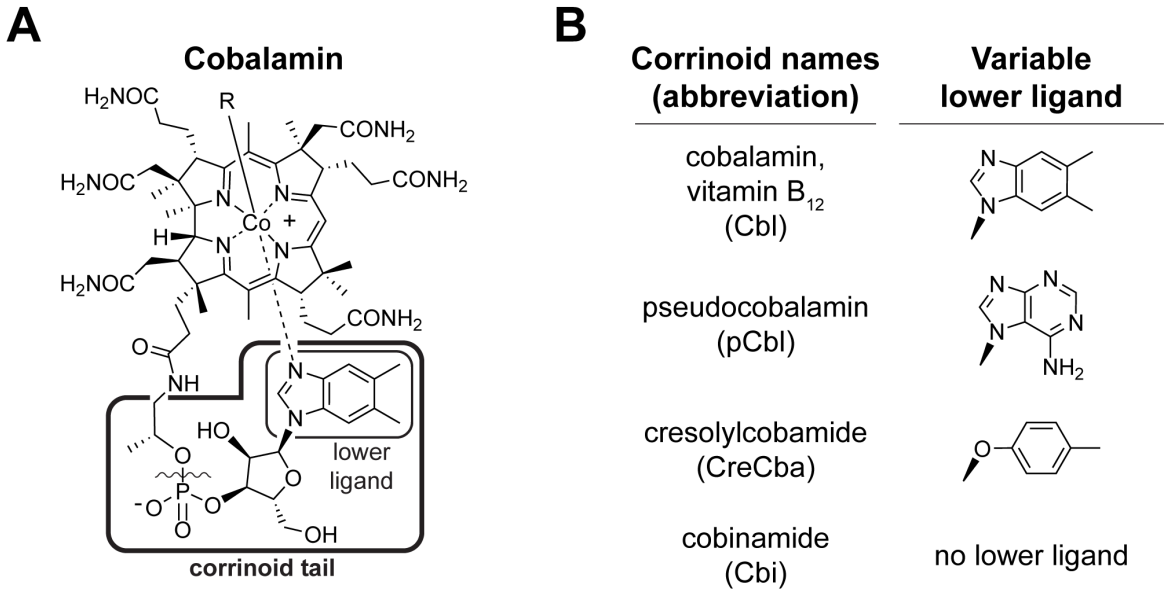


Figure 2.1 – Main corrinoid tail variants examined in this study. (A) Chemical structure of cobalamin, also called vitamin B_{12}. The thick-lined box delineates the corrinoid tail region, which contains the variable lower ligand group in the thin-lined box. The R-group indicates the corrinoid upper ligand which can be 5'-deoxyadenosyl, methyl, hydroxyl or cyanyl moieties. (B) Names and lower ligand structures of the corrinoids used throughout this study. Cobalamin (Cbl), pseudocobalamin (pCbl) and cresolylcobamide (CreCba) are 'complete corrinoids' (cobamides). Cobinamide (Cbi) is an 'incomplete corrinoid' that lacks the phosphoribosyl and lower ligand groups as indicated by the wavy line in the corrinoid tail in panel A.

Corrinoid structural diversity is often overlooked in Cbl-riboswitch studies, likely due to the intricacies of the microbial corrinoid biosynthesis pathway, lability of many corrinoid intermediates, and limited commercial availability of most corrinoids [123]. Bacterial corrinoid biosynthetic pathways produce a variety of end products known as cobamides (or 'complete

corrinoids'), which are fully functional enzyme cofactors (Figure 2.1B) [126]. Early steps of corrinoid biosynthesis produce a corrin ring containing a central cobalt ion and several methyl and propionamide substitutions. In later steps, an 'upper' (β) axial ligand is attached to the cobalt ion by a coordination bond, and a 'tail' structure is extended from the corrin ring. The structures of complete corrinoids vary at the upper axial ligand and at the terminal end of the tail structure, which often functions as the 'lower' or α axial ligand to the cobalt ion) (Figure 2.1). The upper axial ligand varies among three naturally occurring moieties: 5'-deoxyadenosyl, methyl, or hydroxyl, with each of these groups mediating distinct catalytic mechanisms, while the vitamin form contains a cyanyl group at this position [123]. Corrinoid tails vary among at least ten confirmed naturally occurring benzimidazolyl and purinyl base moieties which can form a coordinate bond with the lower axial face of the cobalt ion. Additionally, some bacteria produce corrinoids with a phenolic tail moiety incapable of forming a coordinate bond with the cobalt ion. The combination of reported upper axial ligands and tail structures suggests there exist at least 36 distinct complete corrinoid cofactors naturally produced by microbes, in addition to dozens of corrinoid pathway intermediates [125, 131, 136, 141, 228-230].

It is well established that organisms favor particular corrinoid tail variants through specific corrinoid import, biosynthesis, remodeling, and enzyme-corrinoid compatibility [143, 147, 149, 160, 231-233]. Yet the extent of corrinoid tail-specific gene regulation remains unclear, because nearly all studies of Cbl-riboswitches have focused solely on Cbl as an effector. It is therefore unclear to what extent Cbl-riboswitches are capable of distinguishing among corrinoids with different tail structures. In most cases, high effector selectivity is a key functional feature of a riboswitch. For example, *S*-adenosylmethionine (SAM) riboswitches strongly discriminate SAM from the chemical analog *S*-adenosylhomocysteine (SAH) by specifically detecting the charge state of the sulfur moiety, which differs between SAM and SAH [106]. Finding strongly selective Cbl-riboswitches that respond only to a single naturally occurring analog would be consistent with most models of riboswitch-based gene regulation. On the other hand, corrinoid cross-feeding is widespread amongst microbes [127, 155, 232-235] and enzymes are often compatible with multiple corrinoid tail analogs [139, 140, 146, 148, 231], so perhaps promiscuous sensitivity to corrinoids is a more advantageous gene regulatory strategy. Alternatively, a semi-selective corrinoid response in which a Cbl-riboswitch responds to some but not all corrinoids may balance the benefits of these two extremes. In any case, investigating how Cbl-riboswitches respond to other corrinoid tail variants will yield new insight into how microbes monitor and use these chemically diverse cofactors [127].

A few key studies have examined functional diversity among Cbl-riboswitches. Cbl-riboswitches can distinguish between the variable upper axial ligands of Cbl. The molecular basis of upper axial ligand selectivity, particularly the discrimination between the bulky 5'-deoxyadenosyl and smaller hydroxyl and methyl groups has been well characterized, although tail variants were not examined in these studies [101, 236, 237]. In the only study to our knowledge of corrinoid tail specificity of Cbl-riboswitches, *in vitro* binding measurements showed that the aptamer of the *Escherichia coli btuB* Cbl-riboswitch binds the complete corrinoids Cbl and 2-methyladeninylcobamide ([2-MeAde]Cba) with a 3.2-fold difference in affinity (K_D = 89 and 290 nM, respectively). Furthermore, cobinamide (Cbi), an incomplete corrinoid with a truncated tail, binds the aptamer

with roughly 8000-fold lower affinity than Cbl (K_D = 753 μM), suggesting that this aptamer binds corrinoids in a selective manner [163]. In a different study, it was shown that Cbl-riboswitches in *Desulfitobacterium hafniense* Y51 have distinct binding affinities for Cbl with K_D measurements ranging from 27 nM to 90 μM [150]. A bioinformatic analysis and comparison across several classes of riboswitches indicates that while Cbl-riboswitches retain conserved secondary structures, most of the stem-loop structures that comprise the Cbl-riboswitch aptamer are highly variable in sequence and length [64]. Collectively, these studies lead us to speculate that the sequence diversity among Cbl-riboswitches may harbor a range of distinct corrinoid tail-specific activities.

In this work, we examined how a panel of 38 Cbl-riboswitches derived from 12 bacterial species responds to the distinct tail structures of four corrinoids: Cbl, pseudocbobalamin (pCbl), cresolylcobamide (CreCba), and Cbi, representative of the benzimidazolyl, purinyl, phenolyl and incomplete corrinoids, respectively. To compare activities between several dozen Cbl-riboswitches, we devised a live cell fluorescence-based reporter system in *Bacillus subtilis*. In contrast to conventional *in vitro* biochemical approaches, this riboswitch reporter system captures the complete corrinoid-responsive gene regulatory process and provides rapid functional measurements with multiple effectors in parallel. Our results obtained from experiments in the reporter system in conjunction with comparative structural analyses of Cbl-riboswitches and corrinoid ligands allowed us to develop a mechanistic model for how corrinoid tail-specific gene regulation is achieved. Additionally, we examine a gene regulatory strategy for the corrinoid specificity of a Cbl-riboswitch and discuss the conceptual and practical implications of these findings.

2.3 Results

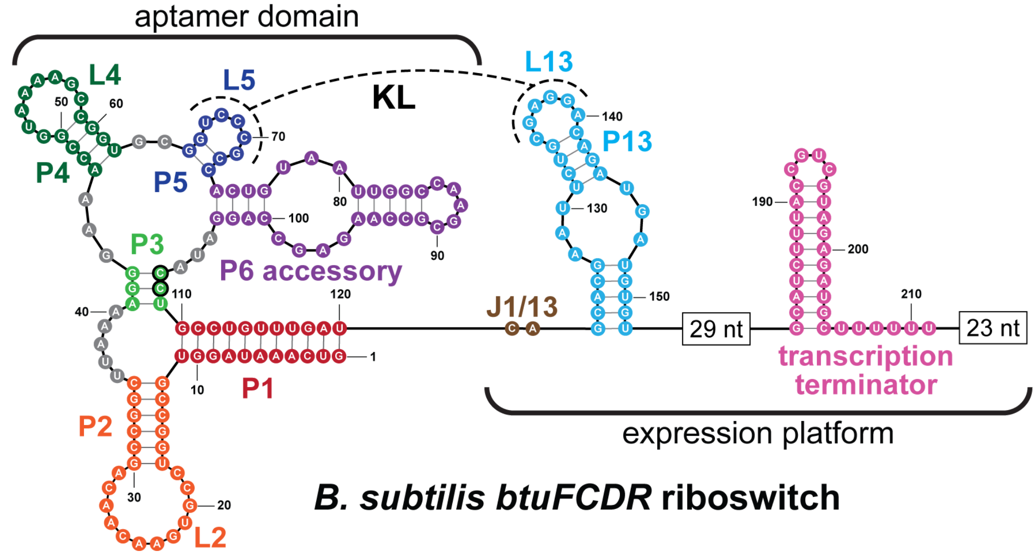
2.3.1 Development of a Cbl-riboswitch reporter system

We developed a GFP-based *in vivo* reporter system to measure and compare the responses of multiple riboswitches to a panel of corrinoids. We first engineered a *B. subtilis* strain for optimal corrinoid uptake and measurements of Cbl-riboswitch activity. The *B. subtilis* genome lacks the cobalamin biosynthesis pathway genes, but contains a putative operon, *btuFCDR*, which encodes homologs of the corrinoid uptake system BtuFCD [238, 239] and the adenosyltransferase BtuR, which installs the adenosyl upper ligand [240, 241]. The only Cbl-riboswitch in *B. subtilis* is located immediately upstream of *btuF* [102] (Figure 2.2A). We replaced the *btuF* promoter and riboswitch with the constitutive synthetic promoter P_{Veg} to ensure that the GFP reporter system is not limited by quantity or specificity of corrinoid import. As expected, constitutive expression of *btuFCDR* increases import of Cbl, pCbl, CreCba and Cbi compared to wild type expression, especially at the highest levels of corrinoid added to the media (2,500 pmol at a concentration of 100 nM) (Figure 2.3A-D). In fact, of 2,500 pmol of corrinoid provided in cell cultures, we recovered 45-65 pmol from the wild type strain and 900-1,900 pmol from the P_{Veg}-*btuFCDR* strain, which is roughly a 20- to 30-fold increase in corrinoid uptake. Thus, the P_{Veg}-*btuFCDR* strain provides a means of testing riboswitch function by addition of different corrinoids over a range of

concentrations up to 100 nM. We also deleted the gene *queG,* which encodes epoxyqueuosine reductase, the only annotated corrinoid-dependent enzyme in *B. subtilis,* to eliminate any interference from corrinoid-binding proteins [57, 65, 155, 166]. In this strain background, fluorescent reporters that contain a Cbl-riboswitch upstream of the *gfp* sequence were integrated into the *B. subtilis* chromosome.

We first validated the functionality of the reporter system by testing the Cbl dose-response of the native Cbl-riboswitch of *B. subtilis*, located upstream of *btuF*, in the strain described above (Figure 2.2A). Addition of CNCbl (the vitamin form of Cbl) at increasing concentrations produced a saturable response of 4.5-fold repression of GFP expression, indicating that, like other characterized Cbl-riboswitches [101, 122, 150, 242, 243], this riboswitch negatively regulates gene expression in response to Cbl (Figure 2.3E). Mutation of two highly conserved residues in the Cbl binding site of the aptamer domain eliminated Cbl-dependent repression of GFP expression, indicating that the response is dependent on a functional aptamer domain (Figures 2.2A, 2.3F). Compared to the strain that overexpresses *btuFCDR*, the Cbl-riboswitch GFP reporter in a wild type strain background displays an attenuated response to Cbl at concentrations above 1 nM, saturating to 2.5-fold repression. (Figure 2.3G). This is consistent with limited corrinoid uptake in wild type *B. subtilis* as shown in Figure 2.3A-D. Notably, deletion of *btuR* rendered this riboswitch reporter insensitive to exogenously supplied CNCbl, MeCbl, and OHCbl, while retaining dose-dependent repression in response to AdoCbl (Figure 2.4). This strongly suggests that the *B. subtilis btuFCDR* Cbl-riboswitch only responds to Cbl containing the 5'-deoxyadenosine upper ligand, in contrast with a report suggesting that this riboswitch aptamer can also bind MeCbl and OHCbl [244]. Based on the results thus far, we chose the P_{Veg}-*btuFCDR* Δ*queG* strain background for all subsequent experiments with riboswitch-GFP reporters examining corrinoid tail specificity of Cbl-riboswitches.

A



B

```
              P1            P2, L2                          P3      P4, L4
2° str  ((((((((((((.((((((((............)))))))).).....(((...(.((((.......)))).)
4GMA    AAGCCUUAUGGU---CGCUACC-AUUGCACUCC-GGUAGCG---UUAAAAGGGAAGACGGGUGAGAAUCCCG-C
4GXY    ---GCGGCAGGUGCUCCCGACC----CUGC----GGUCGGGA-GUUAAAAGGGAAG-CCGGUGCAAGUCCGG-C
B.sub   -GUCAAAUAGGU----GCCGGUCCGUGAACAACAGCCGGC----UUAAAAGGGAAA-CCGGUAAAAG-CCGG-U
B.meg   -AUGCAUCAGGU--GCUGUGUA--ACGAAUG---UACGCAGC--UUAAAAGGGAAGACCGGUGAAAAGCCGG-C
D.haf   -UAUUGACAGGU------UUAC----GUAA----GUAA------UGAAAAGGGAAU-CAGGUGCAAAUCCUG-A
S.ova   ----GGUUAGGU------GCCU----UUAA----GGGC------UUAAUAGGGAAGACCGGUGUAAUACCGG-C
V.par   ----GAAUAGGU-------GUC----GUAA----GAC-------UUAAUAGAGAAA-UCGGUAUAAG-CCGA-U

          P5, L5                 P6 accessory
2° str  ..(([[[[[[[))(((((..(((((........((((((((((......((((.((((((((......
4GMA    GCAGCCCCCGCUACUGUGA-GGGA-GGACGAA-----GCCCU-AGUAA-GCCACUGCC----GAAA-
4GXY    ACGG-UCCCGCCACUGUGA-CGGG------GAGUCGCCCCUCGGGAUGUGCCACUGGCC---CGAA-
B.sub   GCGG-UCCCGCCACUGUAAUUGGC-----------------------------------CAAGC
B.meg   GCGG-UCCCGCCACUGUAA-UGGG------GAGCAACUUAUU----AA--CCACUGUUUCA--UGU-
D.haf   GCAACCCCCGUUACUGUAAGCGCCGUUAAAGA-UUUAUUUAU-CUUCAUGCCACUGG----CGAAGA
S.ova   GCGG-UCCCGCCACUGUAA--UGG--GGAGUA----ACCCAA-GUUAU-ACCACUGAGAU--GAGA-
V.par   GCGG-UCCCGCCACUGUAA-GGAU------GA-GCCUGCUUUCAUAUAUGUCACUGGA----GUAA-

                       P6 accessory                          P3      P1
2° str  ))))))))...)))))))))).))))).........)))))...).))))...))))))))))))))
4GMA    ---GGUGGGAAGGC-AGGGU------GGAGGAUGAGUCCC--GAGCCAGGAGACCUGCCAUAAGGUUU
4GXY    --GGCCGGGAAGGCGGAGGG-GCGGCGAGGA-----UCCG--GAGUCAGGAAACCUGCCUGCCGC---
B.sub   ------------------------------------GCCAA-GAGCCAGGAUACCUGCCUGUUUGAU-
B.meg   UGAAAUGGGAAGG--GAUAA-GAAGCGAUGAA----CCUA---AGUCAGGAGACCUGCCUGUUGUAU-
D.haf   ----CUGGGAAGGCGAUAAA-UAAA-----------GGCGC-GAGCCAGGAGACCUGCCUGUUAAUA-
S.ova   -AUCUUGGGAAGGU-UUGGG-U----GAACAAUGAACCA---GAGCCAGGAGAACUGCCUAACU----
V.par   ---UCUGGGAAGGCGAAAGCGAGGU-GAUG------AUCC-CGAGCCAGGAGAACUGCCUAUUU----

          J1/13                       P13, L13
2° str  ..........................................(((([[[[[[[..)))).............
4GMA    UAGAA---------------------------GUUCGCCUUCGGGGGG-AAGGUGAAC---------
4GXY    GCU-----------------GGCGC-UG-CUC-G-ACCCUCGGGUGAAAGGGU-GAG-GCGUU---
B.sub   CA------------------------GCACG-AAU-UCUGCGAGGA--CAGA-UGA-UGUGU----
B.meg   G-----------------------CACUU-AAG--CUCUACGGGAUA-UAGAG-AG-GAGUG----
D.haf   AAACUACAAU----------------------UCAUUUUCCGGAGGG-GAGGGUGA----------
S.ova   GAUAAUCACCGUUUG----------------UU-GA-CCCACGAGCGA-UGGG-GA-GG--------
V.par   GA-------------UAAUC-A-CC-G-UU-AUGUA-CCUACGAGCGA-UAGG-AA-G-GG-GAUUG

                   transcription terminator
2° str       -(((((((((((((.....)))))))))))))............-
4GMA         ------------------------------------------
4GXY    37nt -----CCCUCGGGC-UGA-GCCCGGGGG----UUUUCUUU----- 49nt
B.sub   29nt ----GCAUCUUUACCGUC-GUAGAGAUGC---UUUUUU------- 23nt
B.meg   56nt ---UCUCUCUUUUCAUUUCGAAAAGAGAGA--UUUUUU------- 36nt
D.haf    0nt ----------------------------UUUAUUUUCUUU- 86nt
S.ova   82nt -UAAAACUCCCGUUAAUG-AGCGGGAGUUUUAUUUUU-------- 107nt
V.par   51nt -----CGUGUCACU-UUA-GGUGGCACG----UUUUUUAUU---- 178nt
```

Figure 2.2 – Cbl-riboswitch secondary structures. (A) Secondary structural model of the Cbl-riboswitch upstream of *B. subtilis btuFCDR*. Aptamer domain contains residues 1-120 and expression platform contains residues 121-211. Base paired stems, loops, and junctions P1, P2, L2, P3, P4, L4, P5, L5, J1/13, P13, and L13 are labeled. The P6 accessory region is highly variable in sequence length and number of stems across Cbl-riboswitch sequences, often containing up to six paired regions. In this example, P6 accessory consists of only two paired regions. A kissing loop interaction (KL, dashed line) occurs between L5 of the aptamer and L13 of the expression platform. Residues C107 and C108, which are mutated in Figure 2.3F, are outlined in black. (B) Secondary structural alignments of seven representative Cbl-riboswitches examined in this study. Row "2° str" indicates paired bases as parentheses, loops and junctions as periods, and kissing loop base pairs by brackets. PDB numbers or abbreviations are given for riboswitches from the following organisms and genes: 4GMA, *Thermoanaerobacter tengcongensis*; 4GXY, *Symbiobacterium thermophilum cblST*; *B.sub*, *B. subtilis btuFCDR; B.meg*, *Bacillus megaterium metE*; *D.haf*, *Desulfitobacterium hafniense* DSY0087; *S.ova*, *Sporomusa ovata btuB2*; *V.par*, *Veillonella parvula mutAB*. Non-conserved sequences flanking the transcription terminator are by nucleotide sequence length (nt), but the actual sequences were omitted for clarity in panels A and B.

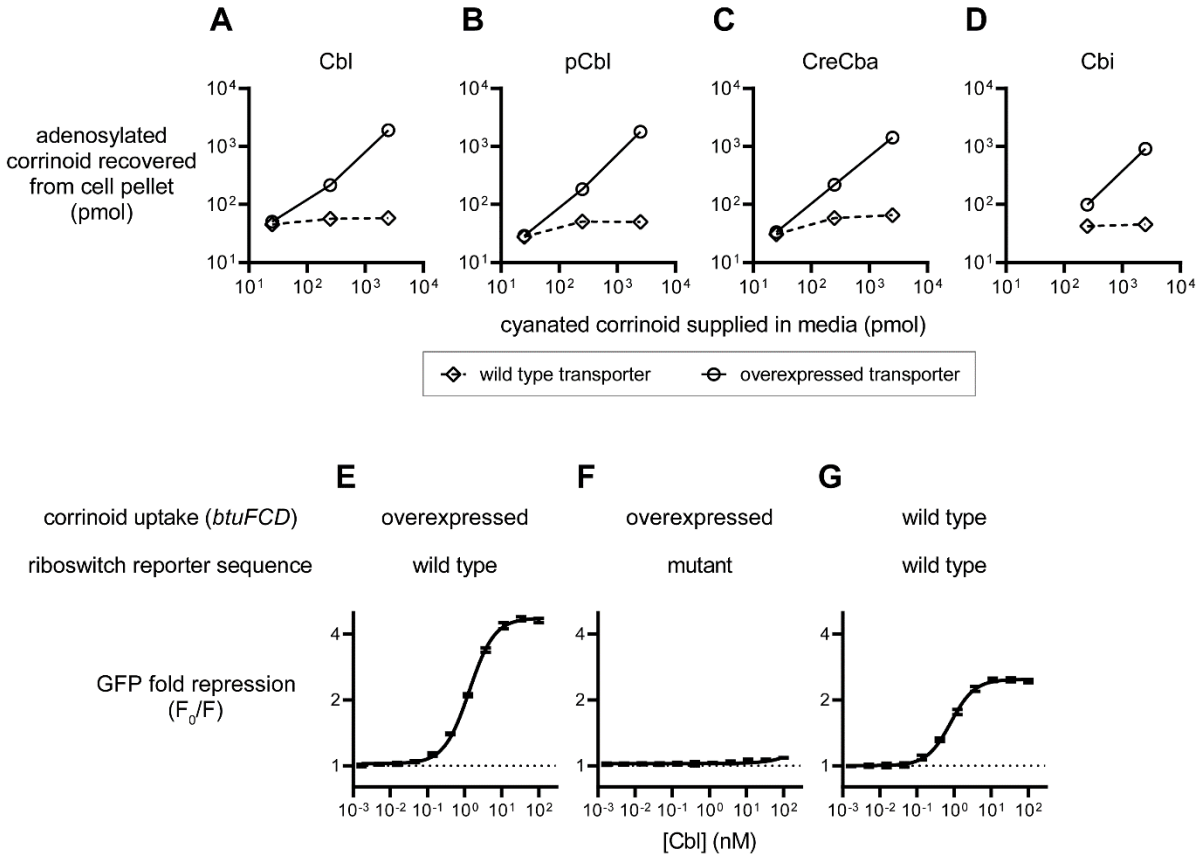


Figure 2.3 – Characterization of a live cell Cbl-riboswitch reporter system. (A-D) Intracellular accumulation of corrinoids in *Bacillus subtilis* strains expressing wild type or constitutively overexpressed levels of corrinoid uptake genes. Diamonds indicate the wild type *btuFCDR* strain, and circles indicate the constitutively overexpressed *btuFCDR* strain. (E-G) Dose responses of *B. subtilis btuF* riboswitch reporters. 'Mutant' riboswitch indicates the C107T,C108T mutant version of the *B. subtilis btuF* riboswitch was assayed. Data points in panels A-D represent single measurements from one representative experiment. Data points and error bars in panels E-G represent mean and standard deviation of four independent replicates, and horizontal dotted lines demarcate no change in expression.

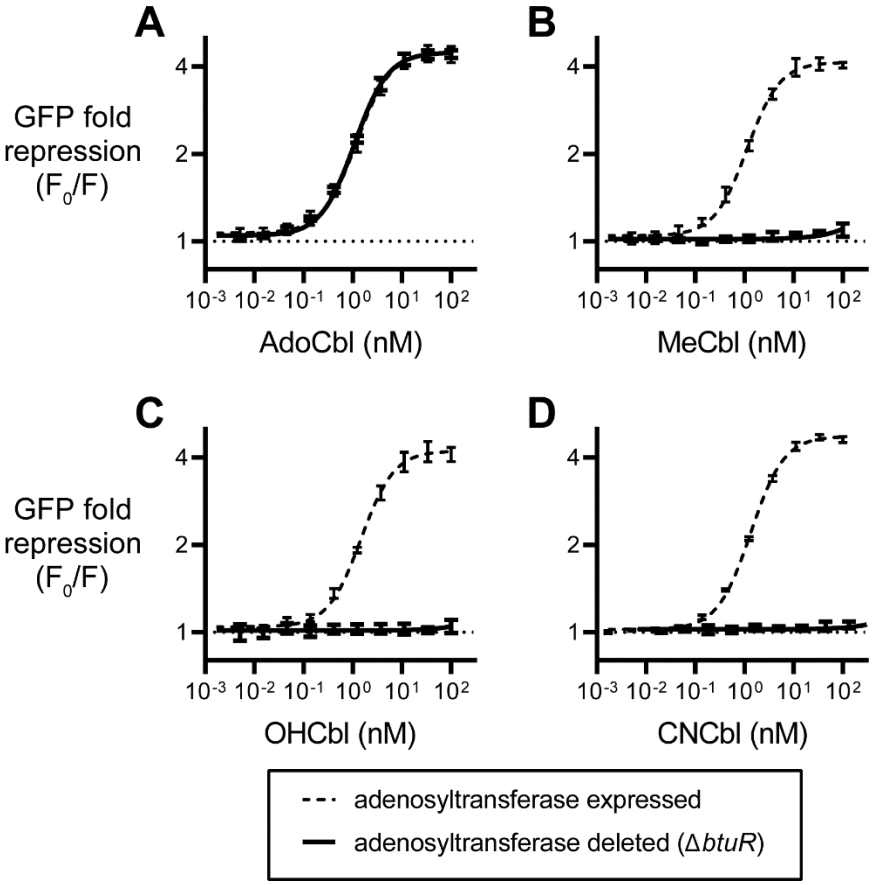


Figure 2.4 – The *B. subtilis btuFCDR* Cbl-riboswitch is sensitive to 5'-deoxyadenosyl, but not methyl, hydroxyl, or cyanyl upper ligand moieties of Cbl. Dose responses of *B. subtilis btuFCDR* Cbl-riboswitch reporter strains with (A) adenosylcobalamin (AdoCbl), (B) methylcobalamin (MeCbl), (C) hydroxocobalamin (OHCbl), and (D) cyanocobalamin (CNCbl). Data points and error bars represent mean and standard deviation of four independent replicates. Horizontal dotted lines demarcate no change in expression. Note that the two dose-response curves in panel A are overlapping.

2.3.2 Comparison of corrinoid specificity among Cbl-riboswitches

We constructed and tested 86 reporter strains to examine riboswitches from 20 bacterial species including 10 known to produce or require different corrinoids. When integrated into the *B. subtilis* chromosome, 38 of the 86 riboswitches showed 0.5-fold or greater repression of GFP expression in response to one or more corrinoids. We observed extensive variations in sequence length and nucleotide composition throughout the aptamer and expression platforms among the 38 riboswitches functional in *B. subtilis* (Figure 2.2B). Among these functional riboswitches are nine from *Bacillus megaterium*, which produces Cbl [125], and 12 from *Sporomusa ovata* and *Veillonella parvula* which both produce CreCba [134, 144, 229, 230]. These results show that the reporter system can be used to measure riboswitches of diverse sequence composition from organisms other than *B. subtilis*. We next measured the responses of the 38 functional Cbl-riboswitches to four corrinoids, Cbl, pCbl, CreCba and Cbi, to determine 1) whether Cbl-riboswitches are capable of responding to corrinoids other than Cbl, 2) whether they can distinguish between different corrinoids, and 3) whether corrinoid selectivity varies among Cbl-riboswitches. Our results demonstrate that all three are true; all of the riboswitches responded to more than one corrinoid, and a subset responded to all four (Figures 2.5-2.7). Strikingly, all of the tested riboswitches are either semi-selective or promiscuous (Figure 2.8A, B). We did not find any riboswitches that respond to only one corrinoid. The semi-selective and promiscuous riboswitches all respond to Cbl and pCbl (Figure 2.8C), and the promiscuous ones additionally respond to Cbi and CreCba (Figure 2.8D and E, points above the horizontal dashed line). Furthermore, the semi-selective riboswitches are generally more sensitive to Cbl than to pCbl, while the promiscuous riboswitches respond similarly to these two corrinoids. Almost all of the riboswitches respond weakly to CreCba compared to the other three corrinoids. In general, corrinoid selectivity appears to be associated with taxonomic origin (Figure 2.8C-E). The riboswitches from the Bacilli class are exclusively semi-selective, whereas ones from Negativicutes are predominantly promiscuous (Figures 2.5 and 2.6). The *S. ovata cobT* riboswitch is a notable exception that will be discussed later on. In contrast to taxonomy, a riboswitch's corrinoid selectivity is not strongly associated with the function of its regulatory target genes (Figure 2.9).

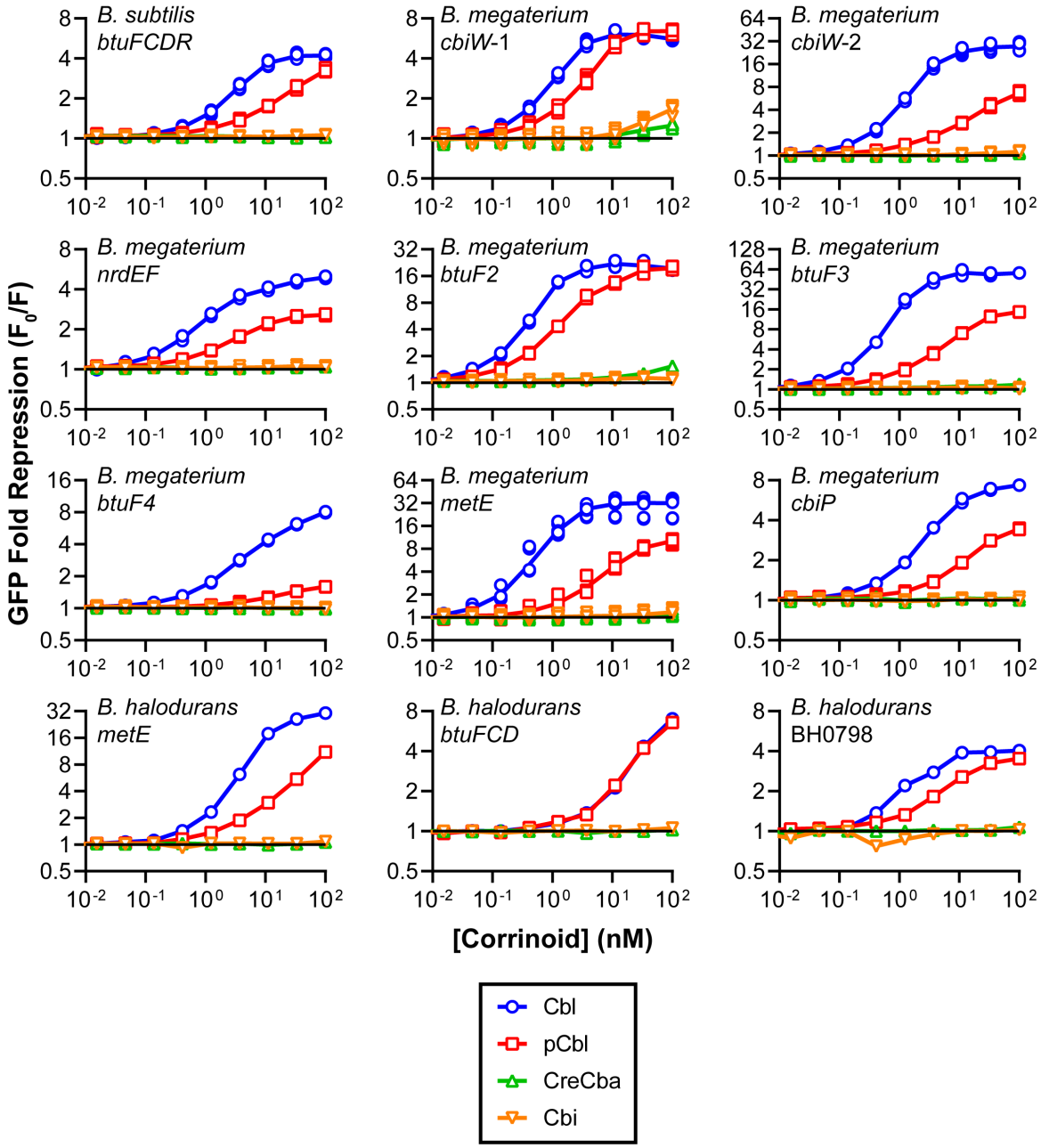


Figure 2.5 – Corrinoid dose-responses of riboswitch reporter strains with riboswitches from the class Bacilli. Species and operon indicated for each panel. Data points are plotted for at least two independent experiments for each strain. Lines connect mean values. *Bacillus subtilis*, *Bacillus megaterium* and *Bacillus halodurans* are members of Bacilli.

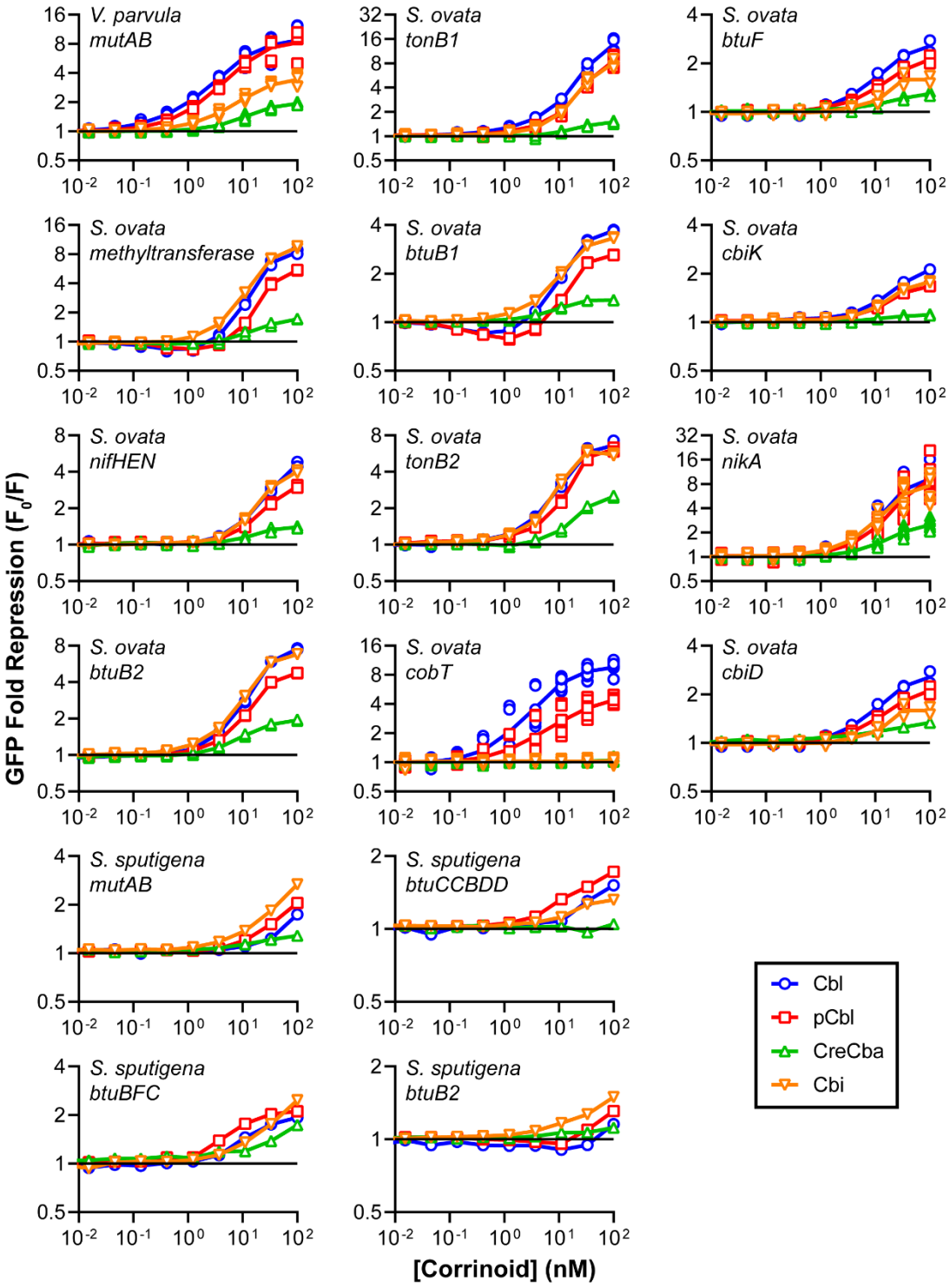


Figure 2.6 – Corrinoid dose-responses of riboswitch reporter strains with riboswitches from the class Negativicutes. Species and operon indicated for each panel. Data points are plotted for at least two independent experiments for each strain. Lines connect mean values. *Veillonella parvula*, *Sporomusa ovata* and *Selenomonas sputigena* are members of Negativicutes.

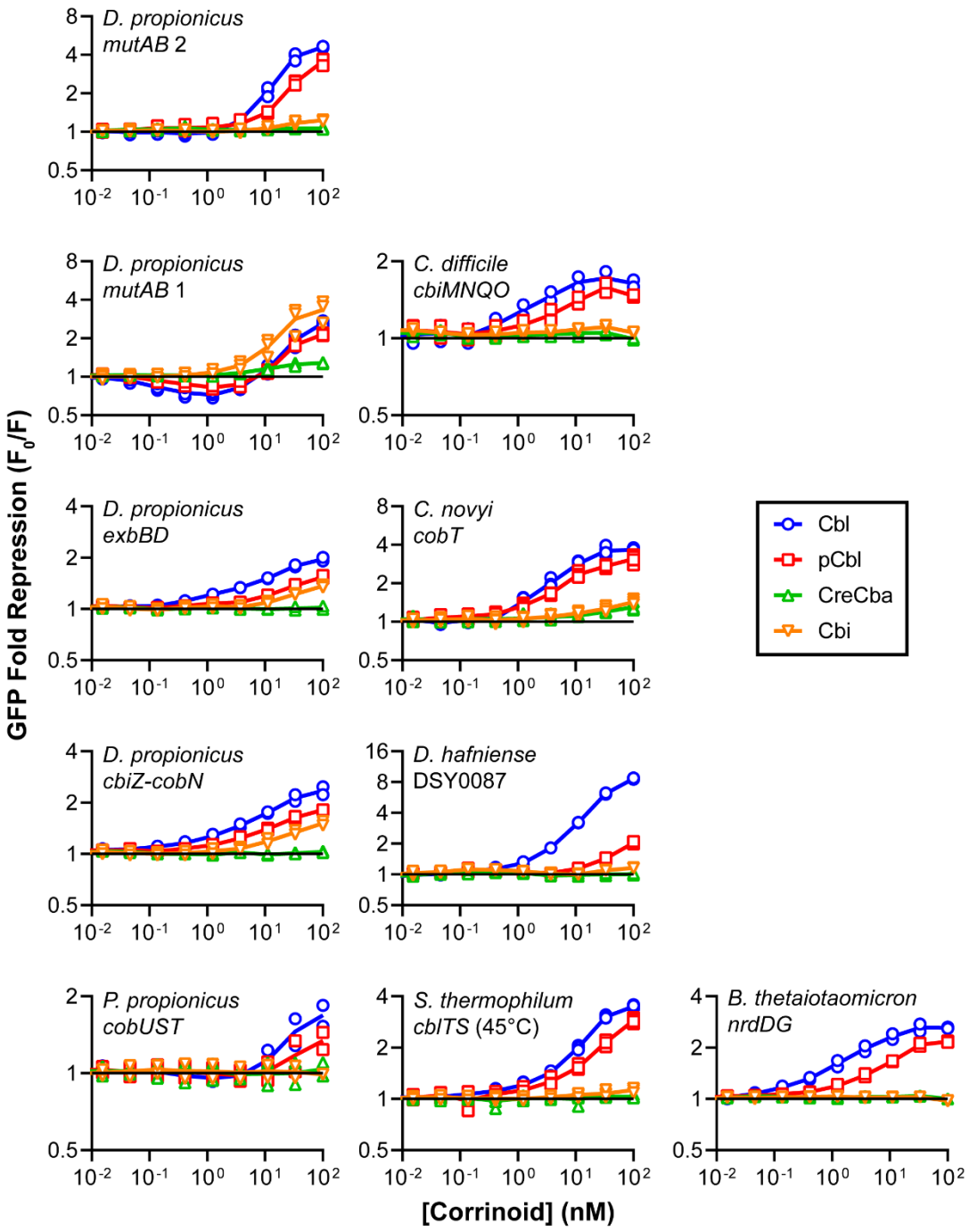


Figure 2.7 – Corrinoid dose-responses of riboswitch reporter strains with riboswitches from the classes Bacteroidia, Clostridia and Deltaproteobacteria. Species and operon indicated for each panel. Data points are plotted for at least two independent experiments for each strain. Lines connect mean values. *Bacteroides thetaiotaomicron* is a member of Bacteroidia. *Clostridium difficile*, *Clostridium novyi*, *Desulfitobacterium hafniense* and *Symbiobacterium thermophilum* are members of Clostridia. *Desulfobulbus propionicus* and *Pelobacter propionicus* are members of Deltaproteobacteria.

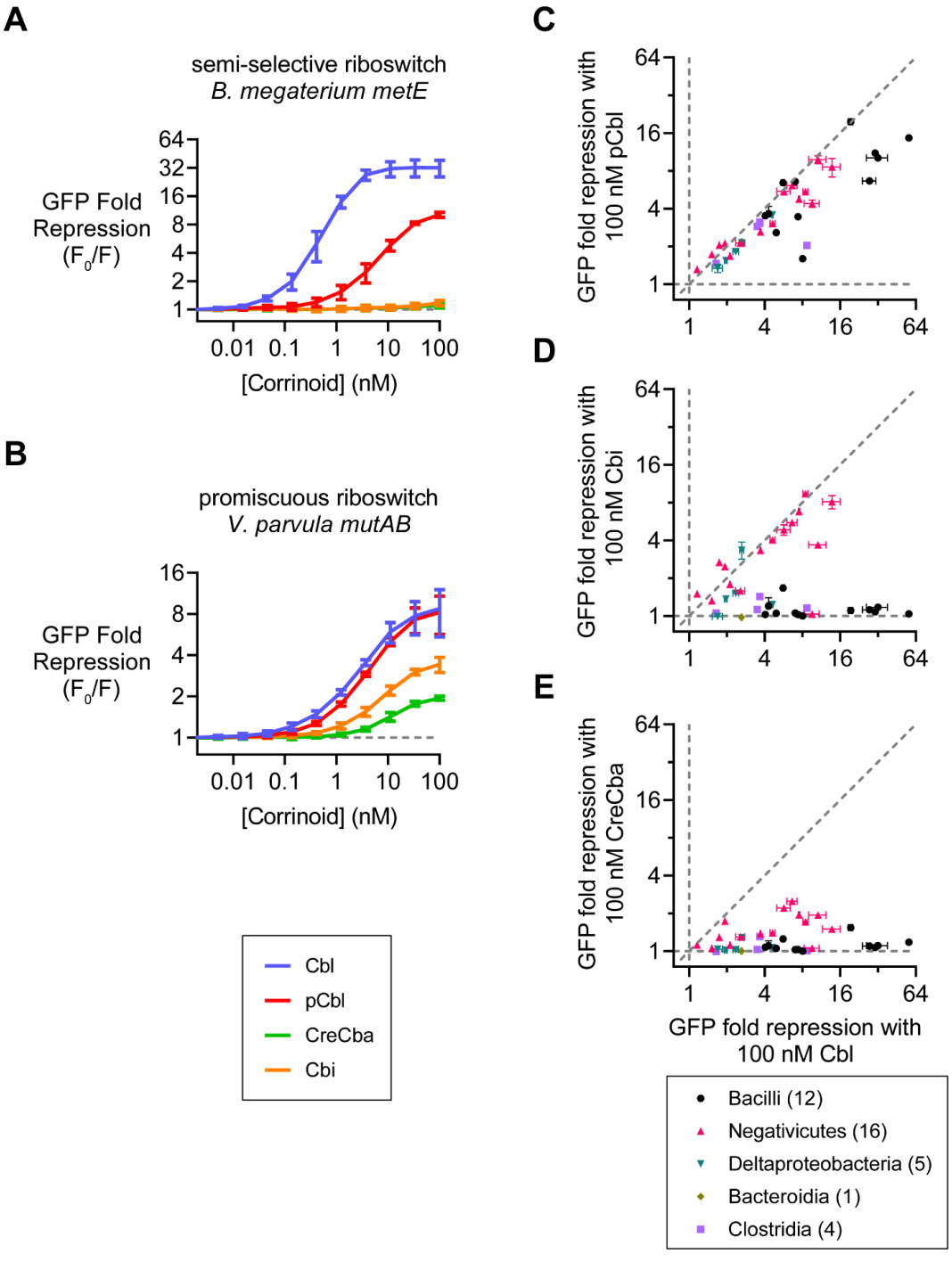

A
semi-selective riboswitch
B. megaterium metE
GFP Fold Repression (F_0/F)
64
32
16
8
4
2
1
0.01 0.1 1 10 100
[Corrinoid] (nM)
B
promiscuous riboswitch
V. parvula mutAB
GFP Fold Repression (F_0/F)
16
8
4
2
1
0.01 0.1 1 10 100
[Corrinoid] (nM)
Cbl
pCbl
CreCba
Cbi
C
GFP fold repression with 100 nM pCbl
D
GFP fold repression with 100 nM Cbi
E
GFP fold repression with 100 nM CreCba
1 4 16 64
GFP fold repression with 100 nM Cbl
Bacilli (12)
Negativicutes (16)
Deltaproteobacteria (5)
Bacteroidia (1)
Clostridia (4)

Figure 2.8 – Trends in corrinoid specificity among 38 Cbl-riboswitches. (A,B) Corrinoid dose-responses of two riboswitch reporter strains with distinct corrinoid specificities. The *B. megaterium metE* riboswitch is a semi-selective riboswitch example (A), whereas the *V. parvula mutAB* is an example of a promiscuous riboswitch (B). (C-E) Pairwise comparisons of GFP fold repression induced by 100 nM doses of Cbl versus (C) pCbl, (D) Cbi, and (E) CreCba. Data points are colored by taxonomic class with the number of riboswitches analyzed from each class indicated in parentheses. Vertical and horizontal gray dashed lines demarcate a response to only one of the two corrinoids. Diagonal line indicates equal repression between two corrinoids. The distance of a datapoint from the diagonal line indicates the bias in response towards one of the two corrinoids. Data points and error bars in panels A and B represent mean and standard deviation of ten experiments for *B. megaterium metE* and six experiments for *V. parvula mutAB*. Data points and error bars in panels C-E represent mean and range of at least two independent experiments for each riboswitch.

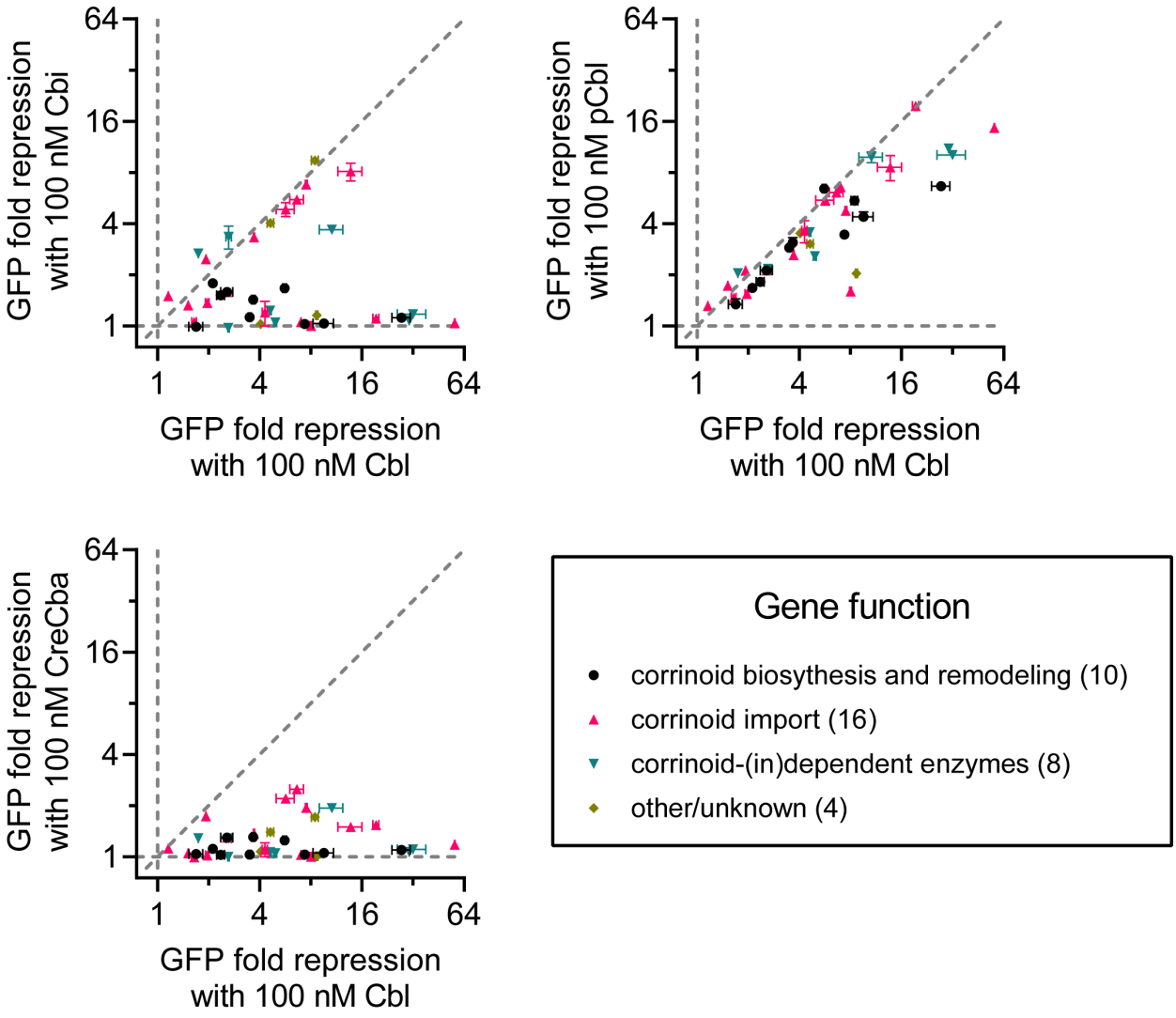


Figure 2.9 – Cbl-riboswitch specificity does not cluster by regulatory gene target function. Pairwise comparisons of GFP fold repression induced by 100 nM doses of Cbl versus Cbi (top left) pCbl (top right), and CreCba (bottom left). Data points are colored by predicted function of downstream regulatory target genes. The number of riboswitches analyzed from each group is indicated in parentheses. Vertical and horizontal gray dashed lines demarcate a lack of response to one corrinoid. Diagonal line indicates equal repression between corrinoids. The distance of a datapoint from the diagonal line indicates the bias in response towards one of the two corrinoids.

The observation that Cbl-riboswitches differ in their corrinoid selectivity led us to investigate the RNA sequence features underpinning this characteristic. We examined the influence of specific domain and subdomain sequences on corrinoid selectivity by constructing chimeric riboswitches containing parts of the semi-selective *B. megaterium metE* riboswitch and the promiscuous *V. parvula mutAB* riboswitch (Figure 2.10A). This allowed us to directly measure the effects of specific domain and subdomain sequences on corrinoid selectivity. A chimera containing the aptamer domain of the *B. megaterium metE* riboswitch fused to the expression platform of the *V. parvula mutAB* riboswitch retained the selectivity of the *B. megaterium metE* riboswitch, while the reciprocal chimera retained the selectivity of the *V. parvula mutAB* riboswitch (Figure 2.10A). This result indicates that the aptamer domain contributes strongly to the corrinoid selectivity. To pinpoint the specific regions of the aptamer domain that confer selectivity, we swapped the stem P1, stem-loop P2-L2, stem-loop P4-L4, and P6 accessory region between the two riboswitches to generate a set of aptamer subdomain chimeras. In the context of the *B. megaterium metE* riboswitch scaffold, replacing the stem-loop P2-L2 or the P6 accessory region with the corresponding structure in the *V. parvula mutAB* riboswitch resulted in a riboswitch that is less selective by gaining sensitivity to Cbi and CreCba. This suggests that these two subdomains contribute to corrinoid promiscuity (Figure 2.10B). In contrast, within the *V. parvula mutAB* riboswitch scaffold, swapping stem P1 increased corrinoid selectivity by retaining sensitivity to Cbl and pCbl, but losing sensitivity to CreCba and Cbi (Figure 2.10C). The remaining chimeric riboswitches partially or completely lost repressive activity (Figure 2.10B and C). These results demonstrate that subdomains distributed throughout the Cbl-riboswitch aptamer domain impact the corrinoid selectivity characteristic; no single conserved substructure completely controls corrinoid selectivity, nor did any single structure fully convert a riboswitch's corrinoid selectivity. Thus, the source of corrinoid selectivity within the riboswitch RNA sequence appears to be a complex phenotype requiring inputs from multiple regions of the RNA scaffold.

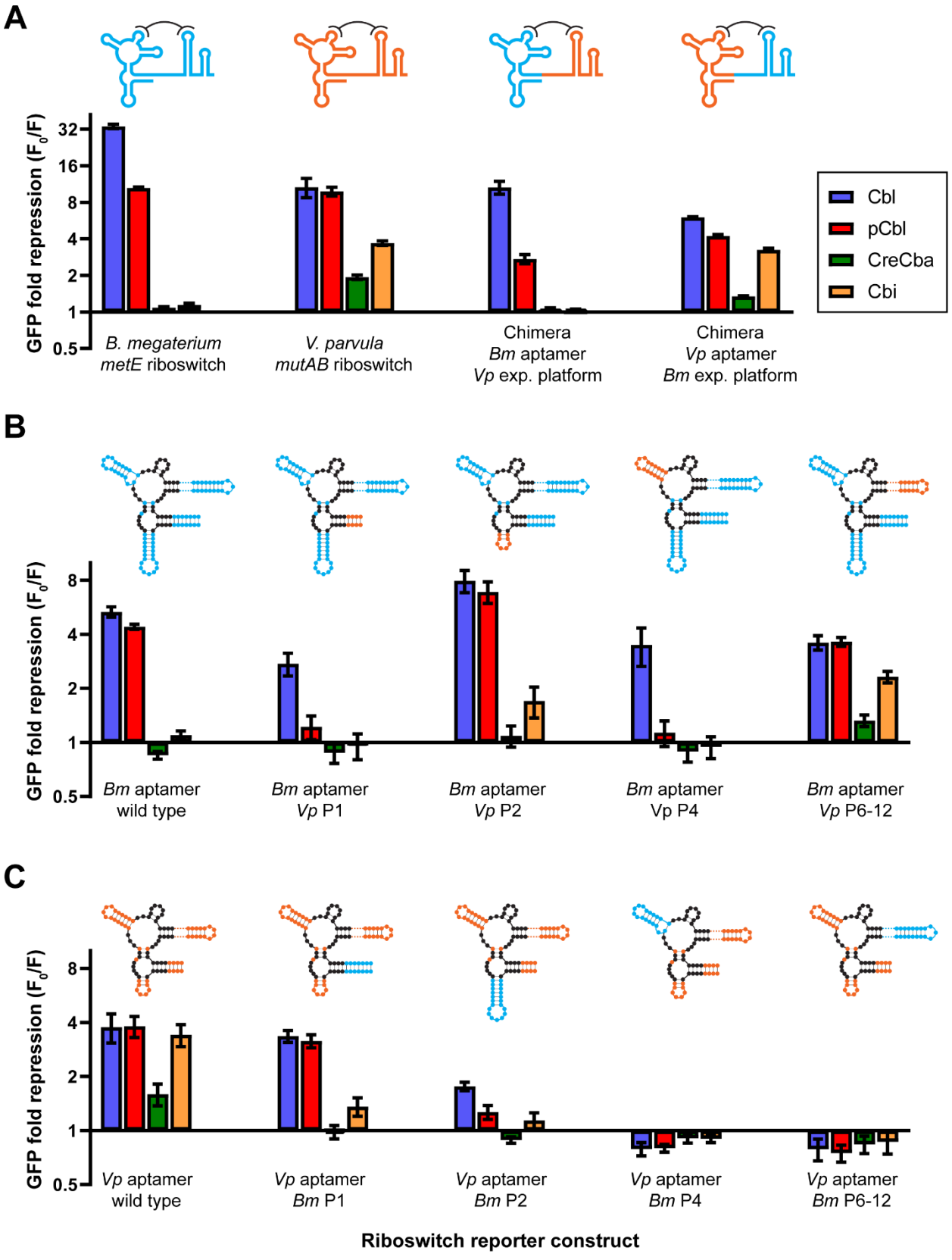

A
GFP fold repression (F0/F)
32
16
8
4
2
1
0.5
B. megaterium metE riboswitch
V. parvula mutAB riboswitch
Chimera Bm aptamer Vp exp. platform
Chimera Vp aptamer Bm exp. platform
Cbl
pCbl
CreCba
Cbi
B
GFP fold repression (F0/F)
8
4
2
1
0.5
Bm aptamer wild type
Bm aptamer Vp P1
Bm aptamer Vp P2
Bm aptamer Vp P4
Bm aptamer Vp P6-12
C
GFP fold repression (F0/F)
8
4
2
1
0.5
Vp aptamer wild type
Vp aptamer Bm P1
Vp aptamer Bm P2
Vp aptamer Bm P4
Vp aptamer Bm P6-12
Riboswitch reporter construct

Figure 2.10 – Chimeric riboswitches show that multiple components contribute to corrinoid specificity. Repression of GFP expression with 100 nM corrinoid is shown for (A) *B. megaterium metE* and *V. parvula mutAB* riboswitches and chimeric fusions of their aptamer and expression platform domains. Aptamer subdomains swaps within the (B) *B. megaterium metE* riboswitch scaffold and (C) *V. parvul*a *mutAB* riboswitch scaffold. Columns and error bars represent mean and standard deviation of 3 replicates. Cartoon riboswitches depict *B. megaterium* riboswitch sequence in cyan and *V. parvula* riboswitch sequence in orange. Sequences in black are identical between riboswitches.

2.3.3 Corrinoid tail structure impacts selectivity of Cbl-riboswitches

We next sought to identify how structural differences in the corrinoid tail affect the response of semi-selective Cbl-riboswitches to corrinoids. Curiously, there are no predicted hydrogen bond interactions between the lower ligand and the RNA in the X-ray crystal structures of Cbl-bound riboswitches, making it difficult to surmise how a Cbl-riboswitch might distinguish between corrinoids [122, 244, 245]. Could the overall structural conformation of the corrinoid, rather than specific interactions between the RNA and the corrinoid tail, influence Cbl-riboswitch activity? Most complete corrinoids exist in an equilibrium between two distinct isomeric states known as base-ON and base-OFF [137, 138]. In the base-ON form, a nitrogen atom in the lower ligand base is coordinated to the central cobalt atom of the corrin ring (as shown in Figure 2.1A). In the base-OFF form, the cobalt ion is not coordinated by the nitrogen atom, allowing the tail to move more freely. All six X-ray crystal structures of Cbl-riboswitches contain Cbl in the base-ON form. Biochemical evidence shows that the *E. coli btuB* riboswitch aptamer binds Cbl with a higher affinity than the primarily base-OFF corrinoids [2-MeAde]Cba and Cbi. Yet, [2-MeAde]Cba and Cbi were previously found to remain primarily base-OFF while bound to the aptamer [163]. Interestingly, we noticed that the semi-selective riboswitches respond strongly to Cbl and pCbl, which are capable of switching between base-ON and base-OFF forms, but respond weakly to Cbi or CreCba, which exist exclusively in the base-OFF form. Additionally, we observe that semi-selective Cbl-riboswitches are consistently more sensitive to Cbl, which has a greater base-ON constitution than pCbl (Figures 2.5-2.7). Based on these observations, we hypothesized that semi-selective riboswitches specifically detect the base-ON/OFF constitution of corrinoids. We therefore used a range of corrinoids with diverse lower ligand structures to test whether the activity of Cbl-riboswitches quantitatively correlates with the base-ON/OFF constitution of the corrinoid.

We selected a panel of sixteen adenosylated corrinoids for this analysis, including both natural and synthetic benzimidazolyl, purinyl, and aza-benzimidazolyl corrinoids that span a range of base-ON tendencies between that of Cbl and pCbl (Figure 2.11A). We observed a strong association between base-ON constitution – measured as the ratio of spectral absorbance at 525 and 458 nm (Figure 2.12) – and riboswitch response in semi-selective riboswitches (Figure 2.11B). Even among promiscuous Cbl-riboswitches, we observe measurable sensitivity to base-ON constitution, albeit to a much smaller degree (Figure 2.11C). These results support the hypothesis that Cbl-riboswitches differentially regulate downstream gene expression by sensing corrinoid base-ON constitution.

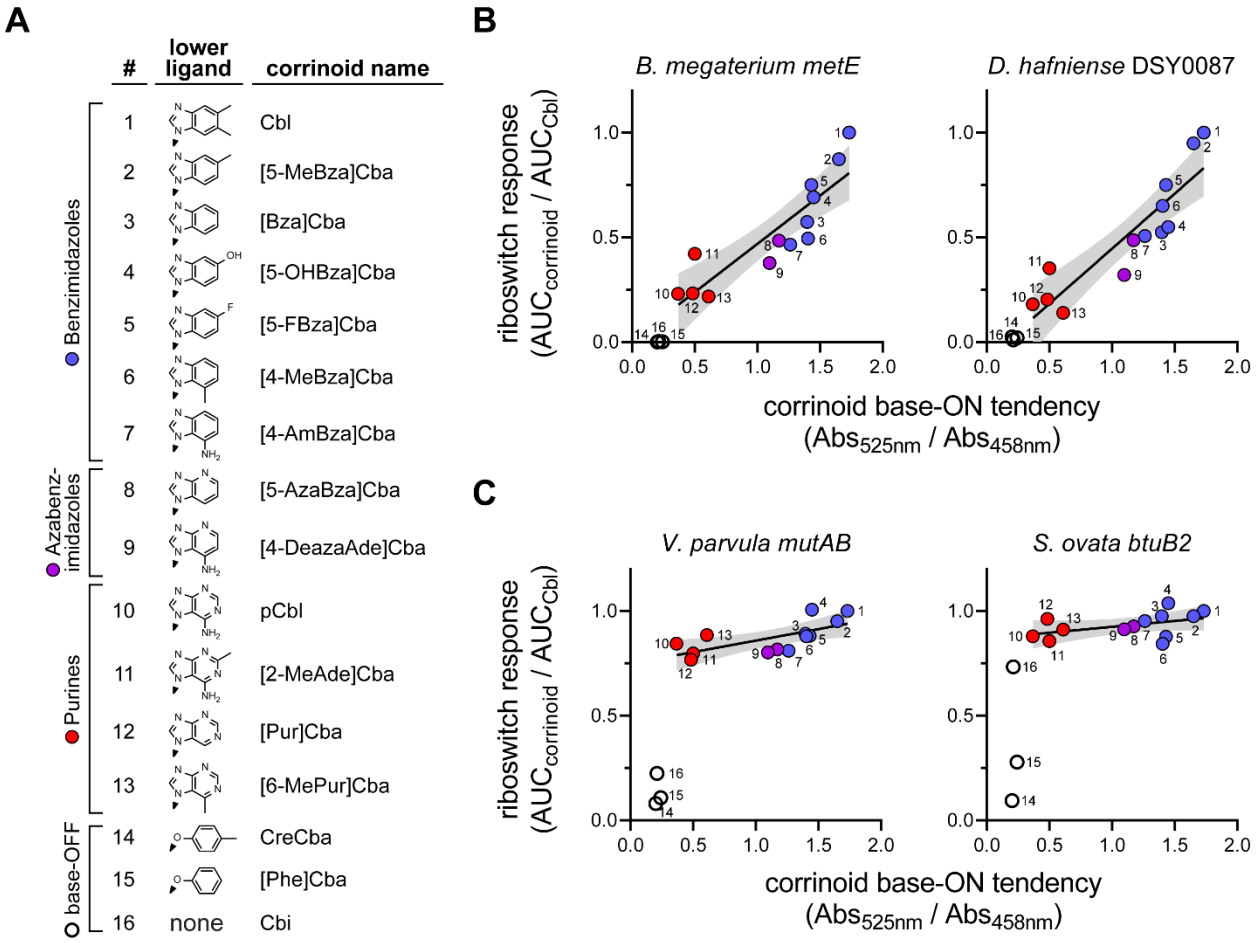


Figure 2.11 – Corrinoid base-ON tendency correlates with corrinoid potency in semi-selective riboswitches. (A) A panel of sixteen adenosylated corrinoids were used to test the relationship between corrinoid base-ON tendency and corrinoid potency. (B) Two semi-selective riboswitches *B. megaterium metE* and *D. hafniense* DSY0087 and (C) two promiscuous riboswitches *V. parvula mutAB* and *S. ovata btuB2* were used to measure each corrinoid's potency. Base-ON tendency was measured as the ratio between spectral absorbance of 525nm and 458 nm wavelengths in a pH 7.3 solution. The potency of a corrinoid to activate a riboswitch was measured as the area under the dose-response curve of the riboswitch reporter strain for that corrinoid ($AUC_{corrinoid}$) relative to its dose-response to Cbl (AUC_{Cbl}). Corrinoids with benzimidazole (blue circles), azabenzimidazole (purple circles), and purine (red circles) lower ligands can adopt the base-ON conformation. The corrinoids that are unable to adopt the base-ON form (CreCba, [Phe]Cba and Cbi) are represented by empty circles. Trendlines in B and C were fit to the data points of corrinoids with benzimidazole, azabenzimidazole and purine lower ligands. Strictly base-OFF corrinoids were excluded from trendline fits. Each data point represents a single measurement of riboswitch dose-response and corrinoid base-ON tendency.

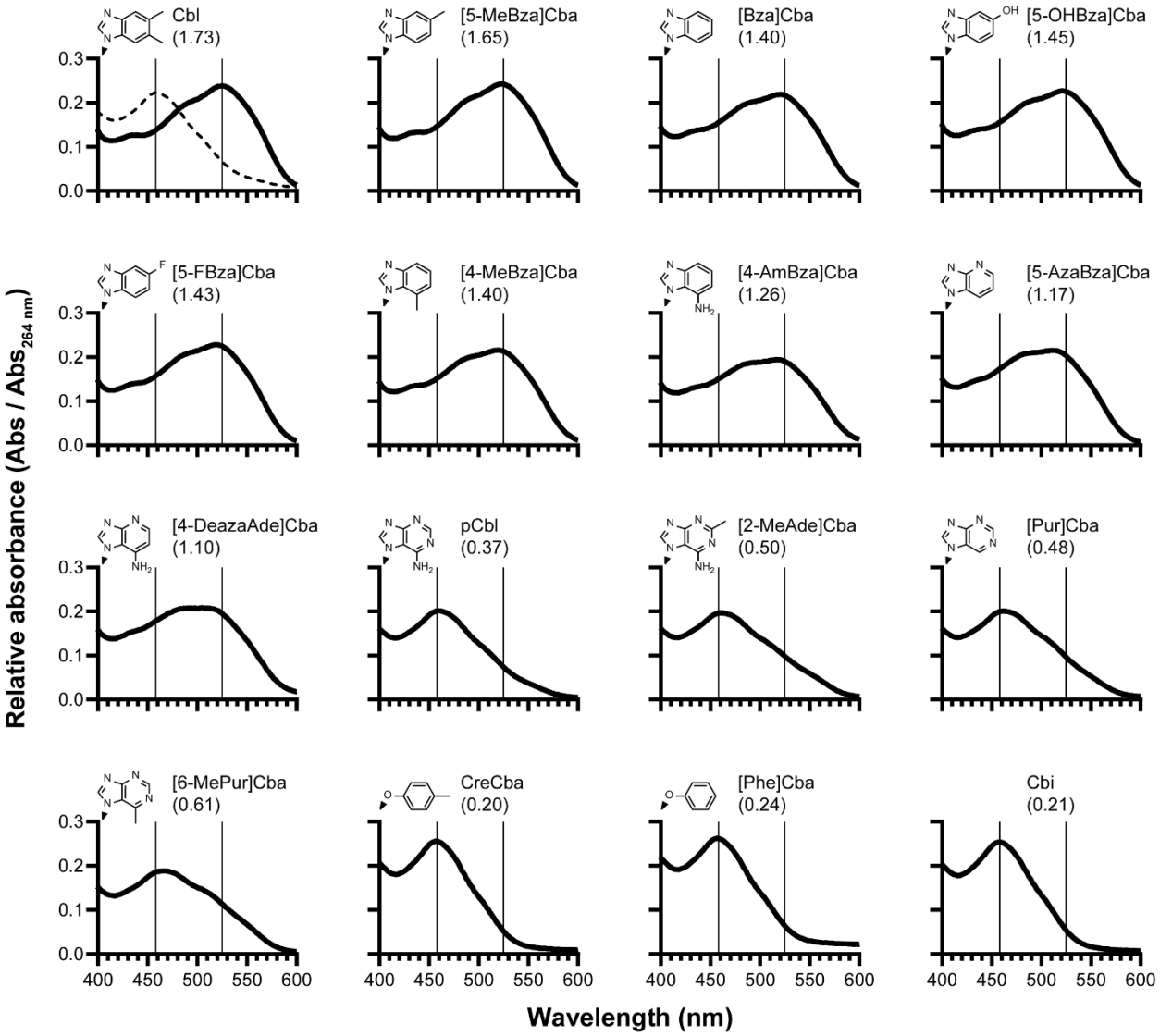


Figure 2.12 – Corrinoid lower ligand structure impacts base-ON tendency. Absorbance spectra of corrinoids were measured in a neutral buffered solution (pH 7.3). Absorbance at 458 nm and 525 nm are associated with base-OFF and base-ON conformations, respectively. Base-ON tendency was measured as the ratio between absorbance at 525 nm and 458 nm (thin vertical lines). Each panel is labeled with corrinoid name, base-ON tendency value (Abs_{525nm} / Abs_{458nm}), and lower ligand chemical structure. For reference, the dashed line in the upper left panel is the absorbance spectrum of nearly complete base-OFF Cbl measured in acidic solution (pH 1.57).

2.3.4 Structural comparisons between base-ON and base-OFF tail orientations

The results presented above led us to speculate about how a Cbl-riboswitch could detect the base-ON constitution of corrinoid. While corrinoid lower ligand bases such as adenine could conceivably be detected by specific base pair or base stacking interactions with RNA residues, we argue that it is unlikely because the Co-N coordinate bond of the base-ON form sequesters the Watson-Crick and Hoogsteen edges of the lower ligand away from the solvent. In all published X-ray crystal structures of Cbl-riboswitches, aptamer-effector binding is achieved mainly through Van Der Waals forces and shape complementarity between the binding site and base-ON Cbl. Only a few hydrogen bonds between the RNA and corrinoid are observed, none of which occur with the lower ligand group of Cbl [122, 244, 245]. One major difference between the base-ON and base-OFF forms of a corrinoid is the positional constraint placed on the corrinoid tail. Specifically, the tail in the base-ON form is spatially constrained due to the Co-N coordinate bond, whereas the tail in the base-OFF form is able to sample a wider range of spatial positions [246]. To develop mechanistic insight into how the base-ON/OFF constitution of a corrinoid could impact Cbl-riboswitch activity, we leveraged the plethora of publicly accessible X-ray crystal structures of macromolecule-bound Cbl, which have accumulated over decades of corrinoid research [57, 65, 122, 169, 172, 226, 244, 245, 247-252]. We first assessed the range of structural conformations that are potentially sampled by corrinoids as they dynamically switch between base-ON and base-OFF states by aligning and visually comparing various structural models of Cbl. Six base-ON Cbl models were obtained from structural studies of Cbl-riboswitches and synthetic Cbl RNA aptamers, whereas base-OFF/His-ON Cbl models were obtained from ten Cbl-dependent enzyme studies (Table 2.1). After aligning and superimposing these molecular models by their central cobalt and coordinating nitrogen atoms, we observed that the corrin rings and their amide and methyl substituents occupy similar spatial positions, but the tails of base-ON and base-OFF Cbl structures occupy distinct positions (Figure 2.13A and B). Moreover, the base-ON Cbl tails appear in very similar positions with lower ligands in close proximity to the central cobalt ion, whereas the base-OFF Cbl tails appear more scattered with lower ligands more distal to the cobalt ion. These structural alignments visually convey the degree to which the conformations of base-ON and base-OFF corrinoids can vary among biomolecular complexes.

Next, we compared the positions of the aligned base-ON and base-OFF Cbl models in the context of 3D Cbl-riboswitch models. We analyzed X-ray crystal structures of the two Cbl-riboswitches that contain resolved kissing loop structures: one from *Thermoanaerobacter tengcongensis* (Figure 2.13C and D) and one identified from a marine metagenome sequence (Figure 2.14A and B) [122]. These structural models show that the base-ON tails are contained within the binding site, whereas the tails of the base-OFF Cbl structures protrude away from the binding site and clash with the L5-L13 kissing loop. Although the *B. subtilis btuF* (Figure 2.14C and D) and *Symbiobacterium thermophilum cblTS* (Figure 2.14E and F) riboswitch models do not contain the L13 structure of the kissing loop, some of the modeled base-OFF tails clash with L5 of the aptamer in these structures [244, 245]. The kissing loop has been shown to play a key mechanistic role of sensing the corrinoid-binding state of the aptamer domain to influence downstream regulatory structures in the expression platform [120, 121]. If kissing loop formation in semi-selective Cbl-riboswitches is sensitive to the base-ON/OFF constitution of the corrinoid, then corrinoid selectivity may be

mediated by either selective binding by the aptamer or selective formation of downstream regulatory structures.

PDB ID	Bound macromolecule	Organism	Base-ON/OFF
4GMA	Cbl-riboswitch	*Thermanaerobacter tengcongensis*	Base-ON
4FRN	Cbl-riboswitch	Unknown (marine metagenome sequence)	Base-ON
4FRG	Cbl-riboswitch	Unknown (marine metagenome sequence)	Base-ON
4GXY	Cbl-riboswitch	*Symbiobacterium thermophilum*	Base-ON
6VMY	Cbl-riboswitch	*Bacillus subtilis subsp. subtilis* str. 168	Base-ON
1ET4	Synthetic RNA aptamer	Not applicable	Base-ON
4REQ	Methylmalonyl-CoA mutase	*Propionibacterium freudenreichii subsp. shermanii*	Base-OFF
2XIJ	Methylmalonyl-CoA mutase	*Homo sapiens*	Base-OFF
1K7Y	Methionine synthase	*Escherichia coli*	Base-OFF
5D08	Epoxyqueuosine reductase	*Bacillus subtilis subsp. subtilis* str. 168	Base-OFF
5D6S	Epoxyqueuosine reductase	*Streptococcus thermophilus*	Base-OFF
5C8D	CarH photoreceptor	*Thermus thermophilus* HB27	Base-OFF
4RAS	Reductive dehalogenase	*Nitratireductor pacificus* pht-3B	Base-OFF
5CJU	Isobutyrl-CoA mutase	*Cupriavidus metallidurans* CH34	Base-OFF
1I9C	Glutamate mutase	*Clostridium cochlearium*	Base-OFF
1XRS	D-lysine 5,6-aminomutase	*Acetoanaerobium sticklandii*	Base-OFF

Table 2.1 – X-ray crystal structural models used for 3D Cbl structural alignments.

We attempted to dissect corrinoid-selective binding from subsequent corrinoid-selective regulation. We tested for promiscuous binding by comparing the Cbl dose-response of the *B. megaterium metE* Cbl-riboswitch in the presence and absence of competing 100 nM Cbi and found that the response to Cbl is unaffected by Cbi (Figure 2.15A and B). This indicates that Cbi does not compete with Cbl for riboswitch binding, supporting corrinoid-selective binding as the mechanism of semi-selectivity. However, when the aptamer of this semi-selective riboswitch is replaced with the aptamer of the promiscuous *S. ovata nikA* riboswitch, it retains semi-selectivity, which suggests that the *B. megaterium* expression platform also plays a role in corrinoid selectivity (Figure 2.15A). Interestingly, the Cbl dose-response of the *S. ovata nikA* / *B. megaterium metE* chimeric riboswitch does become sensitized to competing Cbi addition, confirming that the *S. ovata nikA* aptamer retains sensitivity to base-OFF corrinoid in the context of this chimera (Figure 2.15C). Taken together, these show that base-OFF corrinoids may impede both Cbl-riboswitch binding and formation of regulatory structures, explaining the link between base-ON/OFF constitution and riboswitch activity observed in Figure 2.11.

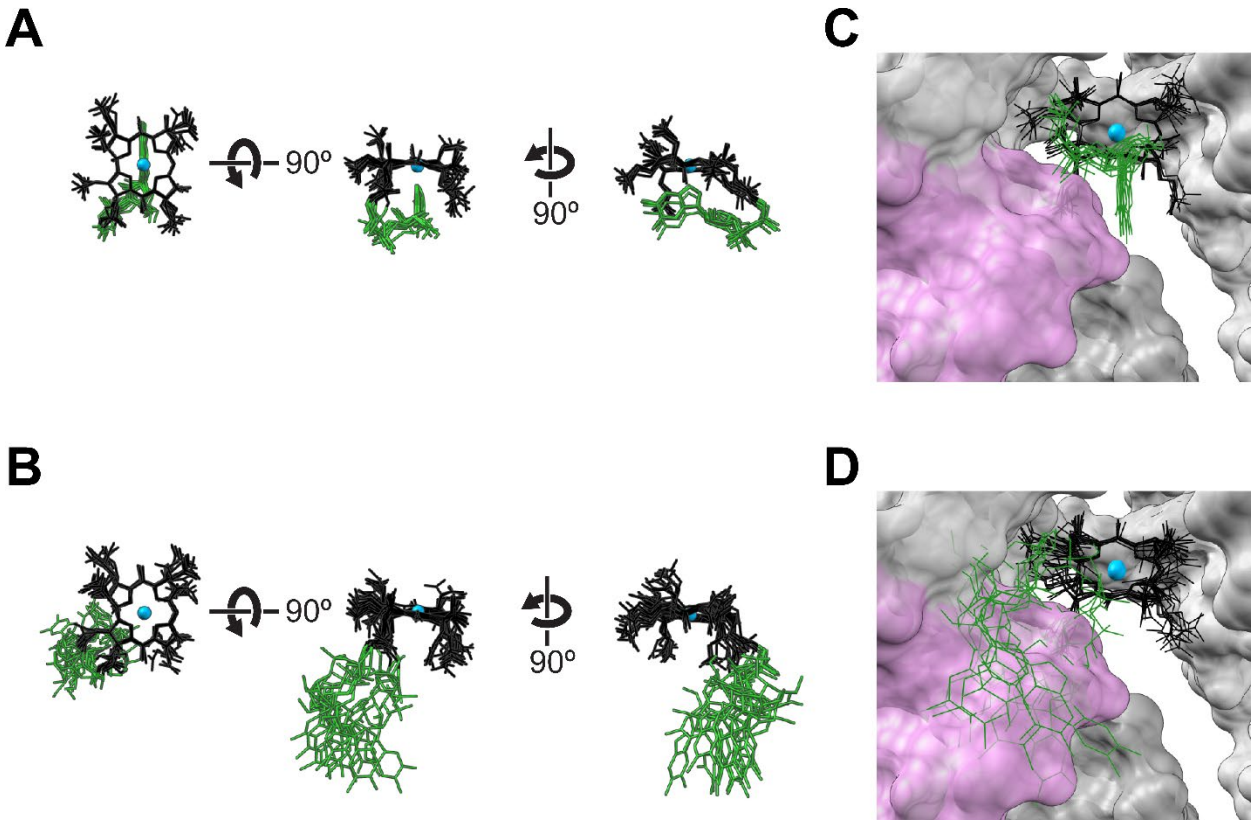


Figure 2.13 – Distinct tail positions among base-ON and base-OFF corrinoids may impact binding of corrinoids to riboswitches. 3D alignments of Cbl structural models derived from published X-ray crystal structures of (A) base-ON Cbl in complex with RNAs and (B) base-OFF Cbl in complex with proteins. The Cbl-binding site in the X-ray crystal structure of the *Thermoanaerobacter tengcongensis* Cbl-riboswitch (PDB ID 4GMA) is depicted with (C) base-ON and (D) base-OFF Cbl alignments. Cbl models were aligned by the cobalt and coordinating nitrogen atoms in the corrin ring. Structures of the corrin ring, cobalt, and tail of Cbl are colored in black, blue, and green, respectively. Upper ligand structures of Cbl were omitted for clarity. Riboswitch RNA structures are depicted as space-filled models with the L5-L13 kissing loop in pink and the rest of the RNA in gray.

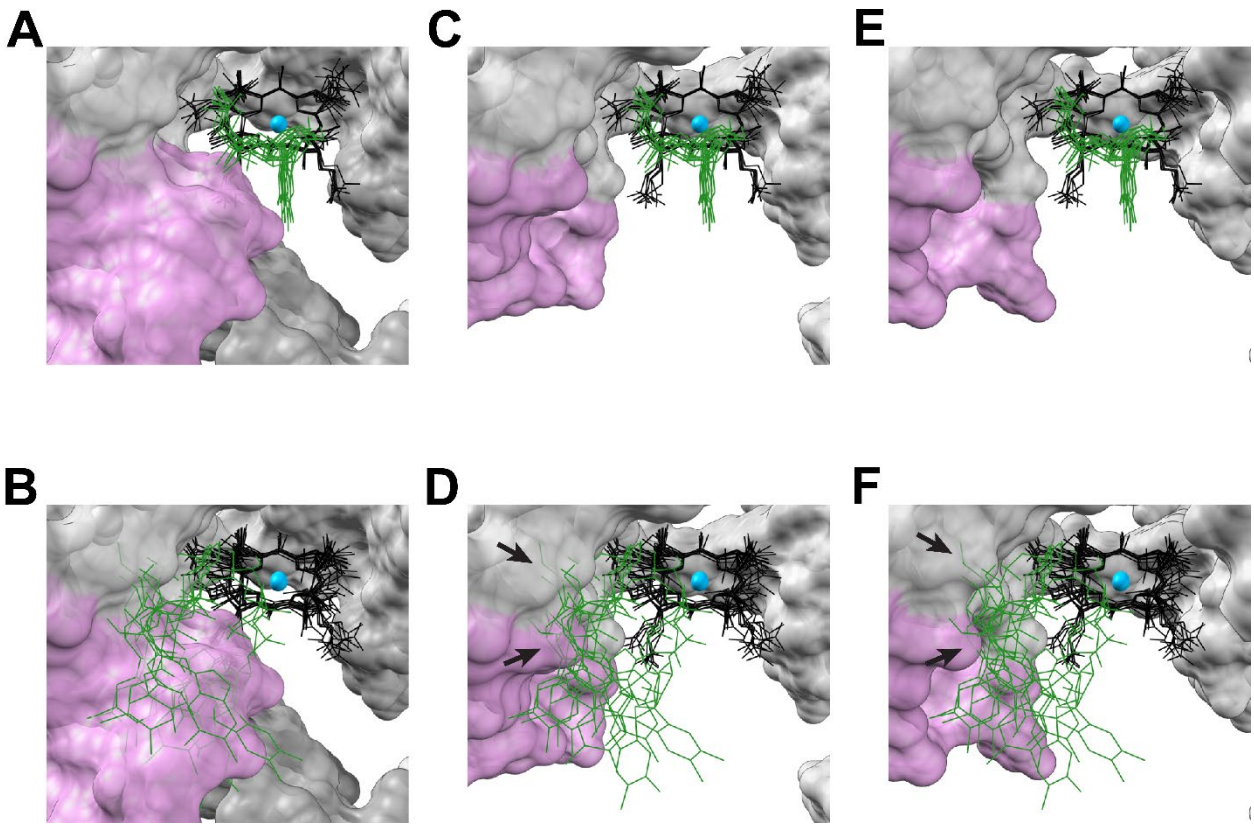


Figure 2.14 – Cbl-binding sites in X-ray crystal structures of various Cbl-riboswitches. The marine metagenome derived riboswitch *env8* (PDB ID 4FRN) (A, B), the *B. subtilis btuFCDR* Cbl-riboswitch (PDB ID 6VMY) (C, D), and the *S. thermophilum cblTS* Cbl-riboswitch (PDB ID 4GXY) are depicted with base-ON (A, C, E) and base-OFF (B, D, F) Cbl alignments. Cbl models were aligned by the cobalt and coordinating nitrogen atoms in the corrin ring. Structures of the corrin ring, cobalt, and tail of Cbl are colored in black, blue and green, respectively. Upper ligand structures of Cbl were omitted for clarity. Riboswitch RNA structures are depicted as space-filled models with the L5-L13 kissing loop in pink in panels A and B, and just the L5 in pink in panels C-F. The rest of the RNA structures in each panel are in gray. Arrows in panels D and F point to clashes between RNA and the Cbl tails.

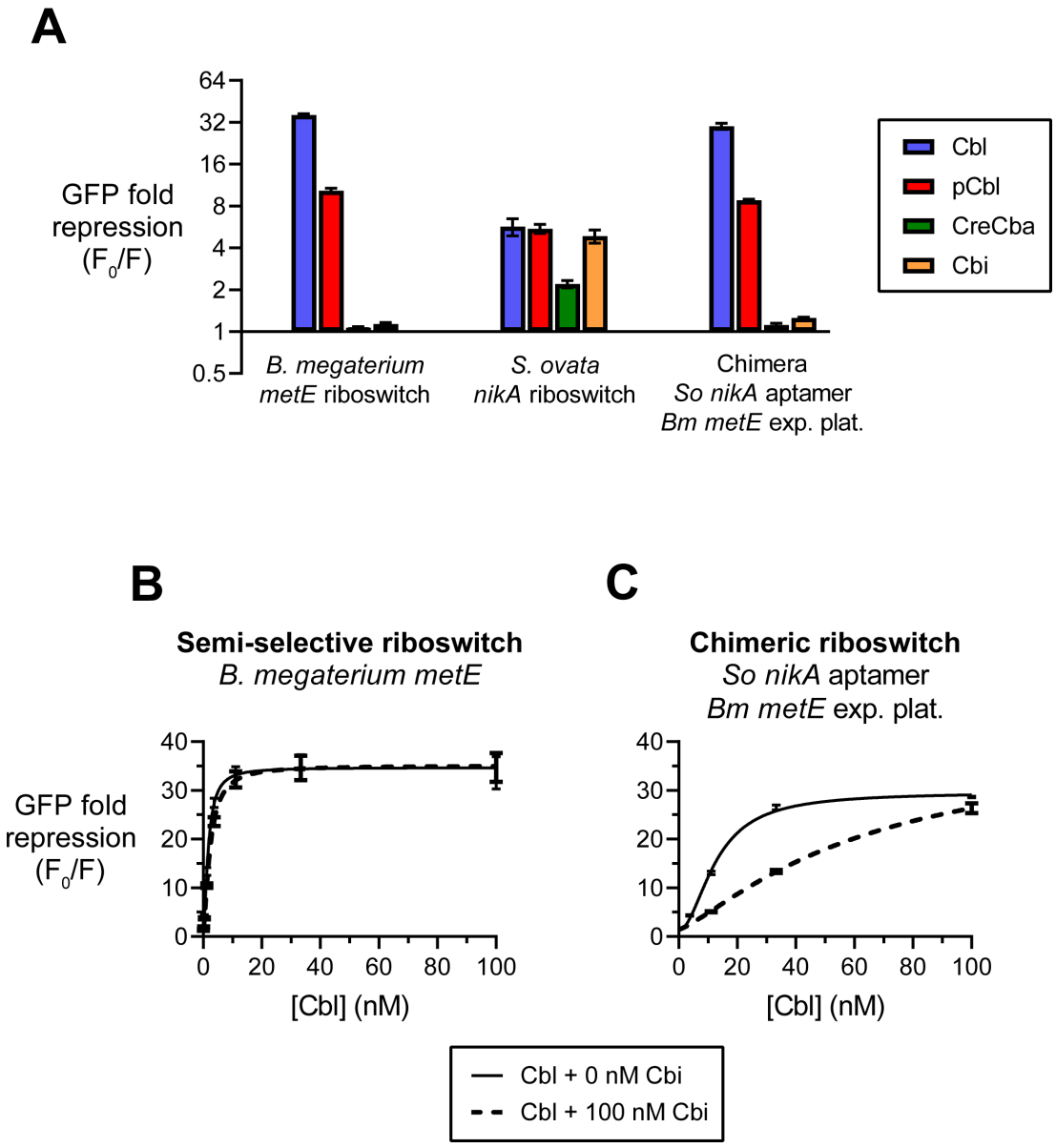


Figure 2.15 – The aptamer and expression platform of the *B. megaterium* riboswitch contribute to corrinoid selectivity. (A) Repression of GFP expression with 100 nM corrinoid is shown for *B. megaterium metE* riboswitch, *S. ovata nikA* riboswitch, and a chimeric fusion of the *S. ovata nikA* aptamer and *B. megaterium metE* expression platform domains. Cbl dose-responses with or without 100 nM Cbi in (B) *B. megaterium metE* and (C) chimeric riboswitches. Data points and error bars represent mean and standard deviation of four independent replicates.

2.3.5 Gene regulatory strategy of corrinoid selectivity

While the prior experiments clearly demonstrate that Cbl-riboswitches are capable of distinguishing between corrinoids, we wondered what purpose Cbl-riboswitch corrinoid selectivity might serve in cells. We posit that Cbl-riboswitch selectivity reflects a regulatory strategy that promotes efficient gene expression while avoiding mis-regulation. As a specific example, we hypothesize that only corrinoids that are functionally compatible with a Cbl-dependent enzyme should repress the expression of its Cbl-independent counterpart. We tested this hypothesis directly in *B. subtilis* by examining the methionine synthase isozymes MetE (Cbl-independent) and MetH (Cbl-dependent), which are commonly found in bacterial genomes with a Cbl-riboswitch upstream of the *metE* gene [47, 242]. The *B. subtilis* genome contains *metE* but lacks *metH*, and no Cbl-riboswitch is located upstream of *metE*. We therefore constructed strains of *B. subtilis* that heterologously express the *metE* or *metH* locus from the related species *B. megaterium*, in a Δ*metE* background (Figure 2.16A). The *B. megaterium metE*-expressing strain contains the entire transcribed region of the *B. megaterium metE* locus, including the SAM and Cbl tandem riboswitch located upstream of the *metE* open reading frame. The *B. megaterium metH*-expressing strain contains the *metH* open reading frame and its predicted ribosome binding site. In each strain, the *B. megaterium* genes are constitutively transcribed from the promoter P_{Veg} and also contain a transcriptionally fused *gfp* to measure expression levels. Growth of the *B. subtilis* strain expressing *B. megaterium metH* in a medium lacking methionine was supported to varying extents by most benzimidazolyl and both phenolyl cobamides, but not by [5-OHBza]Cba, the purinyl cobamides, or Cbi (Figure 2.16B, red squares). This result indicates that MetH-dependent growth is influenced by the corrinoid tail structure, as observed previously in other bacteria [149, 160, 231, 235, 253]. In the *B. subtilis* strain expressing riboswitch-regulated *metE*, growth was suppressed by benzimidazolyl cobamides to different extents (Figure 2.16B, blue circles). This growth pattern coincides with the GFP repression measured for the *metE* riboswitch (Figure 2.17), except that the repression of *metE* by purinyl cobamides is apparently insufficient to suppress growth in this context. Comparison of the two strains in response to a suite of corrinoids reveals a striking correspondence between riboswitch-mediated suppression of growth in the *metE*-containing strain and growth promotion by corrinoids in the *metH*-containing strain (Figure 2.16). [5-OHBza]Cba and the phenolyl cobamides are exceptions to the trend, though in neither case is MetE-dependent growth completely suppressed by a corrinoid incompatible with MetH. This result demonstrates that riboswitch-based repression and cobalamin-dependent isozyme function are largely reciprocal for the *B. megaterium metE-metH* pair and suggests that Cbl-riboswitch specificity may generally adhere to a regulatory strategy reflecting the cell's corrinoid preference.

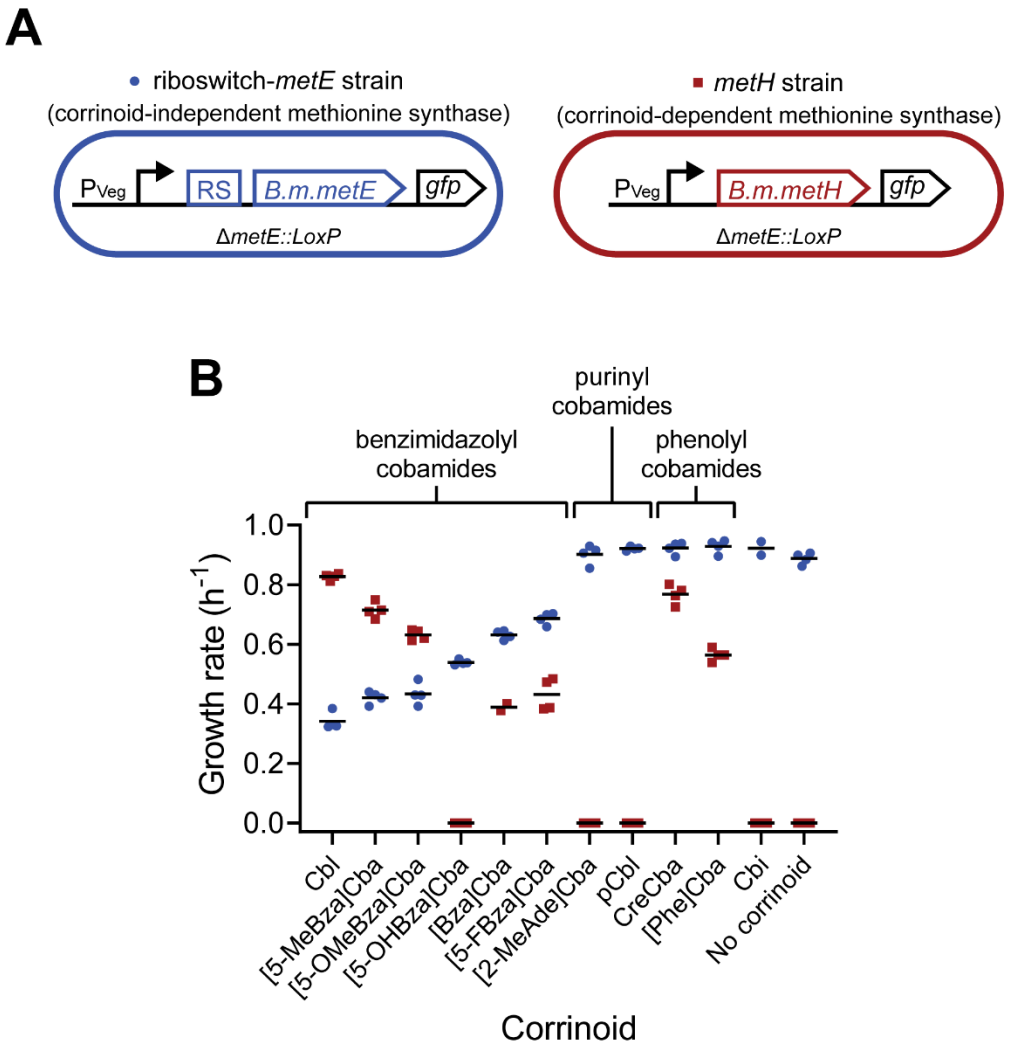


Figure 2.16 – Corrinoid selectivities of *metE* riboswitch and *metH* activity are largely balanced. (A) The 'riboswitch-*metE*' strain (blue) is a *B. subtilis* Δ*metE::loxP* strain heterologously expressing the *B. megaterium metE* operon which encodes a corrinoid-independent methionine synthase downstream from a SAM-Cbl tandem riboswitch (RS). The '*metH*' strain (red) is a *B. subtilis* Δ*metE::loxP* strain heterologously expressing the *B. megaterium metH* gene which encodes a corrinoid-dependent methionine synthase. Each methionine synthase gene is constitutively expressed by the P_{Veg} promoter and is transcriptionally fused to *gfp*. (B) Growth rates of the 'riboswitch-*metE*' (blue circles) and '*metH*' (red squares) strains were measured in methionine-dependent culture conditions. Data points are individual growth rate measurements and black horizontal lines represent the mean of the four replicate measurements.

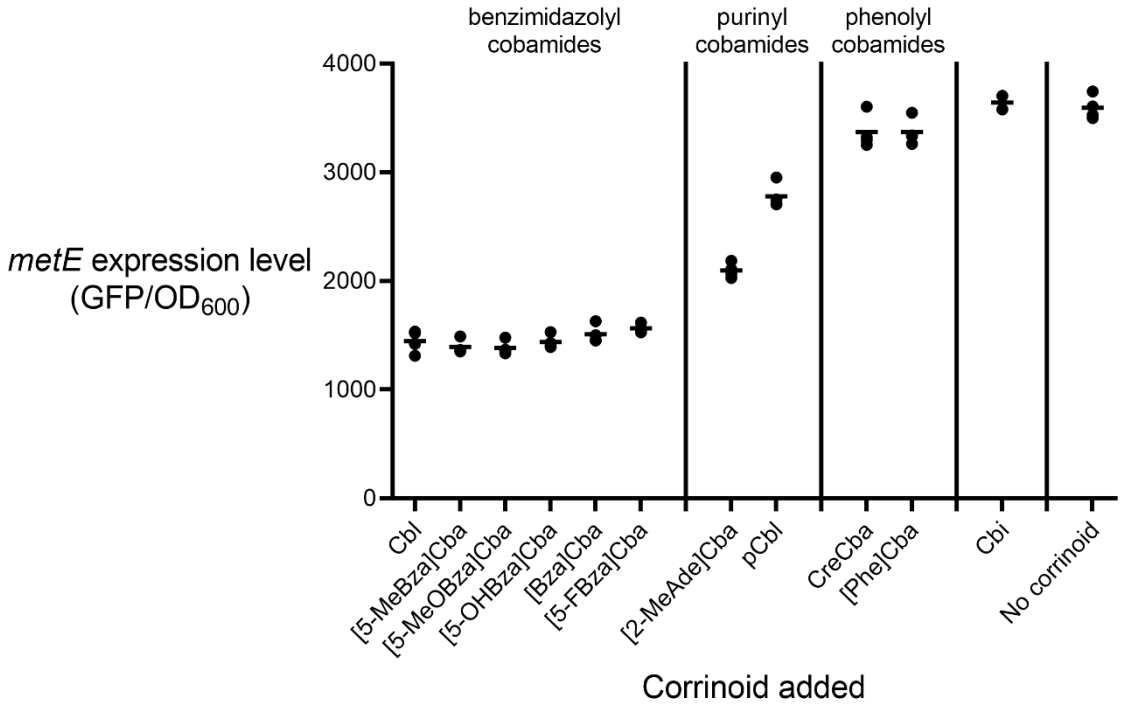


Figure 2.17 – Corrinoid-dependent repression of *B. megaterium metE-gfp* heterologously expressed in *B. subtilis* Δ*metE::LoxP*. GFP fluorescence per OD_{600} was measured in methionine-dependent cultures supplemented with 10 nM corrinoids. Data points are individual measurements and black horizontal lines represent the mean of four replicate measurements.

2.4 Discussion

Riboswitches are key regulators of microbial gene expression. The Cbl-riboswitch was the first type discovered and is among the most widely distributed riboswitch classes in bacteria and archaea [101, 102]. Previous biochemical and structural studies have uncovered the major molecular features of the Cbl-riboswitch response to Cbl [120-122, 236, 237, 242, 244, 245, 254]. Yet few studies have examined how other naturally occurring corrinoids impact gene regulation by Cbl-riboswitches [163]. Here, we found that Cbl-riboswitches vary in their ability to discriminate between corrinoids, with some being semi-selective on the basis of corrinoid base-ON/OFF state, and others being promiscuous. These results were enabled by a carefully designed fluorescent reporter system capable of measuring the responses of dozens of Cbl-riboswitches to multiple corrinoids *in vivo* (Figures 2.3-2.5).

We observed that Cbl-riboswitches display different degrees of corrinoid selectivity, with some that respond to a subset of corrinoids (semi-selective) and others that respond to all tested corrinoids (promiscuous) (Figure 2.5-2.8). Our chimeric riboswitch results suggest that sequence and structural determinants of corrinoid selectivity are dispersed throughout the Cbl-riboswitch aptamer scaffold rather than being confined to a single conserved region (Figure 2.10). This finding contrasts with other studies of riboswitch specificity. For example, in a study of Cbl upper ligand specificity, a few key residues in the Cbl binding site were sufficient to fully convert a Cbl-riboswitch from MeCbl-specific to AdoCbl-specific [236]. In purine riboswitches, effector specificity is achieved by positioning of a critical conserved uracil or cytosine residue in the binding site of the aptamer, which specifically forms a base-pair with the adenine or guanine effector, respectively. Among the various families of SAM riboswitches, highly specific binding of SAM and exclusion of SAH is achieved by specific RNA structures that discriminate the charged sulfonium ion of SAM from the uncharged sulfoether of SAH [106]. In contrast, the SAM/SAH riboswitch class attains effector promiscuity for SAM and SAH by a general lack of interaction between the RNA and the aminocarboxypropyl side chains of these effectors [104, 105]. This is reminiscent of Cbl-riboswitches which similarly have few molecular contacts between the RNA and corrinoid tail [122, 163, 245, 254].

Overall, our structural model analyses support a mechanism in which semi-selective Cbl-riboswitches primarily sense the distinct corrinoid tail orientations of the base-ON and base-OFF forms. In this case, Cbl-riboswitches only indirectly sense the chemical composition of the variable lower ligand group, with differential binding largely determined by steric effects and shape complementarity (Figures 2.13-2.14). Recent molecular dynamics simulations of the *T. tengcongensis* Cbl-riboswitch suggest that the kissing loop structure may form prior to effector binding, which would place even greater constraints on the corrinoid tail orientation to achieve shape complementarity with its binding site [255]. Additionally, some Cbl-riboswitches may in fact bind base-OFF corrinoids but disrupt subsequent formation of downstream regulatory structures of the expression platform, perhaps by disrupting the kissing loop (Figures 2.13-2.15). A similar feature has been observed in tetrahydrofolate (THF) riboswitches where chemical variations in the para-aminobenzoic acid moiety of THF analogs differentially perturb expression platform structures without affecting aptamer binding [256]. The mechanisms of corrinoid

selectivity of Cbl-riboswitches could be directly tested in future structural or biochemical studies of promiscuous Cbl-riboswitches with base-ON and base-OFF corrinoids.

In regard to gene regulatory strategies, it seems sensible that Cbl-riboswitches are not highly effector specific because bacteria are often flexible in their corrinoid usage. A variety of corrinoids have been shown to support growth of *C. difficile*, *S. meliloti*, and *S. ovata* despite each of these organisms displaying strong preferences in their corrinoid production [139, 143, 231]. Also, since corrinoid auxotrophy is prevalent among corrinoid-dependent bacteria, many organisms may need to take advantage of the wide range of corrinoids that may be available in their environment [155]. Thus, the range of effector selectivity that we observe among Cbl-riboswitches may reflect a coevolution between corrinoid-responsive gene regulation and the corrinoid-dependent physiology. Our result demonstrating complementary corrinoid selectivity between *B. megaterium* MetH-dependent growth and Cbl-riboswitch-dependent expression of MetE is consistent with this notion (Figure 2.16).

Alternatively, the preference for corrinoid base-ON constitution among Cbl-riboswitches may function as a proxy to discriminate complete corrinoid coenzymes from incomplete corrinoids such as Cbi, which often function poorly as coenzymes. This idea has been proposed as an explanation for the remarkably high selectivity of the corrinoid uptake system in mammals [257]. Interestingly, we found that all but one of the *S. ovata* riboswitches tested are promiscuous types that can respond to its natively produced CreCba. The notable exception is a semi-selective riboswitch upstream of the gene *cobT* (Supplemental Figure 4), which functions in a late step of corrinoid biosynthesis that occurs after synthesis of Cbi [144, 145, 258-262]. Thus, the Cbl-riboswitch may allow homeostatic regulation of *cobT* in response to complete corrinoids like Cbl, while preventing unproductive repression of *cobT* in the presence of incomplete corrinoids like Cbi. Future studies examining corrinoid-specific gene regulation of riboswitches in the context of their native organisms may help clarify which regulatory strategies are generally at play in corrinoid-related bacterial physiology.

There is significant interest in the fields of bioengineering and synthetic biology to use riboswitches as gene regulatory devices because they act more rapidly and efficiently than protein-based regulatory systems [206, 207, 263, 264]. Riboswitches have also become desirable therapeutic drug targets because they often control essential metabolic pathways in pathogenic microbes [111, 116, 205]. Broader consideration of the conformational dynamics of larger types of effector molecules including organic cofactors, antibiotics, and their analogs could aid efforts to engineer synthetic RNA-based regulatory platforms that function robustly *in vivo*. This could also inform future efforts to develop synthetic antimetabolites that, for example, elicit gene mis-regulation, instead of a more common strategy of creating inhibitors for enzymes [117, 265].

The chemical diversity of corrinoids is intrinsically linked to a vast array of metabolic processes and microbial interactions. Yet it remains unclear how microbes have evolved to cope with and thrive on the assortment of natural corrinoid analogs, especially when compared to other primary metabolites including organic cofactors, nucleotides and amino acids which typically require one specific structural form for precise biological functions. We have gained new appreciation for the impacts of chemical diversity on biological function by focusing on the Cbl-riboswitch with its

distinctively complex structure and regulatory mechanism, and by accounting for the often overlooked biological and ecological roles of corrinoid analogs. Future studies into the evolution of microbial molecular specificity for corrinoids may yield further insight into the nature of these exceptionally versatile coenzymes.

2.5 Materials and methods

2.5.1 Cbl-riboswitch sequence analysis

Cbl-riboswitch sequences, chromosomal coordinates and regulon information were downloaded from the RiboD online database [47]. The genome of *Sporomusa ovata* was not included in the RiboD database, so we used the RiboswitchScanner webserver to search for Cbl-riboswitches in this organism [266, 267]. Cbl-riboswitch aptamer sequences were manually aligned by conserved secondary structures bounded by the 5' and 3' ends of the P1 stem [64, 150]. The P13-L13 stem loop and potential intrinsic transcriptional termination hairpin structures of the expression platform were identified using secondary structure prediction tools in RNAstructure 6.2 [268, 269]. Intrinsic terminators were identified as stem loops directly preceding a sequence of five or more consecutive uracil residues [11, 13]. Sequence alignment and annotation was carried out in JalView 2.11.1.4 [270]. Cartoons of riboswitch secondary structures were constructed using the StructureEditor program of RNAstructure 6.2 [269].

2.5.2 Corrinoid production, extraction, purification, and analysis

Cyanocobalamin, adenosylcobalamin, methylcobalamin, hydroxocobalamin, and dicyanocobinamide were purchased from MilliporeSigma. All other corrinoids used in this study were produced in bacterial cultures and purified in cyanated form as previously described [140, 143, 146, 231]. For the experiments in Figures 2.11 and 2.12, corrinoids other than cobalamin were chemically adenosylated to obtain the coenzyme (5'-deoxyadenosylated) form as previously described [140, 231].

UV/Vis spectra were collected from corrinoid samples in UV/Vis-transparent 96-well microtiter plates (greiner bio-one UV-STAR® 675801) using a BioTek Synergy 2 or Tecan Infinite M1000 Pro plate reader. To measure concentrations of corrinoid stock solutions, corrinoid samples were diluted 10-fold in 10 mM sodium cyanide to obtain the dicyanated base-OFF form of the corrinoid. The concentration of the dicyanated corrinoid was calculated using the extinction coefficient ε580 = 10.1 mM^{-1} cm^{-1} [134, 147]. For adenosylated corrinoids used in Figures 2.11 and 2.12, base-ON/OFF constitution at neutral pH was measured as the ratio of spectral absorbance at 525 nm and 458 nm in phosphate buffered saline solution pH 7.3 at 37°C [130, 131].

2.5.3 Plasmid and strain construction

Plasmids generated in this study were constructed with one-step isothermal assembly [271] and introduced into *E. coli* strain XL1-Blue by heat shock transformation. Riboswitch reporter plasmids were constructed in the shuttle vector pSG13 [272]. Riboswitch DNA sequences were inserted between the transcriptional start site of the constitutive P_{Veg} promoter and the *gfp*

translational start site of pSG13. For riboswitch sequences that resulted in no detectable GFP signal under any conditions, a synthetic ribosome binding site (RBS) sequence R0 (5'-GATTAACTAATAAGGAGGACAAAC-3') from pSG13 was placed between the riboswitch sequence and *gfp* translational start site.

All *B. subtilis* riboswitch fluorescent reporter strains and *B. subtilis* strains expressing *B. megaterium metE* and *metH* used in this study are derived from the high-efficiency transformation strain SCK6, which has a xylose-inducible competence gene cassette [273]. Preparation of competent cells and transformations of all SCK6-derived strains were performed as previously described [273]. The strain KK642, which overexpresses the corrinoid uptake genes, was constructed by deletion of gene *queG* and replacement of the promoter and 5' untranslated region of the *btuFCDR* operon with the P_{Veg} promoter and R0 RBS [272]. *B. subtilis* genes *queG*, *btuR*, and *metE* were targeted for deletion by recombination with kanamycin resistance cassettes containing flanking sequence homology to each respective locus. Kanamycin resistance cassettes were PCR-amplified from genomic DNA of *B. subtilis* strains BKK08910 (Δ*queG*::*kan*^{R}), BKK33150 (Δ*btuR*::*kan*^{R}), and BKK13180 (Δ*metE*::*kan*^{R}) [274]. Kanamycin resistance cassettes were removed by Cre-Lox recombination using plasmid pDR244 as previously described [274].

B. subtilis strains heterologously expressing *metE* and *metH* from *B. megaterium* were constructed as follows. The *metE* and *metH* genes were PCR-amplified from genomic DNA of *B. megaterium* DSM319 and cloned between the transcriptional start site and the *gfp* translational start site of pSG13. The *B. megaterium metE* amplified fragment starts at the SAM-Cbl tandem riboswitch in the 5' UTR and ends at the *metE* stop codon, whereas the *B. megaterium metH* fragment starts at the *metH* RBS, which is composed of 20 nucleotides preceding the *metH* translational start site and ends at the *metH* stop codon.

Riboswitch reporter plasmids and plasmids containing *B. megaterium metE* and *metH* were linearized by restriction enzyme digest with ScaI-HF (New England Biolabs) and selected for integration at the *amyE* locus of *B. subtilis* by plating on lysogeny broth (LB) agar with 100 µg/mL spectinomycin. Colonies were screened for integration at *amyE* by colony PCR. All stocks of bacterial strains were stored in 15% glycerol at -80 ºC.

2.5.4 Intracellular corrinoid accumulation experiments

B. subtilis strains were inoculated from single colonies into 50 mL LB and grown with aeration at 37 ºC in a shaking incubator (Gyromax 737R, Amerex Instruments, Inc.) for 5-6 hours until reaching an optical density at 600 nm (OD_{600}) of 1.0 to 1.5. Each culture was diluted 10-fold in LB and split into 13 25 mL cultures containing 0, 25, 250, or 2500 picomoles of a cyanated corrinoid (Cbl, pCbl, CreCba and Cbi). These cultures were incubated at 37 ºC with aeration for 3-4 hours to a final OD_{600} of 1.5 to 2.0. The cells were pelleted by centrifugation at 4,000 g for 10 min. Cell pellets were rinsed three times by resuspension in 10 mL phosphate buffered saline solution pH 7.3 followed by centrifugation. After the final centrifugation, tubes were wrapped in aluminum foil to protect adenosylated corrinoids from exposure to light.

To extract intracellular corrinoids, cell pellets were resuspended in 5 mL of 100% methanol by vigorous vortexing for 30 secs. Samples were stored at -80 ºC until the next day. Frozen lysates

were heated in an 80 ºC water bath for 1.5 hours, with 15 seconds of vortexing every 30 minutes. Methanol concentration of each sample was diluted to 10% by adding 45 mL of water, and cell debris was pelleted by centrifugation at 4,000 g for 10 mins. The supernatants were used for the subsequent steps.

All of the following steps were carried out in darkened rooms illuminated with red light to preserve light-sensitive adenosylated corrinoid samples. Solid-phase extraction of adenosylated corrinoids with Sep-Pak C18 cartridges (Waters) was performed as previously described [231]. Solvents were evaporated in a vacuum concentrator centrifuge (Savant SPD1010, Thermo Scientific) at 45 ºC and the samples were resuspended in 500 µL deionized water and passed through 0.45 µm pore-size filters (Millex-HV, MilliporeSigma).

Corrinoids were analyzed on an Agilent 1200 series (high-performance liquid chromatography system equipped with a diode array detector (Agilent Technologies). Samples were injected into an Agilent Zorbax SB-Aq column (5-µm pore size, 4.6 by 150 mm). The following HPLC method was used: solvent A, 0.1% formic acid–deionized water; solvent B, 0.1% formic acid–methanol; flow rate of 1 mL/minute at 30°C; 25% to 34% solvent B for 11 minutes, followed by a linear gradient of 34% to 50% solvent B over 2 minutes, followed by a linear gradient of 50% to 75% solvent B over 8 minutes.

2.5.5 Riboswitch fluorescent reporter assays

Corrinoid dose-response assays of riboswitch reporter strains were set up as follows. Saturated cultures of the riboswitch reporter strain in LB were diluted 200-fold in LB and dispensed into 96-well microtiter plates (Corning Costar Assay Plate 3904) containing a range of concentrations of various corrinoids. The plates were sealed with gas diffusible membranes (Breathe-Easy) and incubated at 37 ºC for 4 to 5 hours in a benchtop heated plate shaker (Southwest Science) at 1,200 revolution per minute (rpm). GFP fluorescence (excitation/emission/bandwidth = 485 / 525 / 10 nm) and absorbance at 600 nm (A_{600}) were measured on a Tecan Infinite M1000 Pro plate reader. The A_{600} measurements of uninoculated medium and fluorescence measurements of the parental control strains lacking *gfp* were subtracted from all readings. Data were plotted and analyzed in GraphPad Prism 9.

2.5.6 3D structural analysis of corrinoids and macromolecular models

Molecular models of base-ON and base-OFF/His-ON forms of cobalamin in complex with various proteins and RNAs were downloaded from the Protein DataBank (PDB) (Table 2) [275]. PDB files were analyzed in UCSF Chimera 1.14 [276]. Corrinoid molecular models were aligned with each other by the central cobalt atom and coordinating nitrogen atoms of the corrin ring, using the PDB ID 4GMA Cbl model as a reference. Corrinoid models were aligned and within the binding sites of riboswitch structures PDB IDs 4GMA, 4FRN, 4GXY, and 6VMY [122, 244, 245].

2.5.7 Methionine-dependent growth of B. subtilis *strains*

B. subtilis strains were streaked from frozen stocks onto LB agar plates and incubated overnight at 37 ºC for 14-18 hrs. Single colonies were used to inoculate 3 mL liquid starter cultures containing Spizizen minimal medium supplemented with 0.02% D-glucose and 0.2% L-Histidine

(SMM) [277]. Starter cultures of the *metH*-expressing strains were supplemented with 1 nM CNCbl to support growth. Starter cultures were incubated overnight shaking (250 rpm, 37 ºC) for 20 hours, reaching cell density of about OD_{600} = 1.0. Starter cultures were diluted 500-fold by transferring 50 µL of starter culture to 25 mL of SMM. Then 75 µL of the diluted culture were dispensed into wells of a 96 well microtiter plate supplemented with 20 nM of various corrinoids. Plates were sealed with gas diffusible membranes (Breathe-Easy, MilliporeSigma) and incubated at 37 ºC on the 'high shaking' setting of a BioTek Synergy2 plate reader. Growth kinetics and *metE* and *metH* expression were measured by A_{600} and GFP fluorescence every 15 minutes for 72 hours. Data were plotted and analyzed in GraphPad Prism 9.

Chapter 3 – Examination of atypical regulons in corrinoid specialists reveals potential novel functionalities of cobalamin riboswitches and the genes they control

3.1 Abstract

Cobalamin riboswitches are the second-most widespread riboswitch class in bacteria. Yet relatively few cobalamin riboswitches have been studied in detail compared to the purine, S-adenosylmethionine, and glycine classes. The sequence and structural diversity of cobalamin riboswitches may contain novel functionality, but an effective search strategy is first required to identify potentially rare cases among thousands of cobalamin riboswitch sequences. I used two conceptual approaches to address this problem. In the first approach, I focus on atypical regulon architectures such as tandem riboswitches and riboswitches regulating uncommon target genes, reasoning that these riboswitches and genes are likely to have atypical and potentially novel functions. The second approach is to examine cobalamin riboswitches in organisms that rely heavily on corrinoids, reasoning that these organisms are more likely to have evolved novel corrinoid-related functions. With these search strategies, I identify a novel activator riboswitch, characterize modular functionality of tandem riboswitches, and develop new hypotheses about corrinoid-dependent isoenzyme physiology. Together, these results highlight the utility of leveraging atypicality to uncover new facets of corrinoid-related gene regulation and physiology.

3.2 Introduction

The discovery of metabolite-binding RNA switches (riboswitches) was a landmark finding in the field of noncoding RNA biology [101]. Prior to this, the prevailing assumption was that biomolecular detection of small molecules by direct binding was an exclusive feature of proteins, not RNAs. Subsequent identification of dozens of riboswitch classes demonstrated that riboswitch RNAs are prevalent and widespread gene regulatory elements that organisms use to sense and respond to intracellular metabolites [64]. Riboswitch-based gene regulation controls a wide array of physiological processes in bacteria including several primary metabolisms, antibiotic production, antibiotic resistance, cell signaling, virulence, and biofilm formation. The study of riboswitches continues to yield fundamental insights into how bacteria live.

Basic principles of riboswitch structures and functions have been established through *in vitro* biochemical and *in vivo* biological studies. Yet, for any given riboswitch class, only a few representative examples have been studied in mechanistic detail. Thus, while we may understand the biological roles of riboswitches in broad strokes, it is likely that novel functionality awaits discovery among the thousands of riboswitches that have been bioinformatically identified in bacterial genome sequences [190].

3.2.1 Examples of previously studied atypical Cbl-riboswitch regulons

Examination of atypical riboswitch regulon configuration has been a successful approach for discovering novel functions of riboswitches as well as the genes they regulate. For example, several variants of the guanidine riboswitch class were revealed to respond to distinct effector molecules including guanosine tetraphosphate (ppGpp), phosphoribosyl pyrophosphate (PRPP),

and several types of nucleoside diphosphate molecules [278, 279]. The identification of the true effectors for these distinct riboswitch variants was aided by the uncommon target genes that were associated with each variant.

Among the Cbl-riboswitches, examination of atypical regulons has also proven fruitful in previous studies. The first tandem riboswitch composed of two distinct riboswitch classes was the SAM-Cbl riboswitch in *Bacillus clausii*. Although SAM-Cbl riboswitches are not widespread, this finding revealed for the first time that riboswitches can function as dual input logic gates, evaluating the presence of multiple effectors to control target gene expression [242]. This has informed efforts to engineer riboswitches with complex functionality [280, 281].

Another example is the Cbl-riboswitch regulation of small non-coding RNAs. Variations of this complex mechanism were discovered in Cbl-riboswitches that control uncommon targets, the corrinoid-dependent ethanolamine utilization (*eut*) operons of *Listeria monocytogenes* and *Enterobacter faecalis* [45, 46]. Cbl-riboswitches typically operate as repressors, so Cbl-riboswitch control of corrinoid-dependent enzyme expression is uncommon. In this case, the Cbl-riboswitches indirectly upregulate the *eut* genes by controlling small RNAs in response to Cbl.

3.2.2 Early inspiration from the genome of Desulfobulbus propionicus *DSM 2032*

The bulk of my doctoral research described in Chapter 2 focused on examining riboswitches with typical configurations (singlet riboswitches regulating common target genes) in order to gain a broad understanding of corrinoid specificity of Cbl-riboswitches. However, early on in my studies, I made the initial observation of an atypical regulon in *Desulfobulubus propionicus* DSM 2032*,* which intrigued me enough to further examine other atypical Cbl-riboswitch regulons. I suspected this approach may provide new insights into corrinoid-related gene regulation and physiology.

The genome of *D. propionicus* DSM 2032 contains 13 Cbl-riboswitches in the 5'UTRs of 12 operons. The *mutA* operon has two Cbl-riboswitches in the 5'UTR which potentially function together as a tandem doublet riboswitch. The 11 other Cbl-riboswitches in the *D. propionicus* DSM 2032 genome are standard singlet Cbl-riboswitches. The *mutA* operon contains three structural genes: *mutA* and *mutB*, which encode subunits of the corrinoid-dependent enzyme methylmalonyl-CoA mutase (MCM), and *meaB* which encodes a chaperone protein that supports MCM enzymatic function. MCM catalyzes the interconversion of methylmalonyl-CoA and succinyl-CoA, a reaction required in the catabolism of branched-chain amino acids, certain short chain fatty acids, and certain polyhydroxyalkanoates [167]. Cbl-riboswitches are thought to function only as repressors, so why would a corrinoid-dependent enzyme such as MCM be regulated by a Cbl-riboswitch, let alone two in tandem? Perhaps this Cbl-riboswitch is a novel activator that upregulates MCM expression in the presence of corrinoids. Alternatively, it may selectively repress MCM expression in response to corrinoids that do not support MCM function.

Here, I describe and characterize several atypical Cbl-riboswitches that regulate MCM genes and other tandem doublet Cbl-riboswitch regulons. I discuss the implications of these findings and propose future experiments to examine potentially new facets of corrinoid biology.

3.3 Results and discussion

3.3.1 Examination of Cbl-riboswitches that regulate MCM genes

Cbl-riboswitches that regulate MCM genes are relatively uncommon. Among a curated list of 3,313 bacterial Cbl-riboswitches in the RiboD riboswitch database, only eight are predicted to regulate MCM genes [47]. These riboswitches belong to *D. propionicus*, *Veillonella parvula*, *Selenomonas sputigena*, *Selenomonas ruminatium*, *Pelobacter propionicus*, *Propionibacterium freudenreichii*, and *Cutibacterium acnes* (Table 3.1). Another MCM-regulated operon was identified in the genome *Cloacibacillus porcorum*, a species which was not included in the RiboD database [147].

Organism	**Cbl-riboswitch regulated genes encoding MCM related factors**	**Genes encoding MCM isoenzymes and related factors**
Desulfobulbus propionicus DSM 2032	Despr_3086 (*mutA*), Despr_3085 (*mutB*), Despr_3084 (*meaB*)	Despr_1914 (MCM), Despr_1915 (*meaB*)
Veillonella parvula DSM 2008	Vpar_1763 (*mutA*), Vpar_1762 (*mutB*), Vpar_1761 (*meaB*)	Vpar_1249 (acetyl-CoA hydrolase), Vpar_1248 (MCM), Vpar_1247 (Cbl-binding domain), Vpar_1246 (*meaB*), Vpar_1245 (methylmalonyl-CoA epimerase)
Cutibacterium acnes ATCC 11828	TIIST44_11480 (*mutA*), TIIST44_11475 (*mutB*)	none
Selenomonas sputigena ATCC 35185	Selsp_0914 (*mutA*), Selsp_0915 (*mutB*), Selsp_0916 (*meaB*), Selsp_0917 (*msrA*),	none
Selenomonas ruminantium subsp. *lactilytica* TAM6421	SELR_07780 (*mutA*), SELR_07790 (*mutB*), SELR_07800 (*meaB*), SELR_07810 (*msrA*)	none
Propionibacterium freudenreichii subsp. *shermanii* CIRM-BIA1	PFREUD_07660 (*mutA*), PFREUD_07650 (*mutB*), PFREUD_07640 (*meaB*)	none
Pelobacter propionicus DSM 2379	Ppro_1285 (*mutA*), Ppro_1284 (*mutB*), Ppro_1283 (*meaB*), Ppro_1282 (methylmalonyl-CoA epimerase)	Ppro_0523 (MCM), Ppro_0522 (methylmalonyl-CoA epimerase)
Cloacibacillus porcorum CL-84	BED41_13600 (*mutA*), BED41_13605 (*mutB*), BED41_13610 (*meaB*)	BED41_04280 (MCM), BED41_04285 (Cbl binding domain), BED41_04290 (methylmalonyl-CoA epimerase

Table 3.1. Bacterial species with Cbl-riboswitch regulated MCM genes.

I constructed GFP reporter strains of *Bacillus subtilis* for four of these riboswitches. The *D. propionicus*, *V. parvula*, and *S. sputigena* riboswitches functioned in the reporter system, displaying robust changes in GFP fluorescence in response to corrinoids (Figure 3.1). However, the *C. acnes* riboswitch displayed no change in GFP fluorescence in response to any of the corrinoids tested (data not shown). Each of the functional reporters suggest that these Cbl-riboswitches are typical repressors, not activators. The *D. propionicus* tandem doublet riboswitch responds to Cbl, pCbl, and Cbi with 6-fold or greater repression and the response to CreCba is weak with less than 2-fold repression (Figure 3.1A). The response of the *S. sputigena* riboswitch

is overall much weaker but displays a greater than 2-fold repression in response to Cbi and pCbl at the highest concentrations tested (Figure 3.1B). The *V. parvula* riboswitch displays equal response to Cbl, pCbl, and Cbi (3-fold repression of *gfp*) and a less than 2-fold repression with CreCba. Each of these riboswitches displays a relatively strong response to Cbi, which may suggest that MCM expression is repressed when incomplete corrinoids that are not functional cofactors are present at high concentration. Additionally, *V. parvula* and *D. propionicus* are known produce CreCba *de novo*, so perhaps these regulons allow MCM expression when CreCba is present (Kenny Mok, personal communication) [144]. However, the robust responses to complete corrinoids Cbl and pCbl are less clear. Perhaps these MCMs are incompatible with Cbl and pCbl, but it remains to be tested.

Intriguingly, several of these organisms with Cbl-riboswitches regulating *mutA* possess two sets of MCM isoenzymes (Table 3.1). The predicted MCM isoenzymes in *D. propionicus* and *V. parvula* are predicted to be true MCMs based on conserved amino acids in the substrate binding site that can be used to distinguish between homologous Acyl-CoA mutases [282]. This analysis was done by a former Taga Lab researcher, Olga Sokolovskaya, who previously studied MCM corrinoid specificity (personal communication). The Cbl-riboswitch control of MCM isoenzyme expression is reminiscent of Cbl-riboswitch control of MetE (corrinoid-independent methionine synthase isoenzyme) expression, which allows an organism to switch between usage of MetE and MetH (corrinoid-dependent methionine synthase isoenzyme) based on the type and quantity of corrinoids present in the cell [156]. I hypothesize that a similar kind of regulatory strategy allows for choosing which MCM isoenzyme to use based on corrinoid availability. There are several reasons why it may be beneficial to choose between two corrinoid-dependent MCM isoenzymes. For example, the two isoenzymes could have distinct corrinoid compatibilities, so the Cbl-riboswitch enables the organism to choose which isoenzyme to use based on the types of corrinoids present in the cell. Another possibility is that the Cbl-riboswitch controlled MCM may have a higher corrinoid binding affinity but lower catalytic efficiency, so the Cbl-riboswitch enables the cell to switch to the lower affinity but catalytically efficient MCM in high corrinoid concentrations. In follow up studies, it would be interesting to measure and compare the corrinoid binding affinities, specificities, and catalytic properties of these MCM isoenzymes to understand this atypical regulatory strategy [140, 231].

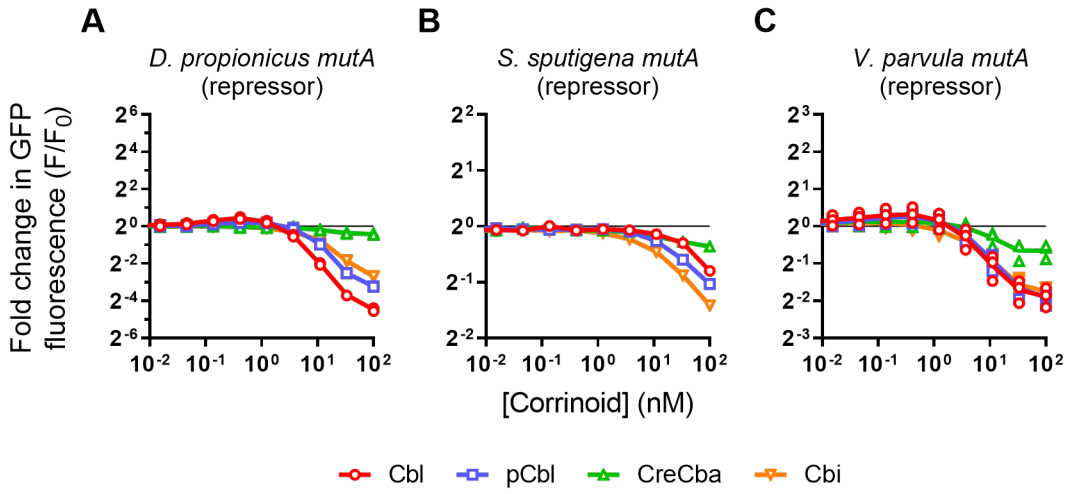


Figure 3.1 – The Cbl-riboswitches that regulate corrinoid-dependent MCM genes function as repressors. Corrinoid dose responses of (A) *D. propionicus mutA* (N=2), (B) *S. sputigena mutA* (N=1), and (C) *V. parvula mutA* (N=3) Cbl-riboswitch reporter strains. Thin horizontal lines demarcate no change in expression. Data points of each replicate are plotted with lines connecting mean values.

3.3.2 The Bacillus halodurans cobT *riboswitch functions as a transcriptional activator*

Each of the MCM Cbl-riboswitch reporters repressed *gfp* expression in response to corrinoids and are thus unlikely to be novel activators of MCM expression in their native organisms. Nonetheless, in the course of my investigations of corrinoid specificity, I came across a novel activator riboswitch in the bacterial species *Bacillus halodurans* C-125. *B. halodurans* is an alkaliphilic soil bacterium closely related to *B. subtilis* [283]. Genomic analysis predicts that this organism has multiple corrinoid-dependent metabolisms and can salvage corrinoids from its environment (Table 3.2) [155]. A corrinoid salvager is an organism that does not have the genes required for *de novo* corrin ring biosynthesis but has genes for corrinoid tail and lower ligand attachment. This allows synthesis of complete corrinoid cofactors from late corrinoid intermediates such as Cbi [142, 152, 284]. Thus, it appears that *B. halodurans* relies upon the corrinoids produced by other species in its environment in order to carry out its corrinoid-dependent metabolisms. I was interested in the Cbl-riboswitches of this species because they may function in discriminating between compatible and incompatible corrinoids.

Operon name	Cbl-riboswitch?	Locus tag (gene)	Function(s)
cobT	Yes	BH0284 (*cobT*)	Activation of lower ligands for corrinoid attachment
metE	Yes	BH0437 (*metE*)	Corrinoid-independent methionine synthesis
btuF	Yes	BH1585 (*btuF*), BH1586 (*btuC*), BH1587 (*btuD*), BH1588 (*cobD*), BH1589 (*cobC*), BH1590 (*cobP*), BH1591 (*cobQ*), BH1592 (*cobS*), BH1593 (*cobC*), BH1594 (*cobU*), BH1595 (*btuR*)	Corrinoid uptake (BtuF, BtuC, BtuD) Corrinoid salvaging (CobD, CobC, CobP, CobQ, CobS, CobC, CobU), Corrinoid adenosylation (BtuR)
nrdA	Yes	BH0501 (*nrdA*), BH0502 (*nrdB*)	Corrinoid-independent deoxyribonucleotide synthesis
BH0798	Yes	BH0798	Acyl-CoA thioesterase (unknown function)
metB	No	BH1627 (*metB*), BH1628 (*metC*), BH1629 (*metF*), BH1630 (*metH*)	Corrinoid-dependent methionine synthesis, L-homocysteine synthesis
nrdJ	No	BH2810 (*nrdJ*)	Corrinoid-dependent deoxyribonucleotide synthesis
mutA	No	BH2958 (TetR family transcription factor), BH2957 (*mmdA*), BH2956 (*mutB*), BH2955 (*mutA*), BH2954 (*meaB*)	Corrinoid-dependent methylmalonyl-CoA mutase (MutA, MutB)
queG	No	BH1020 (*queG*) BH1021	Corrinoid-dependent tRNA modification (QueG)

Table 3.2. Operons of *Bacillus halodurans* C-125 containing corrinoid-related genes.

I first tested for corrinoid-specificity of MCM by culturing *B. halodurans* with branched chain amino acids (isoleucine and valine) as the sole carbon source [231, 285]. I found that *B. halodurans* grows to different densities depending on the type and quantity of corrinoid supplemented. Cbl provided the largest growth yields, followed by Cbi, CreCba and pCbl (Figure 3.2). Notably, even without corrinoid supplementation, *B. halodurans* grew to a substantial density of OD_{600} = 0.4. This is most likely due to the presence of the methylcitrate pathway in *B. halodurans* which bypasses MCM-dependent conversion of methylmalonyl-CoA to succinyl-CoA. Still, it appears that corrinoid supplementation is beneficial in these growth conditions. Similar corrinoid preference of MCM has been observed in another bacterial species, *Sinorhizobium meliloti*, as well as the human MCM ortholog. Lastly, another interesting aspect of this result is that Cbi provided a substantial growth benefit presumably by enabling salvaging of Cbi to produce a complete corrinoid cofactor. However, the *B. halodurans* genome does not contain any annotated lower ligand synthesis genes such as *bluB* or the *bza* genes to make the 5,6-dimethylbenzimidazolyl lower ligand of Cbl [141, 191-193]. Since the growth benefit of Cbi supplementation is greater than pCbl and CreCba, it is likely producing a different corrinoid which may be worth identifying in future experiments.

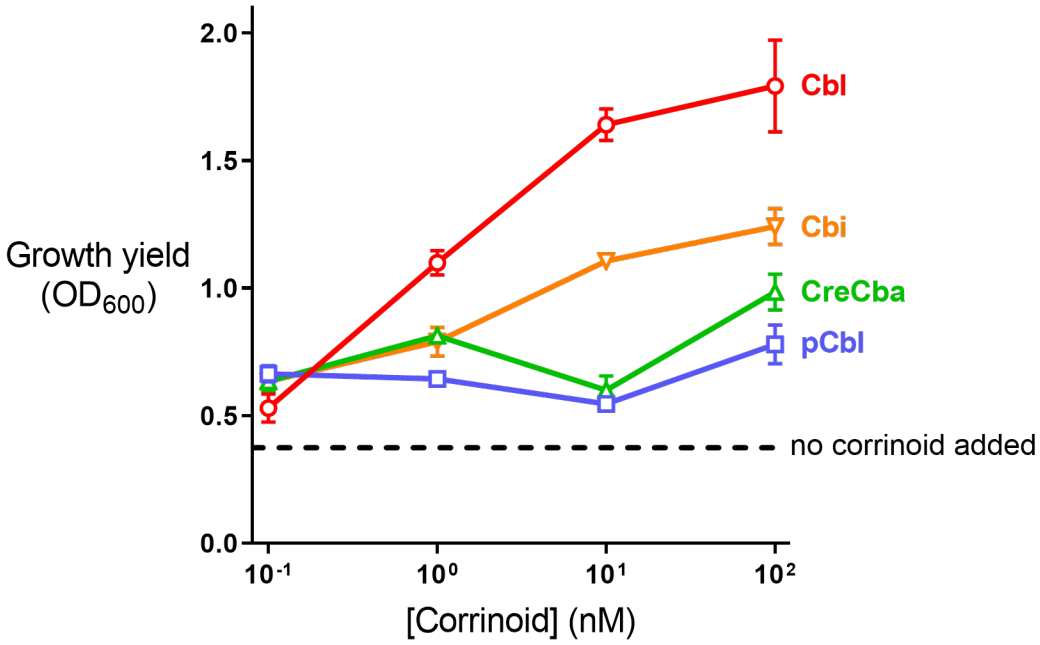


Figure 3.2 – *B. halodurans* displays corrinoid-specific growth in defined minimal medium with Ile and Val as the sole carbon source. Cultures were inoculated at an OD_{600} of 0.001 into media with the indicated corrinoid concentrations and the OD_{600} was measured after 48 hours. Data points and error bars represent mean and standard deviation of three replicates.

Next, I tested for corrinoid specificity of methionine synthesis by the MetE and MetH isozymes of *B. halodurans*. Initial growth experiments indicated that pCbl and Cbi had no effect on the growth kinetics of *B. halodurans* (data not shown). However, in a medium lacking methionine, Cbl addition results in slightly faster growth compared to no corrinoid supplementation. This is seen as shorter mean doubling time of 76 minutes with concentrations of Cbl greater than 4 nM compared to a mean doubling time of 86 minutes without corrinoid supplement (Figure 3.3A). Strikingly, supplementation of 4 nM or greater CreCba causes a substantially longer mean doubling time of 115 minutes. These corrinoid-dependent growth effects were suppressed by supplementation of methionine to the cultures which strongly indicates that this is indeed specifically related to methionine synthesis (Figure 3.4B). In Chapter 2, I showed that the MetH enzyme from the closely related species *B. megaterium* can use CreCba as a cofactor for methionine synthesis. Perhaps *B. halodurans* MetH is also compatible with CreCba but cannot downregulate MetE because the riboswitch in insensitive to CreCba. I hypothesize that the growth defect of *B. halodurans* is due to overproduction of methionine by MetE and MetH. In future studies, this hypothesis can be tested by measuring the corrinoid specificity of a *B. subtilis* Δ*metE* mutant complemented with the *B. halodurans metE* and *metH* orthologs.

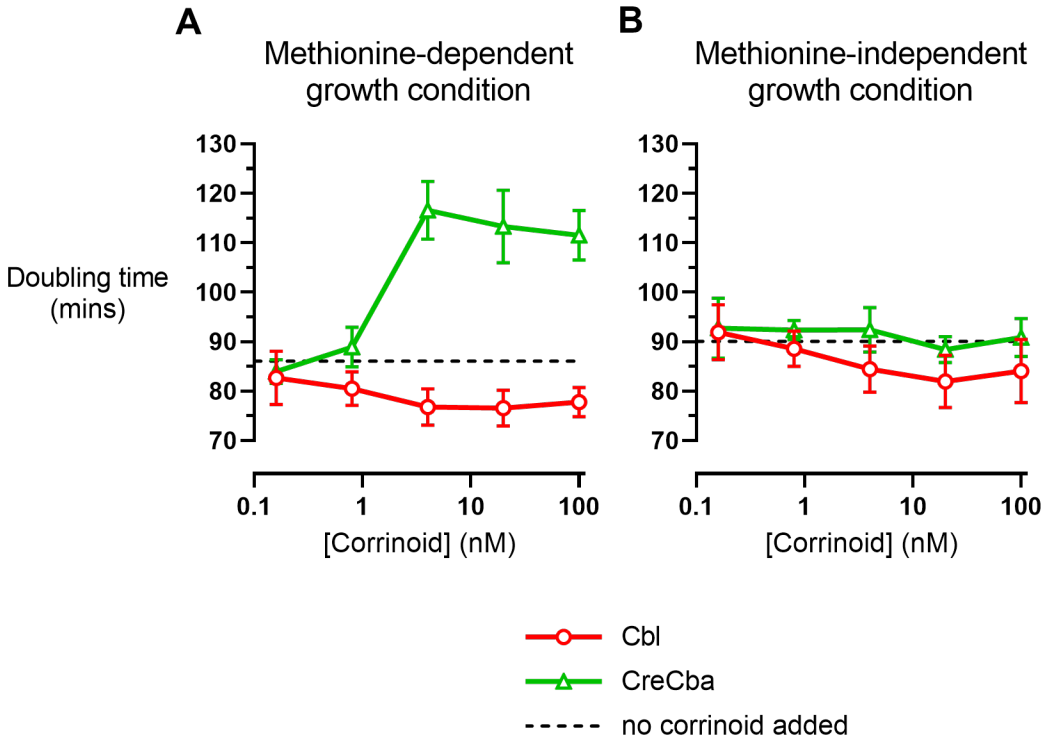


Figure 3.3 – Cbl supports and CreCba inhibits methionine-dependent growth of *B. halodurans*. *B. halodurans* was cultured in defined minimal medium supplemented with (A) 0.1% methionine, and (B) no methionine. Data points and error bars represent mean and standard deviation of four replicates.

After establishing that *B. halodurans* has corrinoid-specific metabolism, I returned to measuring the specificity of its Cbl-riboswitches in the *B. subtilis* reporter system. *B. halodurans* has five Cbl-riboswitch regulons (Table 3.2). The *metE*, *btuF* and BH0798 Cbl-riboswitches are typical semi-selective repressors that respond to Cbl and pCbl, but not CreCba or Cbi (Figure 3.4A-C). The *nrdA* Cbl-riboswitch does not respond to corrinoids in the reporter assay (Figure 3.4D). Lastly and most strikingly, the *cobT* Cbl-riboswitch upregulates *gfp* in response to all four corrinoids, suggesting that this Cbl-riboswitch is a novel promiscuous activator (Figure 3.4E).

Cbi is the most potent corrinoid for the *cobT* riboswitch, causing a 7.8-fold increase in *gfp* expression with 100 nM Cbi added. In contrast, Cbl causes a 4.5-fold increase in expression. The enzyme CobT is required for activating benzimidazolyl and purinyl lower ligands for attachment to the late corrinoid intermediate GDP-cobinamide [262, 286]. Therefore, it makes sense for the accumulation of the Cbi to act as a signal for CobT upregulation. However, it remains unclear why the *cobT* riboswitch reporter also increases *gfp* expression in response to complete corrinoids Cbl, pCbl, and CreCba (Figure 3.4E).

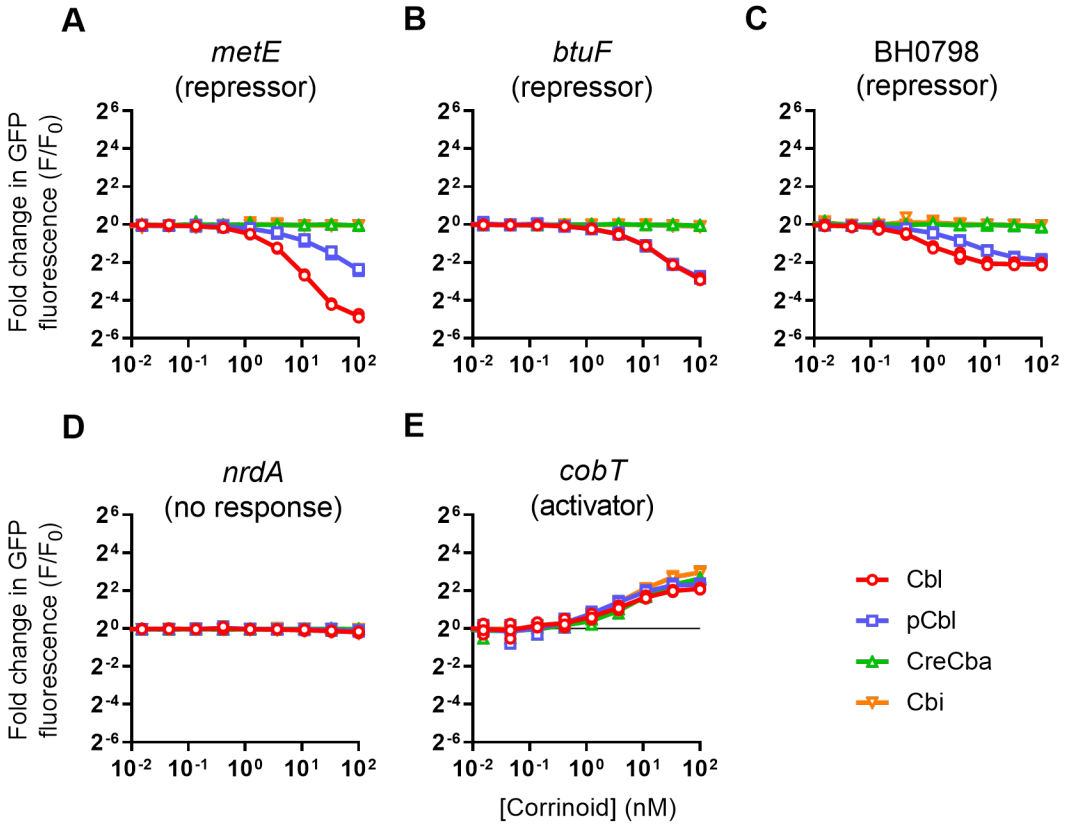


Figure 3.4 – The Cbl-riboswitches of *B. halodurans* include repressors and a novel activator. Corrinoid dose responses of fluorescent reporter strains for *B. halodurans* Cbl-riboswitches that regulate the (A) *metE*, (B) *btuF*, (C) BH0798 (acyl-CoA hydrolase), (D) *nrdA*, and (E) *cobT* operons. Thin horizontal lines demarcate no change in expression. Data points of each replicate are plotted with lines connecting mean values. N=3 for each panel.

In the purine riboswitch class, both activator and repressor mechanisms are common. Effector binding causes repression or activation of downstream genes depending on whether the bound aptamer conformation stabilizes a terminator or anti-terminator in the expression platform [49]. Thus, I analyzed the secondary structures of the *B. halodurans cobT* Cbl-riboswitch expression platform to develop a model for the activator mechanism. I predict that in the unbound state, the expression platform adopts a secondary structure containing an intrinsic transcriptional terminator (a hairpin followed by a polyuracil tract) (Figure 3.5A). In the effector bound state, a kissing loop interaction between the conserved loop 5 and a loop structure in the expression platform prevents the intrinsic terminator hairpin from forming (Figure 3.5B). In typical Cbl-riboswitches, the anti-terminator structure is favored in the effector-unbound state, whereas the kissing loop in effector bound state stabilizes the intrinsic terminator hairpin structure (Figure 3.6). A fellow graduate student in the Taga Lab, Rebecca Procknow, is continuing the investigation of this novel activation mechanism as part of her ongoing dissertation research. In future experiments, it would be interesting to systematically examine expression platform secondary structures of other Cbl-riboswitches, especially those that regulate *cobT* or other lower ligand synthesis and attachment genes, to identify other activating riboswitches.

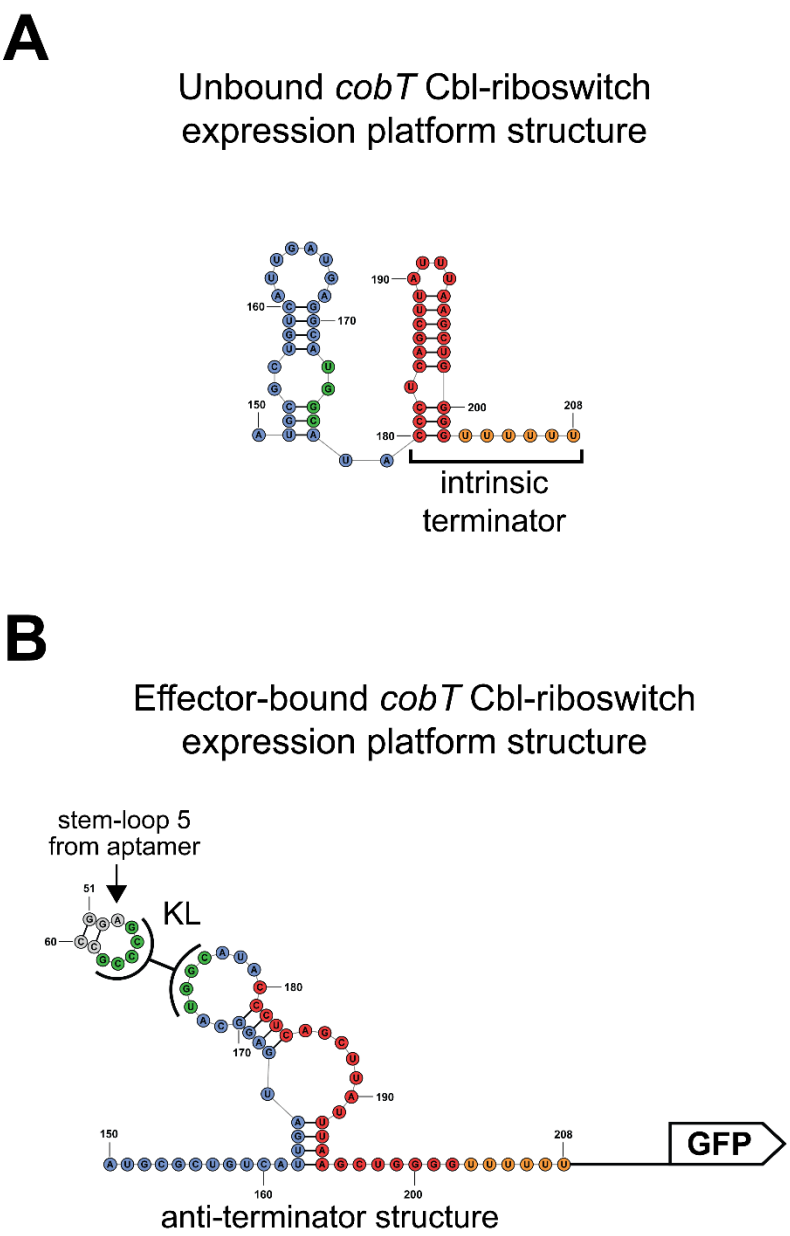


Figure 3.5 – Predicted regulatory structures in the expression platform of the *B. halodurans cobT* Cbl-riboswitch, a novel activator. (A) Expression platform secondary structure of the unbound state. (B) Expression platform secondary structure of the effector bound state. The stem-loop 5 of the aptamer is displayed to illustrate the kissing loop interaction. Red residues indicate intrinsic terminator hairpin. Orange residues indicate polyuracil tract of the intrinsic terminator structure. Blue residues indicate expression platform sequence preceding the terminator. Green residues form the kissing loop.

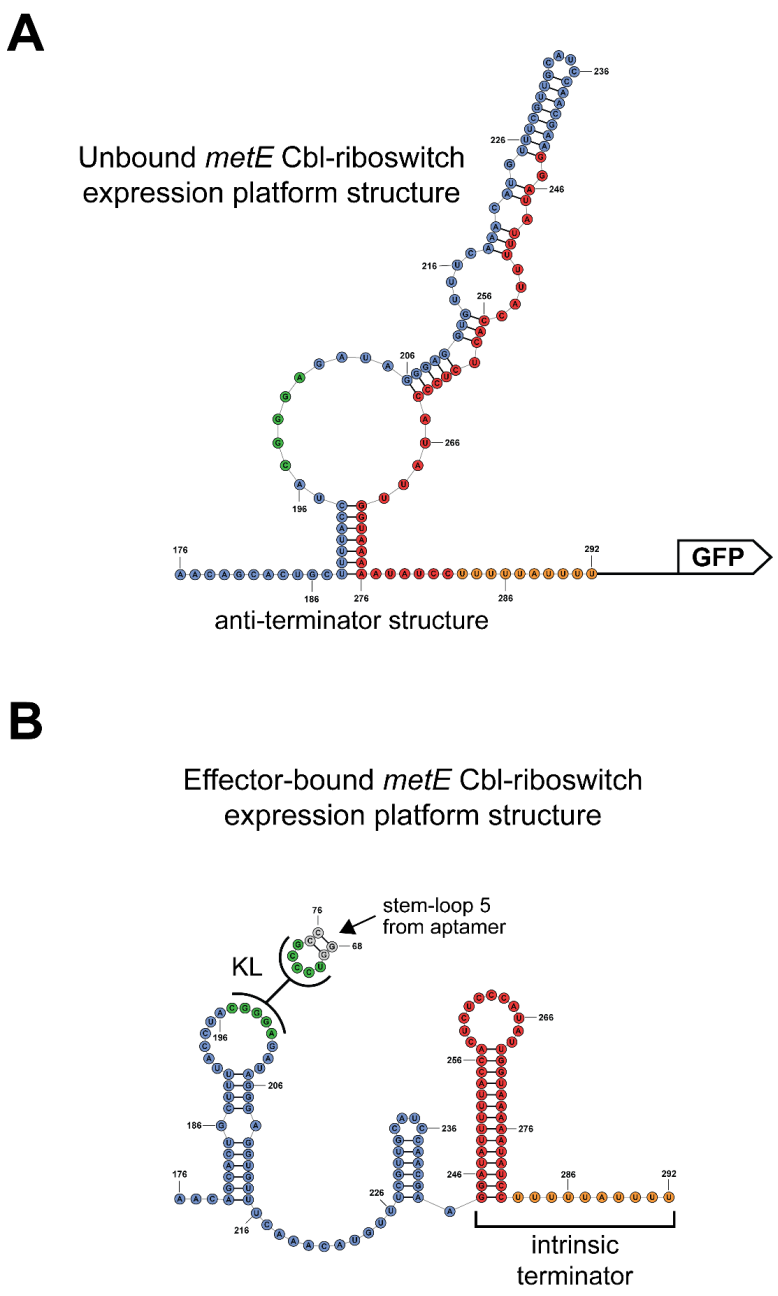


Figure 3.6 – Predicted regulatory structures in the expression platform of the *B. halodurans metE* Cbl-riboswitch, a typical repressor. (A) Expression platform secondary structure of the unbound state. (B) Expression platform secondary structure of the effector bound state. The stem-loop 5 of the aptamer is displayed to illustrate the kissing loop interaction. Red residues indicate intrinsic terminator hairpin. Orange residues indicate polyuracil tract of the intrinsic terminator structure. Blue residues indicate expression platform sequence preceding the terminator. Green residues form the kissing loop.

3.3.3 Modularity of tandem doublet Cbl-riboswitches

Let us return to the *D. propionicus mutA* riboswitch to consider the implications of tandem Cbl-riboswitch configurations. Fusing two riboswitches that respond to the same effector in series is predicted to increase the dynamic range of the regulatory response and potentially produce a more switch-like response [242, 281, 287]. However, these effects have not been experimentally tested for any tandem doublet Cbl-riboswitches. In fact, other than the glycine riboswitches, very few studies have examined the modularity of naturally occurring tandem riboswitches in great detail [48, 242, 278].

I first chose a set of tandem Cbl-riboswitches to examine in the *in vivo* reporter system. Of 3,313 Cbl-riboswitches listed in the RiboD database, 38 (1.1%) are predicted to be in a tandem configuration [47]. Four of these Cbl-riboswitches are paired with a riboswitch of a different class. The remaining 34 occur in tandem doublet Cbl-riboswitches. I also identified five tandem doublet Cbl-riboswitches in the genome of *S. ovata*, a species which was not included in the RiboD database. Tandem Cbl-riboswitches regulate typical gene targets such as corrinoid biosynthesis, corrinoid transport, and corrinoid-independent isoenzyme genes. In this regard, the *D. propionicus mutA* operon appears to be an exceptional case even among tandem Cbl-riboswitch regulated operons. I constructed a set of riboswitch reporter strains to dissect four tandem doublet Cbl-riboswitches from *D. propionicus*, *B. megaterium*, *S. ovata*, and *D. acetoxidans* (Figure 3.7).

Each of these tandem riboswitches repressed *gfp* in response to corrinoids. To examine the modularity of tandem riboswitch configuration, I measured the activity of each individual riboswitch (referred to as subunits A and B) separately in the reporter system. Each subunit A lacked an RBS and thus was fused to a strong synthetic RBS (R0) to allow translation of *gfp*. The reporter strains of the subunit B riboswitches retained their native RBS.

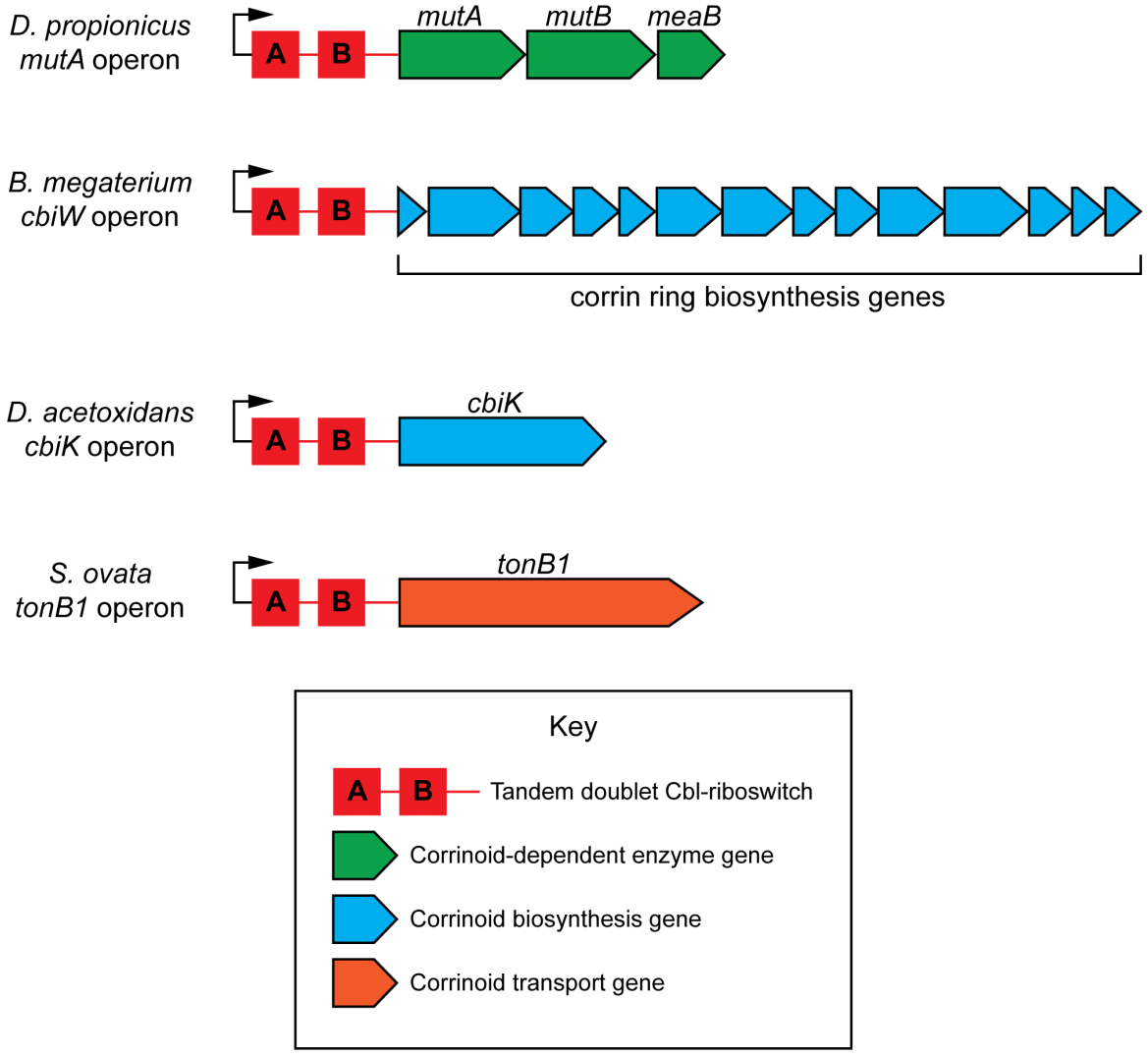


Figure 3.7 – Regulons of the tandem doublet Cbl-riboswitches examined in this study.

As previously shown, the *D. propionicus mutA* tandem riboswitch reporter represses *gfp* to different degrees with Cbl, pCbl, CreCba, and Cbi (Figure 3.1). Interestingly, subunits A and B of the *D. propionicus mutA* riboswitch have distinct corrinoid selectivities. Subunit A responds strongly to Cbl, pCbl and Cbi, and weakly to CreCba. In contrast, subunit B is a semi-selective riboswitch responding only to Cbl and to a lesser degree pCbl (Figure 3.8). If the subunits can fully function independently of each other, then the tandem riboswitch configuration should behave in a predictable additive manner by multiplying the regulatory activities of the individual subunits. Indeed, the calculated combined activities of the *D. propionicus mutA* subunits A and B closely match their measured activity in tandem (Figure 3.8). This demonstrates that the modular arrangement of two singlet Cbl-riboswitches with distinct activities can be combined to produce a single predictable output in an additive fashion. A similar additive combined activity of riboswitch subunits A and B is seen in the *B. megaterium cbiW* riboswitch (Figure 3.9). However, both of these subunits are corrinoid semi-selective types, responding only to Cbl and pCbl. It appears that this tandem riboswitch arrangement achieves a greater dynamic range of repression and likely provides strong feedback inhibition for Cbl biosynthesis in *B. megaterium* [185].

Interestingly, the *S. ovata tonB1* and *D. acetoxidans cbiK* tandem riboswitch activities are not predicted by the combined activities of their subunit riboswitches (Figures 3.10 and 3.11). The response to each corrinoid in the tandem riboswitch is much greater than predicted. However, a major limitation of these results is that the inconsistencies between the predicted and actual tandem riboswitch activities may reflect an artificial loss of activity caused by separating the singlet subunits. This is particularly important in the case of the *D. acetoxidans* tandem riboswitch in which both subunits alone appear to be nearly nonfunctional (Figure 3.11). In follow up experiments, it will be useful to introduce point mutations to the corrinoid binding sites of the aptamers in the tandem doublet riboswitch. This should rule out any potential artifacts of truncating the riboswitches. This approach would be particularly helpful to determine whether the subunits of the *D. acetoxidans* and *S. ovata* tandem riboswitches operate in a novel complex, non-additive manner.

Still, the tandem configurations of *D. propionicus mutA* and *B. megaterium cbiW* riboswitches behave in a predictable additive fashion based on the activities of their individual subunits. This suggests that it may be possible to engineer complex corrinoid-dependent gene regulation using the diverse functionalities of the natural occurring singlet Cbl-riboswitches. As a proof of principle, I propose future experiments to construct synthetic tandem riboswitches to create new regulatory outputs. I predict that fusing the *B. halodurans cobT* activator riboswitch with the *B. halodurans metE* riboswitch will create a regulator with robust repression in response to Cbl and activation in response to CreCba and Cbi (Figure 3.12A). In contrast, fusing the *B. halodurans cobT* with the *B. halodurans btuF* riboswitch should effectively create a semi-selective activator for Cbi and CreCba (Figure 3.12B). It is challenging to envision a direct approach to generate effector-selective activation and repression simultaneously in a singlet riboswitch. Thus, given the potential sophistication enabled by tandem riboswitch configurations, I argue that sampling the activities of naturally occurring riboswitches of all classes may be worthwhile for synthetic biology and engineering purposes.

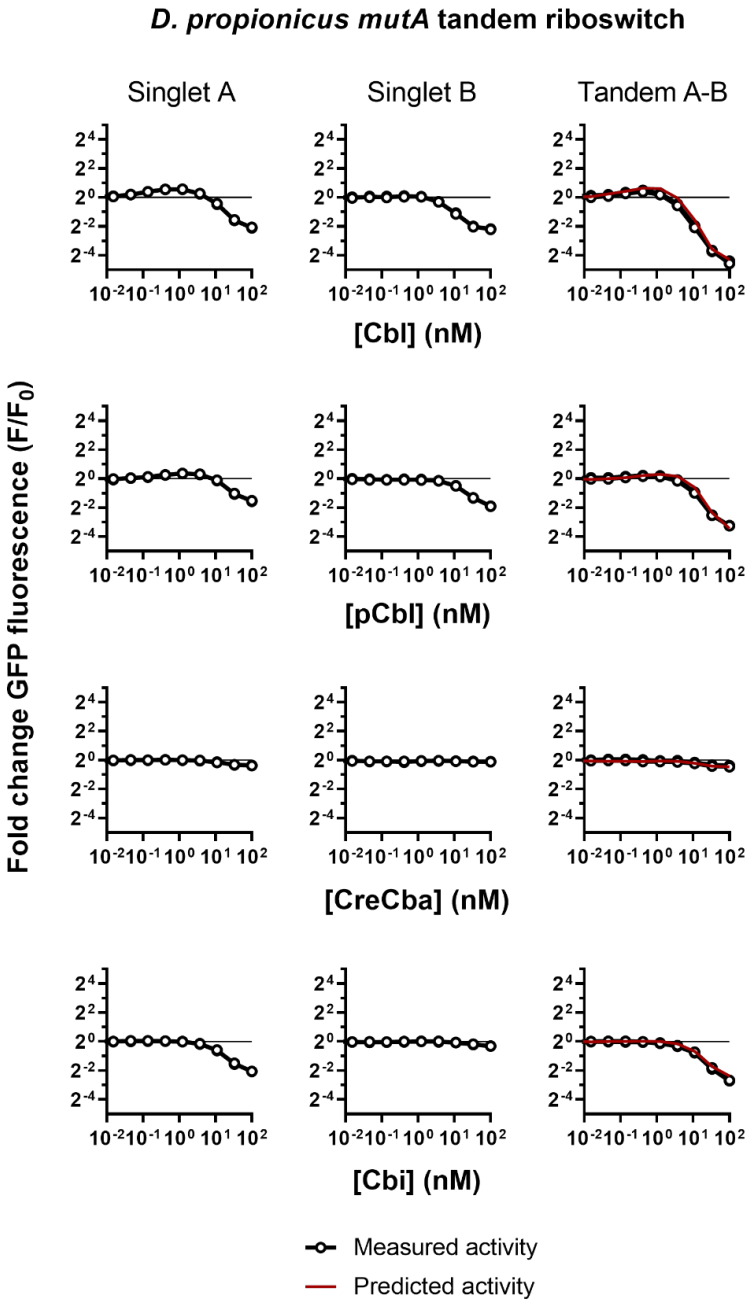


Figure 3.8 – Dissection of the *D. propionicus mutA* tandem doublet Cbl-riboswitch. Thin horizontal lines demarcate no change in expression. Data points of each replicate are plotted with lines connecting mean values. N=2 for each panel.

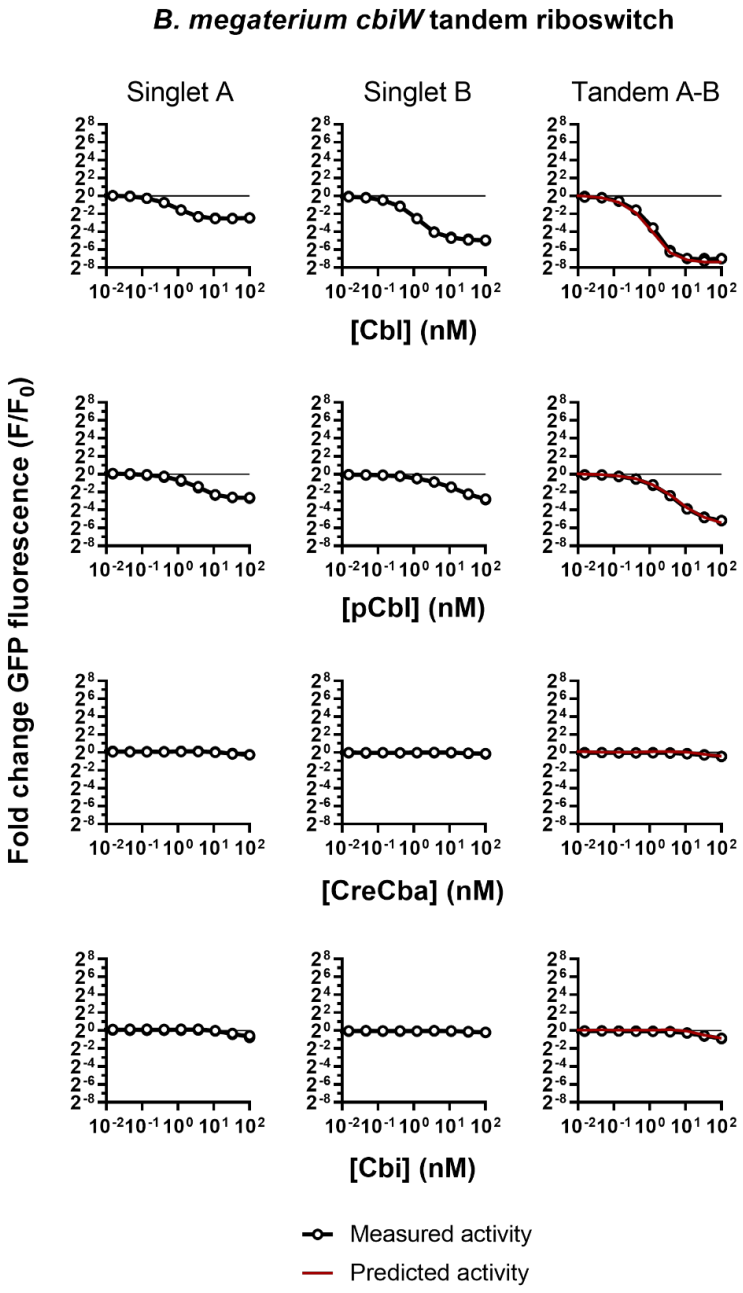


Figure 3.9 – Dissection of the *B. megaterium cbiW* tandem doublet Cbl-riboswitch. Data points of each replicate are plotted with lines connecting mean values. N=2 for each panel.

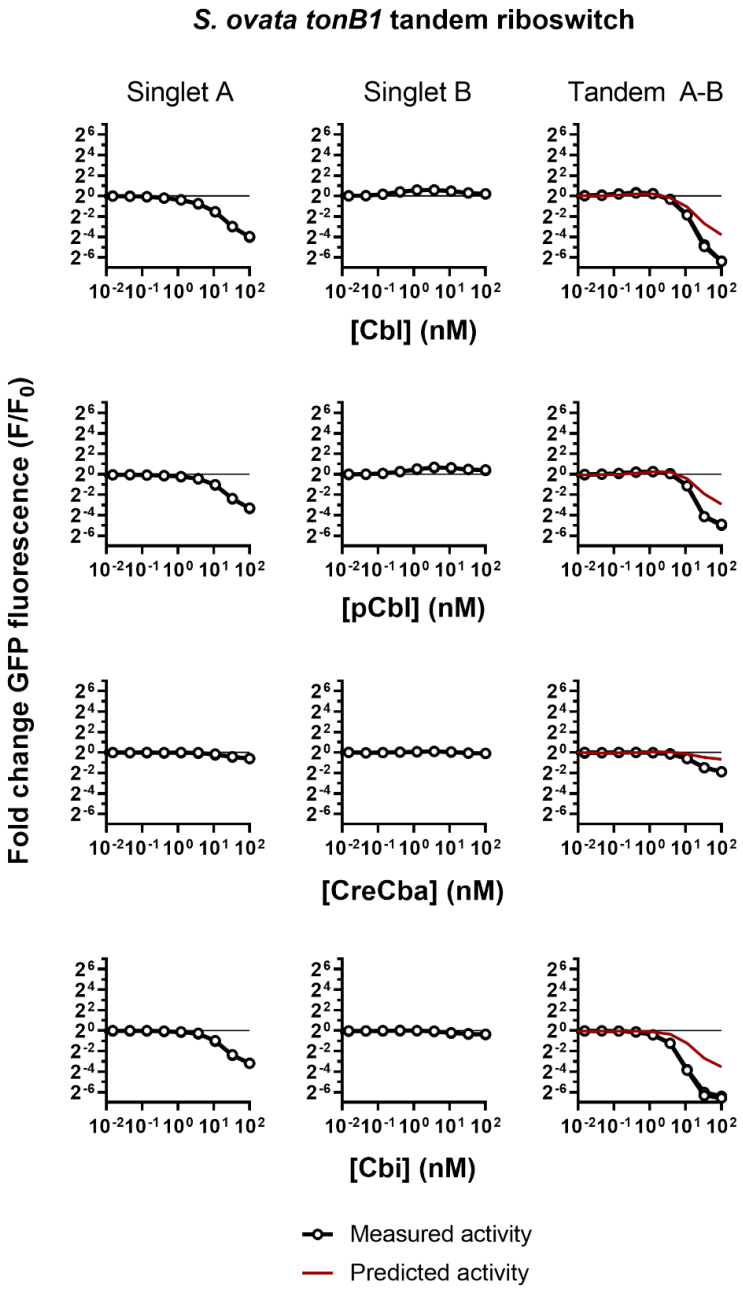


Figure 3.10 – Dissection of the *S. ovata tonB1* tandem doublet Cbl-riboswitch. Data points of each replicate are plotted with lines connecting mean values. N=2 for each panel.

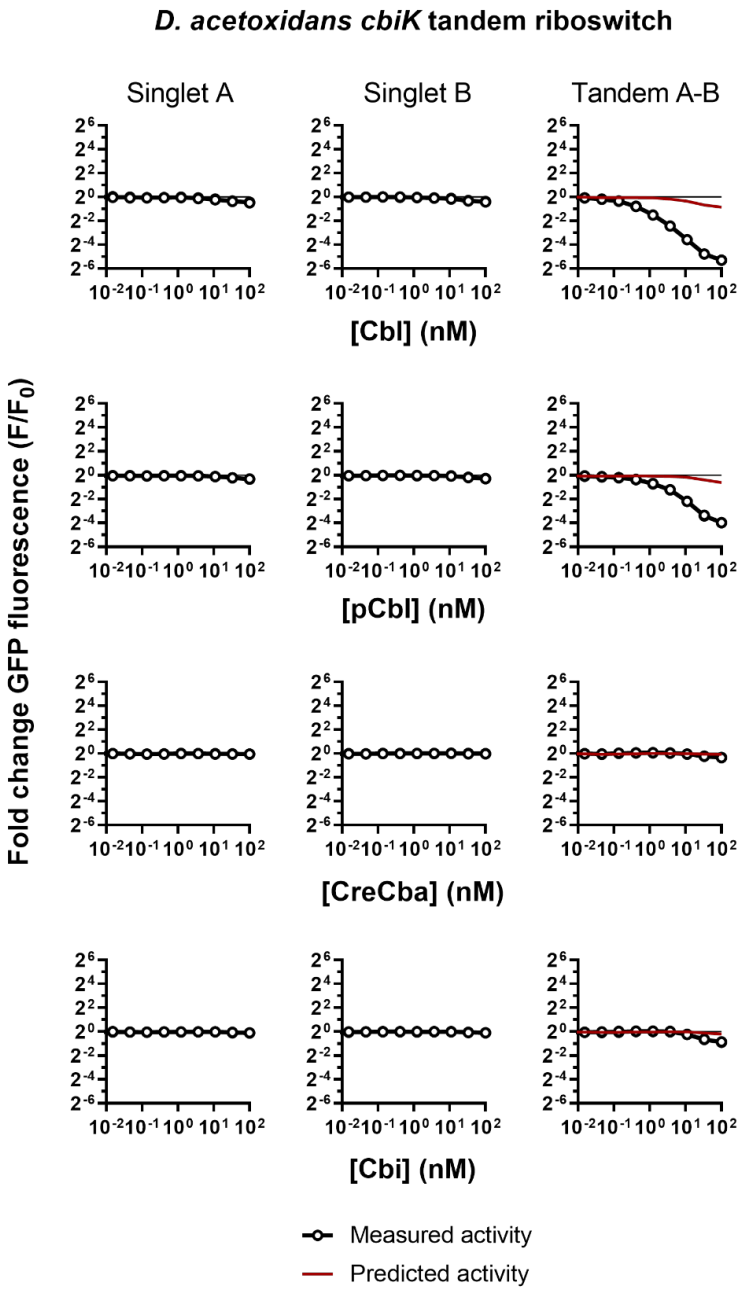


Figure 3.11 – Dissection of the *D. acetoxidans cbiK* tandem doublet Cbl-riboswitch. Data points of each replicate are plotted with lines connecting mean values. N=2 for each panel.

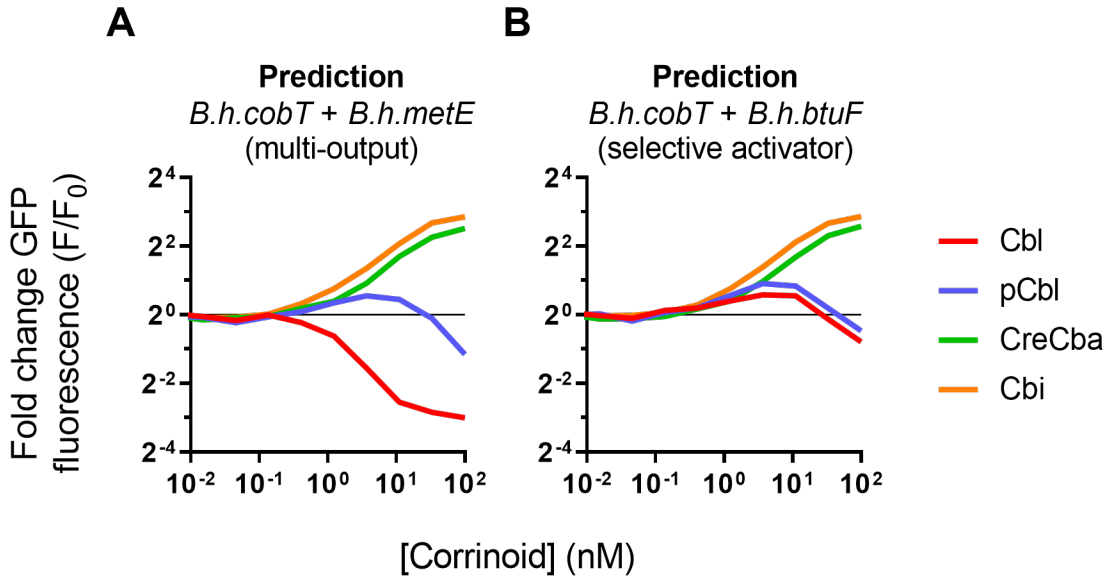


Figure 3.12 – Predicted activities of synthetic tandem doublet Cbl-riboswitches. (A) Tandem fusion of the *B. halodurans cobT* and *metE* Cbl-riboswitches. (B) Tandem fusion of the *B. halodurans cobT* and *btuF* Cbl-riboswitches.

3.4 Chapter 3 summary

In this study, I explore novel functionality among Cbl-riboswitches by examining atypical Cbl-riboswitch regulon configurations. In the course of these experiments, I propose new hypotheses about the physiological roles of Cbl-riboswitch regulation of MCM genes. I also described a Cbl-riboswitch that operates as an activator in response to corrinoids, which to my knowledge is the first ever reported. I end with a dissection of modular configurations of tandem Cbl-riboswitches and propose simple ways that artificial tandem configurations can produce complex modes of regulation. Together, these results demonstrate the potential utility of studying rare cases to gain novel biological insight.

In the previous chapter, the primary goal was to measure corrinoid specificity among several Cbl-riboswitches to potentially gain generalizable knowledge about how Cbl-riboswitches respond to chemically diverse effectors. In this chapter, I focused on Cbl-riboswitches with seemingly exceptional features in order to extend our knowledge of the functional diversity that exists among Cbl-riboswitches. The study of exceptional cases is a viable research approach which has been successfully applied to riboswitch gene regulation in several contexts.

Most of the results presented here are preliminary leads for new projects. Also, any results suggesting novel functionality like the activator Cbl-riboswitch needs to be confirmed in the organism from which it derived to make a strong claim about the role of that riboswitch in that organism's physiology. Likewise, any application of a riboswitch for synthetic biology would need to be validated in the engineered organism of interest. Nevertheless, the results from this reporter

system are proving useful for gaining insight into basic principles of riboswitch functionality, and their potential application to engineering of *B. subtilis*.

Among the many Cbl-riboswitch sequences that exist across bacterial genomes, there are likely more functional properties of both Cbl-riboswitches and the genes that they regulate. Further exploration of Cbl-riboswitch regulons will likely lead to new facets of corrinoid metabolism in bacteria.

3.5 Materials and methods

The following procedures were done as described in Chapter 2:

a) Cbl-riboswitch sequence analysis
b) Corrinoid production, extraction, purification, and analysis
c) Plasmid and strain construction
d) Riboswitch fluorescent reporter assays

3.5.1 *MCM-dependent growth of* B. halodurans *C-125*

B. halodurans was streaked from frozen stocks onto LB agar supplemented with 0.1 M sodium sesquicarbonate and incubated overnight at 30 ºC for 14-18 hours. Single colonies were used to inoculate 2 mL liquid starter cultures containing Spizizen minimal medium supplemented with 0.1 M sodium sesquicarbonate (alkaline SMM) without any carbon source. Starter cultures were incubated shaking overnight (250 rpm, 30 ºC) for 16 hours, reaching cell density of about OD_{600} = 0.1. Starter cultures were diluted 100-fold in 8 mL of alkaline SMM supplemented with 0.04% L-Isoleucine and 0.04% L-valine as the carbon source. 75 µL of diluted culture was dispensed into wells of a 96-well microtiter plate. Each well was further supplemented with 75 µL of alkaline SMM containing various concentrations of Cbl, pCbl, CreCba or Cbi. Plates were sealed with gas diffusible membranes (Breathe-Easy, MilliporeSigma) and incubated at 30 ºC in a high-speed plate shaker. After 48 hours, final cell densities were measured on a BioTek Synergy2 plate reader.

3.5.2 *Methionine-dependent growth of* B. halodurans *C-125*

B. halodurans was streaked from frozen stocks onto LB agar supplemented with 0.1 M sodium sesquicarbonate and incubated overnight at 30 ºC for 14-18 hrs. Single colonies were used to inoculate 2 mL starter cultures of alkaline SMM with 30 mM sodium citrate as a carbon source. Starter cultures were incubated shaking overnight (250 rpm, 30 ºC) for 14-18 hours, reaching cell density of roughly OD_{600} = 1. Starter cultures were diluted to OD600 = 0.01 in 10 mL of alkaline SMM with 30 mM sodium citrate. 75 µL of diluted culture was dispensed into wells of a 96-well microtiter plate. Each well was further supplemented with 75 µL of alkaline SMM containing various concentrations of Cbl and CreCba, with and without 0.1% L-methionine. Plates were sealed with gas diffusible membranes and incubated at 30 ºC in a BioTek Synergy2 plate reader set to ‘high-speed’ shaking. Absorbance at 600 nm was measured every 20 minutes for 60 hours. Growth kinetics were analyzed and plotted using GraphPad Prism 9.

Chapter 4 – Summary

Riboswitches are extraordinary for the degree of efficiency by which they transduce a molecular sensory input to a gene expression output. Aside from the essential transcription and translation machinery, there are no other protein, RNA or other biomolecular factors required to implement this type of genetic switch. A riboswitch fully encapsulates the principle of the structure-function relationship. Through my dissertation research, I have grown to appreciate the simultaneous simplicity and sophistication of this pervasive control mechanism in bacteria.

My work has focused on the Cbl-riboswitch. In my opinion, the relatively small field of riboswitch-based gene regulation has greatly benefited from being heavily populated by chemists and biophysicists since its inception. Quantitative *in vitro* biochemical and structural approaches have propelled our understanding of these complex molecular machines to the point that fully artificial riboswitches are actively developed for biotechnological purposes. While the edges of this field have rapidly expanded over the past twenty years, this has left many opportunities for general biologists (like me) to attempt to resolve some of the more curious biological aspects of these naturally evolved control systems.

Despite being the first type of riboswitch discovered over twenty years ago, basic aspects of Cbl-riboswitch functionality have remained uninvestigated. In particular, the Cbl-riboswitch is unmatched by other riboswitches in terms of the natural chemical diversity of its potential effector molecules. Given the widespread importance of corrinoids in microbial communities, it is remarkable how little we know about the bacterial gene regulatory response to corrinoids other than Cbl. My research represents an early step to address this fundamental problem. I can confidently conclude from my studies that Cbl-riboswitches respond to multiple types of corrinoids. Furthermore, Cbl-riboswitches vary in their corrinoid specificity. This alone should provide sufficient motivation to consider the role of Cbl-riboswitches in any future work that examines the impacts of Cbl and other corrinoids on microbial communities.

Based on my results in Chapter 2, I proposed that the propensity of a corrinoid to adopt either the base-ON or base-OFF state is the key molecular feature that Cbl-riboswitches detect to produce distinct regulatory outputs. Future *in vitro* biochemical approaches will be useful for testing this model. Structural studies of semi-selective and promiscuous aptamers in complex with various corrinoids should yield further mechanistic insight into how bacteria differentially sense corrinoid cofactors. Complementary *in vitro* binding assays and *in vitro* transcription/translation assays can further disentangle the potentially complex role of the kissing loop interaction in corrinoid specificity. In addition, new technologies are emerging that repurpose the frameworks of highly parallelized sequencing platforms to conduct highly parallelized sequence probing and biochemical analyses of arrayed RNAs [288]. The complexity of sequence, structure, and functionality inherent to Cbl-riboswitches makes them an attractive subject for these emerging approaches.

In my studies in Chapters 2 and 3, I examined the physiological implications of corrinoid specific gene regulation. These experiments provide proof of principle and hypotheses about the matching of corrinoid specificity of gene regulation and physiology. In future efforts, it will be important to

examine corrinoid-specific gene regulation within the native organisms for at least some riboswitches in order to corroborate some of my major findings. However, there will likely be major challenges to implementing and interpreting these types of experiments. Corrinoid-dependent organisms often possess multiple corrinoid-dependent processes which are further regulated by Cbl-riboswitches. This complexity makes it difficult to predict how an organism will respond to a perturbation to its corrinoid-dependent physiology or gene regulation. As an alternative approach to this problem, it may be possible to engineer corrinoid-dependent physiology within a corrinoid-independent organism in order to precisely dissect the interplay between corrinoid-dependent processes.

More broadly, a puzzling question about corrinoids remains: Why do so many corrinoids exist? For most of the other organic cofactors like SAM, THF, and flavins, it appears that nature has settled on a single or small number of molecules to carry out essential biological reactions. What drove some organisms to preferentially produce and use distinct corrinoids? Like most ‘why?’ questions, these are likely to remain unanswered by scientific approaches, but will continue to capture our scientific interests.

References

1. Crick, F., *Central dogma of molecular biology.* Nature, 1970. **227**(5258): p. 561-3.
2. Jacob, F. and J. Monod, *[Genes of structure and genes of regulation in the biosynthesis of proteins].* C R Hebd Seances Acad Sci, 1959. **249**: p. 1282-4.
3. Jacob, F. and J. Monod, *Genetic regulatory mechanisms in the synthesis of proteins.* J Mol Biol, 1961. **3**: p. 318-56.
4. Gilbert, W. and B. Müller-Hill, *Isolation of the lac repressor.* Proc Natl Acad Sci U S A, 1966. **56**(6): p. 1891-8.
5. Gilbert, W. and B. Müller-Hill, *The lac operator is DNA.* Proc Natl Acad Sci U S A, 1967. **58**(6): p. 2415-21.
6. Bervoets, I. and D. Charlier, *Diversity, versatility and complexity of bacterial gene regulation mechanisms: opportunities and drawbacks for applications in synthetic biology.* FEMS Microbiology Reviews, 2019. **43**(3): p. 304-339.
7. Hirschman, J., et al., *Products of nitrogen regulatory genes ntrA and ntrC of enteric bacteria activate glnA transcription in vitro: evidence that the ntrA product is a sigma factor.* Proc Natl Acad Sci U S A, 1985. **82**(22): p. 7525.
8. Merrick, M.J., *In a class of its own — the RNA polymerase sigma factor σ;54 (σN).* Molecular Microbiology, 1993. **10**(5): p. 903-909.
9. Gruber, T.M. and C.A. Gross, *Multiple Sigma Subunits and the Partitioning of Bacterial Transcription Space.* Annual Review of Microbiology, 2003. **57**(1): p. 441-466.
10. Eames, M. and T. Kortemme, *Cost-benefit tradeoffs in engineered lac operons.* Science, 2012. **336**(6083): p. 911-915.
11. Wilson, K.S. and P.H. von Hippel, *Transcription termination at intrinsic terminators: the role of the RNA hairpin.* Proc Natl Acad Sci U S A, 1995. **92**(19): p. 8793-7.
12. Peters, J.M., A.D. Vangeloff, and R. Landick, *Bacterial transcription terminators: the RNA 3'-end chronicles.* J Mol Biol, 2011. **412**(5): p. 793-813.
13. Gusarov, I. and E. Nudler, *The mechanism of intrinsic transcription termination.* Mol Cell, 1999. **3**(4): p. 495-504.
14. Mandell, Z.F., et al., *NusG is an intrinsic transcription termination factor that stimulates motility and coordinates gene expression with NusA.* Elife, 2021. **10**.
15. Naville, M. and D. Gautheret, *Premature terminator analysis sheds light on a hidden world of bacterial transcriptional attenuation.* Genome Biol, 2010. **11**(9): p. R97.
16. Yanofsky, C., *Attenuation in the control of expression of bacterial operons.* Nature, 1981. **289**(5800): p. 751-758.
17. Gollnick, P., et al., *Complexity in Regulation of Tryptophan Biosynthesis in Bacillus subtilis.* Annual Review of Genetics, 2005. **39**(1): p. 47-68.
18. Naville, M. and D. Gautheret, *Transcription attenuation in bacteria: theme and variations.* Brief Funct Genomics, 2010. **9**(2): p. 178-89.
19. Laursen, B.S., et al., *Initiation of protein synthesis in bacteria.* Microbiol Mol Biol Rev, 2005. **69**(1): p. 101-23.
20. Simonetti, A., et al., *Structure of the 30S translation initiation complex.* Nature, 2008. **455**(7211): p. 416-420.
21. Simonetti, A., et al., *A structural view of translation initiation in bacteria.* Cell Mol Life Sci, 2009. **66**(3): p. 423-36.

22. Jacob, W.F., M. Santer, and A.E. Dahlberg, *A single base change in the Shine-Dalgarno region of 16S rRNA of Escherichia coli affects translation of many proteins.* Proc Natl Acad Sci U S A, 1987. **84**(14): p. 4757.
23. Duval, M., et al., *Multiple ways to regulate translation initiation in bacteria: Mechanisms, regulatory circuits, dynamics.* Biochimie, 2015. **114**: p. 18-29.
24. Samatova, E., et al., *Translational Control by Ribosome Pausing in Bacteria: How a Non-uniform Pace of Translation Affects Protein Production and Folding.* Frontiers in Microbiology, 2021. **11**(3428).
25. Storz, G., J. Vogel, and K.M. Wassarman, *Regulation by small RNAs in bacteria: expanding frontiers.* Mol Cell, 2011. **43**(6): p. 880-91.
26. Vogel, J. and B.F. Luisi, *Hfq and its constellation of RNA.* Nat Rev Microbiol, 2011. **9**(8): p. 578-89.
27. Lenz, D.H., et al., *The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in Vibrio harveyi and Vibrio cholerae.* Cell, 2004. **118**(1): p. 69-82.
28. Valentin-Hansen, P., M. Eriksen, and C. Udesen, *The bacterial Sm-like protein Hfq: a key player in RNA transactions.* Mol Microbiol, 2004. **51**(6): p. 1525-33.
29. Chao, Y. and J. Vogel, *The role of Hfq in bacterial pathogens.* Curr Opin Microbiol, 2010. **13**(1): p. 24-33.
30. Gottesman, S. and G. Storz, *Bacterial small RNA regulators: versatile roles and rapidly evolving variations.* Cold Spring Harb Perspect Biol, 2011. **3**(12).
31. Djapgne, L. and A.G. Oglesby, *Impacts of Small RNAs and Their Chaperones on Bacterial Pathogenicity.* Front Cell Infect Microbiol, 2021. **11**: p. 604511.
32. Frieda, K.L. and S.M. Block, *Direct observation of cotranscriptional folding in an adenine riboswitch.* Science, 2012. **338**(6105): p. 397-400.
33. Belogurov, G.A. and I. Artsimovitch, *Regulation of Transcript Elongation.* Annual Review of Microbiology, 2015. **69**(1): p. 49-69.
34. Chauvier, A., et al., *Transcriptional pausing at the translation start site operates as a critical checkpoint for riboswitch regulation.* Nature Communications, 2017. **8**(1): p. 13892.
35. Chauvier, A., et al., *Monitoring RNA dynamics in native transcriptional complexes.* Proc Natl Acad Sci U S A, 2021. **118**(45): p. e2106564118.
36. Breaker, R.R., *Riboswitches and Translation Control.* Cold Spring Harb Perspect Biol, 2018. **10**(11).
37. Hollands, K., et al., *Riboswitch control of Rho-dependent transcription termination.* Proc Natl Acad Sci U S A, 2012. **109**(14): p. 5376.
38. Johnson, G.E., et al., *Functionally uncoupled transcription-translation in Bacillus subtilis.* Nature, 2020. **585**(7823): p. 124-128.
39. Breaker, R.R., *Complex riboswitches.* Science, 2008. **319**(5871): p. 1795-7.
40. Winkler, W.C., et al., *Control of gene expression by a natural metabolite-responsive ribozyme.* Nature, 2004. **428**(6980): p. 281-6.
41. Cheah, M.T., et al., *Control of alternative RNA splicing and gene expression by eukaryotic riboswitches.* Nature, 2007. **447**(7143): p. 497-500.
42. Wachter, A., et al., *Riboswitch control of gene expression in plants by splicing and alternative 3′ end processing of mRNAs.* Plant Cell, 2007. **19**(11): p. 3437-50.
43. Li, S. and R.R. Breaker, *Eukaryotic TPP riboswitch regulation of alternative splicing involving long-distance base pairing.* Nucleic Acids Res, 2013. **41**(5): p. 3022-31.

44. Mellin, J.R., et al., *A riboswitch-regulated antisense RNA in Listeria monocytogenes.* Proc Natl Acad Sci U S A, 2013. **110**(32): p. 13132-7.
45. DebRoy, S., et al., *Riboswitches. A riboswitch-containing sRNA controls gene expression by sequestration of a response regulator.* Science, 2014. **345**(6199): p. 937-40.
46. Mellin, J.R., et al., *Riboswitches. Sequestration of a two-component response regulator by a riboswitch-regulated noncoding RNA.* Science, 2014. **345**(6199): p. 940-3.
47. Mukherjee, S., et al., *RiboD: a comprehensive database for prokaryotic riboswitches.* Bioinformatics, 2019. **35**(18): p. 3541-3543.
48. Zhou, H., et al., *Characterization of a natural triple-tandem c-di-GMP riboswitch and application of the riboswitch-based dual-fluorescence reporter.* Sci Rep, 2016. **6**: p. 20871.
49. Batey, R.T., *Structure and mechanism of purine-binding riboswitches.* Q Rev Biophys, 2012. **45**(3): p. 345-81.
50. Mandal, M., et al., *Riboswitches control fundamental biochemical pathways in Bacillus subtilis and other bacteria.* Cell, 2003. **113**(5): p. 577-586.
51. Batey, R.T., S.D. Gilbert, and R.K. Montange, *Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine.* Nature, 2004. **432**(7015): p. 411-5.
52. Serganov, A., et al., *Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs.* Chem Biol, 2004. **11**(12): p. 1729-41.
53. Nygaard, P. and H.H. Saxild, *The purine efflux pump PbuE in Bacillus subtilis modulates expression of the PurR and G-box (XptR) regulons by adjusting the purine base pool size.* J Bacteriol, 2005. **187**(2): p. 791-4.
54. Wickiser, J.K., et al., *The kinetics of ligand binding by an adenine-sensing riboswitch.* Biochemistry, 2005. **44**(40): p. 13404-14.
55. Johansen, L.E., et al., *Definition of a second Bacillus subtilis pur regulon comprising the pur and xpt-pbuX operons plus pbuG, nupG (yxjA), and pbuE (ydhL).* J Bacteriol, 2003. **185**(17): p. 5200-9.
56. Reader, J.S., et al., *Identification of four genes necessary for biosynthesis of the modified nucleoside queuosine.* J Biol Chem, 2004. **279**(8): p. 6280-5.
57. Payne, K.A., et al., *Epoxyqueuosine Reductase Structure Suggests a Mechanism for Cobalamin-dependent tRNA Modification.* J Biol Chem, 2015. **290**(46): p. 27572-81.
58. Zallot, R., et al., *Plant, animal, and fungal micronutrient queuosine is salvaged by members of the DUF2419 protein family.* ACS Chem Biol, 2014. **9**(8): p. 1812-25.
59. Fergus, C., et al., *The queuine micronutrient: charting a course from microbe to man.* Nutrients, 2015. **7**(4): p. 2897-929.
60. Barrick, J.E. and R.R. Breaker, *The distributions, mechanisms, and structures of metabolite-binding riboswitches.* Genome Biology, 2007. **8**(11): p. R239.
61. Eichhorn, C.D., M. Kang, and J. Feigon, *Structure and function of preQ(1) riboswitches.* Biochim Biophys Acta, 2014. **1839**(10): p. 939-950.
62. Roth, A., et al., *A riboswitch selective for the queuosine precursor preQ1 contains an unusually small aptamer domain.* Nat Struct Mol Biol, 2007. **14**(4): p. 308-17.
63. Meyer, M.M., et al., *Confirmation of a second natural preQ1 aptamer class in Streptococcaceae bacteria.* RNA, 2008. **14**(4): p. 685-95.
64. McCown, P.J., et al., *Riboswitch diversity and distribution.* RNA, 2017. **23**(7): p. 995-1011.

65. Dowling, D.P., et al., *Molecular basis of cobalamin-dependent RNA modification.* Nucleic Acids Res, 2016. **44**(20): p. 9965-9976.
66. Zallot, R., et al., *Identification of a Novel Epoxyqueuosine Reductase Family by Comparative Genomics.* ACS Chem Biol, 2017. **12**(3): p. 844-851.
67. Li, Q., et al., *Epoxyqueuosine Reductase QueH in the Biosynthetic Pathway to tRNA Queuosine Is a Unique Metalloenzyme.* Biochemistry, 2021. **60**(42): p. 3152-3161.
68. Noguchi, S., et al., *Isolation and characterization of an Escherichia coli mutant lacking tRNA-guanine transglycosylase. Function and biosynthesis of queuosine in tRNA.* J Biol Chem, 1982. **257**(11): p. 6544-50.
69. Kim, P.B., J.W. Nelson, and R.R. Breaker, *An ancient riboswitch class in bacteria regulates purine biosynthesis and one-carbon metabolism.* Mol Cell, 2015. **57**(2): p. 317-28.
70. Jones, C.P. and A.R. Ferré-D'Amaré, *Recognition of the bacterial alarmone ZMP through long-distance association of two RNA subdomains.* Nat Struct Mol Biol, 2015. **22**(9): p. 679-85.
71. Trausch, J.J., et al., *Metal Ion-Mediated Nucleobase Recognition by the ZTP Riboswitch.* Chem Biol, 2015. **22**(7): p. 829-37.
72. Anderson, B.W., D.K. Fung, and J.D. Wang, *Regulatory Themes and Variations by the Stress-Signaling Nucleotide Alarmones (p)ppGpp in Bacteria.* Annu Rev Genet, 2021.
73. Sudarsan, N., et al., *Riboswitches in eubacteria sense the second messenger cyclic di-GMP.* Science, 2008. **321**(5887): p. 411-3.
74. Nelson, J.W., et al., *Riboswitches in eubacteria sense the second messenger c-di-AMP*. Nat Chem Biol, 2013. **9**(12): p. 834-9.
75. Lee, E.R., et al., *An allosteric self-splicing ribozyme triggered by a bacterial second messenger.* Science, 2010. **329**(5993): p. 845-848.
76. Kellenberger, C.A., et al., *GEMM-I riboswitches from Geobacter sense the bacterial second messenger cyclic AMP-GMP.* Proc Natl Acad Sci U S A, 2015. **112**(17): p. 5383-8.
77. Nelson, J.W., et al., *Control of bacterial exoelectrogenesis by c-AMP-GMP.* Proc Natl Acad Sci U S A, 2015. **112**(17): p. 5389-94.
78. Serganov, A. and D.J. Patel, *Amino acid recognition and gene regulation by riboswitches.* Biochim Biophys Acta, 2009. **1789**(9-10): p. 592-611.
79. Garst, A.D., et al., *Crystal structure of the lysine riboswitch regulatory mRNA element.* J Biol Chem, 2008. **283**(33): p. 22347-51.
80. Serganov, A., L. Huang, and D.J. Patel, *Structural insights into amino acid binding and gene control by a lysine riboswitch.* Nature, 2008. **455**(7217): p. 1263-7.
81. Sudarsan, N., et al., *An mRNA structure in bacteria that controls gene expression by binding lysine.* Genes Dev, 2003. **17**(21): p. 2688-97.
82. Ren, A., et al., *Structural and Dynamic Basis for Low-Affinity, High-Selectivity Binding of L-Glutamine by the Glutamine Riboswitch.* Cell Rep, 2015. **13**(9): p. 1800-13.
83. Huang, L., et al., *Structure and ligand binding of the glutamine-II riboswitch.* Nucleic Acids Research, 2019. **47**(14): p. 7666-7675.
84. Kwon, M. and S.A. Strobel, *Chemical basis of glycine riboswitch cooperativity.* RNA, 2008. **14**(1): p. 25-34.
85. Babina, A.M., N.E. Lea, and M.M. Meyer, *In Vivo Behavior of the Tandem Glycine Riboswitch in Bacillus subtilis.* mBio, 2017. **8**(5).

86. Torgerson, C.D., D.A. Hiller, and S.A. Strobel, *The asymmetry and cooperativity of tandem glycine riboswitch aptamers.* RNA, 2020. **26**(5): p. 564-580.
87. Ruff, K.M., et al., *Singlet glycine riboswitches bind ligand as well as tandem riboswitches.* RNA, 2016. **22**(11): p. 1728-1738.
88. Ruff, K.M. and S.A. Strobel, *Ligand binding by the tandem glycine riboswitch depends on aptamer dimerization but not double ligand occupancy.* RNA, 2014. **20**(11): p. 1775-88.
89. Crum, M., N. Ram-Mohan, and M.M. Meyer, *Regulatory context drives conservation of glycine riboswitch aptamers.* PLoS Comput Biol, 2019. **15**(12): p. e1007564.
90. Tezuka, T. and Y. Ohnishi, *Two glycine riboswitches activate the glycine cleavage system essential for glycine detoxification in Streptomyces griseus.* J Bacteriol, 2014. **196**(7): p. 1369-76.
91. Ferré-D'Amaré, A.R. and W.C. Winkler, *The roles of metal ions in regulation by riboswitches.* Met Ions Life Sci, 2011. **9**: p. 141-73.
92. Wedekind, J.E., et al., *Metalloriboswitches: RNA-based inorganic ion sensors that regulate genes.* J Biol Chem, 2017. **292**(23): p. 9441-9450.
93. Ramesh, A. and W.C. Winkler, *Magnesium-sensing riboswitches in bacteria.* RNA Biol, 2010. **7**(1): p. 77-83.
94. Waters, L.S., *Bacterial manganese sensing and homeostasis.* Curr Opin Chem Biol, 2020. **55**: p. 96-102.
95. Dambach, M., et al., *The ubiquitous yybP-ykoY riboswitch is a manganese-responsive regulatory element.* Mol Cell, 2015. **57**(6): p. 1099-1109.
96. Price, I.R., et al., *Mn(2+)-sensing mechanisms of yybP-ykoY orphan riboswitches.* Mol Cell, 2015. **57**(6): p. 1110-1123.
97. Furukawa, K., et al., *Bacterial riboswitches cooperatively bind Ni(2+) or Co(2+) ions and control expression of heavy metal transporters.* Mol Cell, 2015. **57**(6): p. 1088-1098.
98. Chawla, M., et al., *Structural stability, acidity, and halide selectivity of the fluoride riboswitch recognition site.* J Am Chem Soc, 2015. **137**(1): p. 299-306.
99. Ren, A., K.R. Rajashankar, and D.J. Patel, *Fluoride ion encapsulation by Mg2+ ions and phosphates in a fluoride riboswitch.* Nature, 2012. **486**(7401): p. 85-9.
100. Nudler, E. and A.S. Mironov, *The riboswitch control of bacterial metabolism.* Trends Biochem Sci, 2004. **29**(1): p. 11-7.
101. Nahvi, A., et al., *Genetic control by a metabolite binding mRNA.* Chem Biol, 2002. **9**(9): p. 1043.
102. Nahvi, A., J.E. Barrick, and R.R. Breaker, *Coenzyme B12 riboswitches are widespread genetic control elements in prokaryotes.* Nucleic Acids Res, 2004. **32**(1): p. 143-50.
103. Markham, G.D., *S-Adenosylmethionine*, in *Encylopedia of Life Sciences (ELS)*. 2010, John Wiley & Sons: Chichester.
104. Huang, L., et al., *Crystal structure and ligand-induced folding of the SAM/SAH riboswitch.* Nucleic Acids Res, 2020. **48**(13): p. 7545-7556.
105. Weickhmann, A.K., et al., *The structure of the SAM/SAH-binding riboswitch.* Nucleic Acids Res, 2019. **47**(5): p. 2654-2665.
106. Price, I.R., J.C. Grigg, and A. Ke, *Common themes and differences in SAM recognition among SAM riboswitches.* Biochim Biophys Acta, 2014. **1839**(10): p. 931-938.
107. Edwards, A.L., et al., *Structural basis for recognition of S-adenosylhomocysteine by riboswitches.* RNA, 2010. **16**(11): p. 2144-55.

108. Ghisla, S. and D.E. Edmondson, *Flavin Coenzymes*, in *Encyclopedia of Life Sciences (ELS)*. 2014, John Wiley & Sons.
109. Serganov, A., L. Huang, and D.J. Patel, *Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch.* Nature, 2009. **458**(7235): p. 233-7.
110. Wickiser, J.K., et al., *The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch.* Mol Cell, 2005. **18**(1): p. 49-60.
111. Mansjö, M. and J. Johansson, *The riboflavin analog roseoflavin targets an FMN-riboswitch and blocks Listeria monocytogenes growth, but also stimulates virulence gene-expression and infection.* RNA Biol, 2011. **8**(4): p. 674-80.
112. Lee, E.R., K.F. Blount, and R.R. Breaker, *Roseoflavin is a natural antibacterial compound that binds to FMN riboswitches and regulates gene expression.* RNA Biol, 2009. **6**(2): p. 187-94.
113. Grill, S., et al., *The bifunctional flavokinase/flavin adenine dinucleotide synthetase from Streptomyces davawensis produces inactive flavin cofactors and is not involved in resistance to the antibiotic roseoflavin.* J Bacteriol, 2008. **190**(5): p. 1546-53.
114. Grill, S., et al., *Identification and characterization of two Streptomyces davawensis riboflavin biosynthesis gene clusters.* Arch Microbiol, 2007. **188**(4): p. 377-87.
115. Pedrolli, D.B., et al., *A highly specialized flavin mononucleotide riboswitch responds differently to similar ligands and confers roseoflavin resistance to Streptomyces davawensis.* Nucleic Acids Res, 2012. **40**(17): p. 8662-73.
116. Panchal, V. and R. Brenk, *Riboswitches as Drug Targets for Antibiotics.* Antibiotics (Basel), 2021. **10**(1).
117. Kräutler, B., *Antivitamins B12 - Some Inaugural Milestones.* Chemistry, 2020. **26**(67): p. 15438-15445.
118. Sudarsan, N., J.E. Barrick, and R.R. Breaker, *Metabolite-binding RNA domains are present in the genes of eukaryotes.* RNA, 2003. **9**(6): p. 644-7.
119. Winkler, W., A. Nahvi, and R.R. Breaker, *Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression.* Nature, 2002. **419**(6910): p. 952-6.
120. Polaski, J.T., et al., *Mechanistic Insights into Cofactor-Dependent Coupling of RNA Folding and mRNA Transcription/Translation by a Cobalamin Riboswitch.* Cell Rep, 2016. **15**(5): p. 1100-1110.
121. Lussier, A., et al., *A kissing loop is important for btuB riboswitch ligand sensing and regulatory control.* J Biol Chem, 2015. **290**(44): p. 26739-51.
122. Johnson, J.E., Jr., et al., *B12 cofactors directly stabilize an mRNA regulatory switch.* Nature, 2012. **492**(7427): p. 133-7.
123. Bryant, D.A., C.N. Hunter, and M.J. Warren, *Biosynthesis of the modified tetrapyrroles-the pigments of life.* J Biol Chem, 2020. **295**(20): p. 6888-6925.
124. Santos, F., et al., *The complete coenzyme B12 biosynthesis gene cluster of Lactobacillus reuteri CRL1098.* Microbiology, 2008. **154**(Pt 1): p. 81-93.
125. Moore, S.J., et al., *Elucidation of the anaerobic pathway for the corrin component of cobalamin (vitamin B12).* Proc Natl Acad Sci U S A, 2013. **110**(37): p. 14906-11.
126. Kennedy, K.J. and M.E. Taga, *Cobamides.* Curr Biol, 2020. **30**(2): p. R55-r56.
127. Sokolovskaya, O.M., A.N. Shelton, and M.E. Taga, *Sharing vitamins: Cobamides unveil microbial interactions.* Science, 2020. **369**(6499).
128. Guimarães, D.H., et al., *Guanylcobamide and hypoxanthylcobamide - Corrinoids formed by Desulfovibrio vulgaris.* Archives of Microbiology, 1994. **162**(4): p. 272-276.

129. Anderson, P.J., et al., *One pathway can incorporate either adenine or dimethylbenzimidazole as an alpha-axial ligand of B12 cofactors in Salmonella enterica.* J Bacteriol, 2008. **190**(4): p. 1160-71.
130. Fieber, W., et al., *Pseudocoenzyme B12 and Adenosyl-Factor A: Electrochemical Synthesis and Spectroscopic Analysis of Two Natural B12 Coenzymes with Predominantly 'Base-off' Constitution.* Helvetica Chimica Acta, 2002. **85**(3): p. 927-944.
131. Hoffmann, B., et al., *Native corrinoids from Clostridium cochlearium are adeninylcobamides: spectroscopic analysis and identification of pseudovitamin B12 and factor A.* J Bacteriol, 2000. **182**(17): p. 4773-82.
132. Stupperich, E., H.J. Eisinger, and B. Kräutler, *Identification of phenolyl cobamide from the homoacetogenic bacterium Sporomusa ovata.* Eur J Biochem, 1989. **186**(3): p. 657-61.
133. Kräutler, B., H.P. Kohler, and E. Stupperich, *5'-Methylbenzimidazolyl-cobamides are the corrinoids from some sulfate-reducing and sulfur-metabolizing bacteria.* Eur J Biochem, 1988. **176**(2): p. 461-9.
134. Stupperich, E., H.J. Eisinger, and B. Kräutler, *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, 1988. **172**(2): p. 459-64.
135. Kräutler, B., J. Moll, and R.K. Thauer, *The corrinoid from Methanobacterium thermoautotrophicum (Marburg strain). Spectroscopic structure analysis and identification as Co beta-cyano-5'-hydroxybenzimidazolyl-cobamide (factor III).* Eur J Biochem, 1987. **162**(2): p. 275-8.
136. Yan, J., et al., *Purinyl-cobamide is a native prosthetic group of reductive dehalogenases.* Nat Chem Biol, 2018. **14**(1): p. 8-14.
137. Ladd, J.N., H.P. Hogenkamp, and H.A. Barker, *Structure of cobamide coenzymes: influence of pH on absorption spectra and ionophoretic mobilities.* J Biol Chem, 1961. **236**: p. 2114-8.
138. Barker, H.A., et al., *Isolation and properties of crystalline cobamide coenzymes containing benzimidazole or 5,6-dimethylbenzimidazole.* J Biol Chem, 1960. **235**: p. 480-8.
139. Shelton, A.N., X. Lyu, and M.E. Taga, *Flexible Cobamide Metabolism in Clostridioides (Clostridium) difficile 630 Δerm.* J Bacteriol, 2020. **202**(2).
140. Sokolovskaya, O.M., et al., *Naturally occurring cobalamin (B12) analogs can function as cofactors for human methylmalonyl-CoA mutase.* Biochimie, 2020.
141. Hazra, A.B., et al., *Anaerobic biosynthesis of the lower ligand of vitamin B12.* Proc Natl Acad Sci U S A, 2015. **112**(34): p. 10792-7.
142. Men, Y., et al., *Sustainable growth of Dehalococcoides mccartyi 195 by corrinoid salvaging and remodeling in defined lactate-fermenting consortia.* Appl Environ Microbiol, 2014. **80**(7): p. 2133-41.
143. Mok, K.C. and M.E. Taga, *Growth Inhibition of Sporomusa ovata by Incorporation of Benzimidazole Bases into Cobamides.* J Bacteriol, 2013. **195**(9): p. 1902-11.
144. Crofts, T.S., et al., *Cobamide structure depends on both lower ligand availability and CobT substrate specificity.* Chem Biol, 2013. **20**(10): p. 1265-74.
145. Hazra, A.B., et al., *Analysis of substrate specificity in CobT homologs reveals widespread preference for DMB, the lower axial ligand of vitamin B12.* Chem Biol, 2013. **20**(10): p. 1275-85.

146. Yi, S., et al., *Versatility in corrinoid salvaging and remodeling pathways supports corrinoid-dependent metabolism in Dehalococcoides mccartyi.* Appl Environ Microbiol, 2012. **78**(21): p. 7745-52.
147. Mok, K.C., et al., *Identification of a novel cobamide remodeling enzyme in the beneficial human gut bacterium Akkermansia muciniphila.* mBio, 2020. **11**(6): p. e02507-20.
148. Keller, S., et al., *Selective Utilization of Benzimidazolyl-Norcobamides as Cofactors by the Tetrachloroethene Reductive Dehalogenase of Sulfurospirillum multivorans.* J Bacteriol, 2018. **200**(8).
149. Helliwell, K.E., et al., *Cyanobacteria and Eukaryotic Algae Use Different Chemical Variants of Vitamin B12.* Curr Biol, 2016. **26**(8): p. 999-1008.
150. Choudhary, P.K., et al., *Diversity of cobalamin riboswitches in the corrinoid-producing organohalide respirer Desulfitobacterium hafniense.* J Bacteriol, 2013. **195**(22): p. 5186-95.
151. Keller, S., et al., *Exogenous 5,6-dimethylbenzimidazole caused production of a non-functional tetrachloroethene reductive dehalogenase in Sulfurospirillum multivorans.* Environ Microbiol, 2013.
152. Gray, M.J. and J.C. Escalante-Semerena, *In vivo analysis of cobinamide salvaging in Rhodobacter sphaeroides strain 2.4.1.* J Bacteriol, 2009. **191**(12): p. 3842-51.
153. Santos, F., et al., *Pseudovitamin B12 is the corrinoid produced by Lactobacillus reuteri CRL1098 under anaerobic conditions.* FEBS Lett, 2007. **581**(25): p. 4865-70.
154. Stupperich, E., H.J. Eisinger, and S. Schurr, *Corrinoids in anaerobic bacteria.* FEMS Microbiol Lett, 1990. **87**(3â€□4): p. 355-360.
155. Shelton, A.N., et al., *Uneven distribution of cobamide biosynthesis and dependence in bacteria predicted by comparative genomics.* ISME J, 2019. **13**(3): p. 789-804.
156. Rodionov, D.A., et al., *Comparative genomics of the vitamin B12 metabolism and regulation in prokaryotes.* J Biol Chem, 2003. **278**(42): p. 41148-59.
157. Gude, S. and M.E. Taga, *Multi-faceted approaches to discovering and predicting microbial nutritional interactions.* Curr Opin Biotechnol, 2019. **62**: p. 58-64.
158. Abreu, N.A. and M.E. Taga, *Decoding molecular interactions in microbial communities.* FEMS Microbiol Rev, 2016. **40**(5): p. 648-63.
159. Men, Y., et al., *Identification of specific corrinoids reveals corrinoid modification in dechlorinating microbial communities.* Environ Microbiol, 2015. **17**(12): p. 4873-84.
160. Degnan, P.H., et al., *Human gut microbes use multiple transporters to distinguish vitamin B12 analogs and compete in the gut.* Cell Host Microbe, 2014. **15**(1): p. 47-57.
161. Seth, E.C. and M.E. Taga, *Nutrient cross-feeding in the microbial world.* Front Microbiol, 2014. **5**: p. 350.
162. Degnan, P.H., M.E. Taga, and A.L. Goodman, *Vitamin B12 as a modulator of gut microbial ecology.* Cell Metab, 2014. **20**(5): p. 769-78.
163. Gallo, S., et al., *The corrin moiety of coenzyme B12 is the determinant for switching the btuB riboswitch of E. coli.* Chembiochem, 2008. **9**(9): p. 1408-14.
164. Drennan, C.L., R.G. Matthews, and M.L. Ludwig, *Cobalamin-dependent methionine synthase: the structure of a methylcobalamin-binding fragment and implications for other B12-dependent enzymes.* Curr Opin Struct Biol, 1994. **4**(6): p. 919-29.
165. Fontecave, M. and E. Mulliez, *Ribonucleotide reductases*, in *Chemistry and Biochemistry of B12*, R. Banerjee, Editor. 1999, John Wiley & Sons, Inc.: New York. p. 731-756.

166. Frey, B., et al., *New function of vitamin B12: cobamide-dependent reduction of epoxyqueuosine to queuosine in tRNAs of Escherichia coli and Salmonella typhimurium.* J Bacteriol, 1988. **170**(5): p. 2078-82.
167. Banerjee, R. and S. Chowdhury, *Methylmalonyl-CoA mutase*, in *Chemistry and Biochemistry of B12*, R. Banerjee, Editor. 1999, John Wiley & Sons, Inc.: New York. p. 707-729.
168. Kim, H.J., et al., *GenK-catalyzed C-6' methylation in the biosynthesis of gentamicin: isolation and characterization of a cobalamin-dependent radical SAM enzyme.* J Am Chem Soc, 2013. **135**(22): p. 8093-6.
169. Gruber, K., R. Reitzer, and C. Kratky, *Radical Shuttling in a Protein: Ribose Pseudorotation Controls Alkyl-Radical Transfer in the Coenzyme B12 Dependent Enzyme Glutamate Mutase.* Angewandte Chemie International Edition, 2001. **40**(18): p. 3377-3380.
170. Chen, H.P. and E.N. Marsh, *How enzymes control the reactivity of adenosylcobalamin: effect on coenzyme binding and catalysis of mutations in the conserved histidine-aspartate pair of glutamate mutase.* Biochemistry, 1997. **36**(25): p. 7884-9.
171. Barker, H.A., *beta-Methylaspartate-glutamate mutase from Clostridium tetanomorphum.* Methods Enzymol, 1985. **113**: p. 121-33.
172. Berkovitch, F., et al., *A locking mechanism preventing radical damage in the absence of substrate, as revealed by the x-ray structure of lysine 5,6-aminomutase.* Proc Natl Acad Sci U S A, 2004. **101**(45): p. 15870-5.
173. Chen, H.P., et al., *Cloning, sequencing, heterologous expression, purification, and characterization of adenosylcobalamin-dependent D-ornithine aminomutase from Clostridium sticklandii.* J Biol Chem, 2001. **276**(48): p. 44744-50.
174. Chang, C.H. and P.A. Frey, *Cloning, sequencing, heterologous expression, purification, and characterization of adenosylcobalamin-dependent D-lysine 5, 6-aminomutase from Clostridium sticklandii.* J Biol Chem, 2000. **275**(1): p. 106-14.
175. Roof, D.M. and J.R. Roth, *Ethanolamine utilization in Salmonella typhimurium.* J Bacteriol, 1988. **170**(9): p. 3855-63.
176. Scarlett, F.A. and J.M. Turner, *Microbial metabolism of amino alcohols. Ethanolamine catabolism mediated by coenzyme B12-dependent ethanolamine ammonia-lyase in Escherichia coli and Klebsiella aerogenes.* J Gen Microbiol, 1976. **95**(1): p. 173-6.
177. Jeter, R.M., *Cobalamin-dependent 1,2-propanediol utilization by Salmonella typhimurium.* J Gen Microbiol, 1990. **136**(5): p. 887-96.
178. Schutz, H. and F. Radler, *Propanediol-1,2-dehydratase and metabolism of glycerol of Lactobacillus brevis.* Archives of Microbiology, 1984. **139**(4): p. 366-370.
179. Parks, J.M., et al., *The genetic basis for bacterial mercury methylation.* Science, 2013. **339**(6125): p. 1332-5.
180. Choi, S.C., T. Chase, Jr., and R. Bartha, *Enzymatic catalysis of mercury methylation by Desulfovibrio desulfuricans LS.* Appl Environ Microbiol, 1994. **60**(4): p. 1342-6.
181. Choi, S.C. and R. Bartha, *Cobalamin-mediated mercury methylation by Desulfovibrio desulfuricans LS.* Appl Environ Microbiol, 1993. **59**(1): p. 290-5.
182. Neumann, A., et al., *Tetrachloroethene reductive dehalogenase of Dehalospirillum multivorans: substrate specificity of the native enzyme and its corrinoid cofactor.* Arch Microbiol, 2002. **177**(5): p. 420-6.

183. Krasotkina, J., et al., *Characterization of the B12- and iron-sulfur-containing reductive dehalogenase from Desulfitobacterium chlororespirans.* J Biol Chem, 2001. **276**(44): p. 40991-7.
184. Banerjee, R. and S.W. Ragsdale, *The many faces of vitamin B12: catalysis by cobalamin-dependent enzymes.* Annu Rev Biochem, 2003. **72**: p. 209-47.
185. Moore, S.J., et al., *Towards a cell factory for vitamin B12 production in Bacillus megaterium: bypassing of the cobalamin riboswitch control elements.* N Biotechnol, 2014. **31**(6): p. 553-61.
186. Biedendieck, R., et al., *Metabolic engineering of cobalamin (vitamin B12) production in Bacillus megaterium.* Microb Biotechnol, 2010. **3**(1): p. 24-37.
187. Li, J., et al., *Regulating vitamin B12 biosynthesis via the cbiMCbl riboswitch in Propionibacterium strain UF1.* Proc Natl Acad Sci U S A, 2020. **117**(1): p. 602-609.
188. González, J.C., et al., *Cobalamin-independent methionine synthase from Escherichia coli: a zinc metalloenzyme.* Biochemistry, 1996. **35**(38): p. 12228-34.
189. González, J.C., et al., *Comparison of cobalamin-independent and cobalamin-dependent methionine synthases from Escherichia coli: two solutions to the same chemical problem.* Biochemistry, 1992. **31**(26): p. 6045-56.
190. Sherlock, M.E. and R.R. Breaker, *Former orphan riboswitches reveal unexplored areas of bacterial metabolism, signaling, and gene control processes.* RNA, 2020. **26**(6): p. 675-693.
191. Campbell, G.R.O., et al., *Sinorhizobium meliloti bluB is necessary for production of 5,6-dimethylbenzimidazole, the lower ligand of B12.* Proc Natl Acad Sci USA, 2006. **103**: p. 4634-4639.
192. Taga, M.E., et al., *BluB cannibalizes flavin to form the lower ligand of vitamin B12.* Nature, 2007. **446**(7134): p. 449-53.
193. Gray, M.J. and J.C. Escalante-Semerena, *Single-enzyme conversion of FMNH2 to 5,6-dimethylbenzimidazole, the lower ligand of B12.* Proc Natl Acad Sci U S A, 2007. **104**(8): p. 2921-6.
194. Wurm, R., R. Weyhenmeyer, and P. Renz, *On the biosynthesis of 5-methoxybenzimidazole. Precursor-function of 5-hydroxybenzimidazole, benzimidazole and riboflavin.* Eur J Biochem, 1975. **56**(2): p. 427-32.
195. Lamm, L., et al., *Biosynthesis of vitamin B12. Experiments with the anaerobe Eubacterium limosum and some labelled substrates.* Eur J Biochem., 1980. **109**(1): p. 115-8.
196. Hollriegl, V., et al., *Biosynthesis of vitamin B12. Different pathways in some aerobic and anaerobic microorganisms.* Arch Microbiol., 1982. **132**(2): p. 155-8.
197. Lamm, L., G. Heckmann, and P. Renz, *Biosynthesis of vitamin B12 in anaerobic bacteria. Mode of incorporation of glycine into the 5,6-dimethylbenzimidazole moiety in Eubacterium limosum.* Eur J Biochem., 1982. **122**(3): p. 569-71.
198. Vogt, J.R., L. Lamm-Kolonko, and P. Renz, *Biosynthesis of vitamin B-12 in anaerobic bacteria. Experiments with Eubacterium limosum and D-erythrose 14C-labeled in different positions.* Eur J Biochem., 1988. **174**(4): p. 637-40.
199. Vogt, J.R. and P. Renz, *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., 1988. **171**(3): p. 655-9.

200. Munder, M., et al., *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., 1992. **204**(2): p. 679-83.
201. Renz, P., et al., *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, 1993. **217**(3): p. 1117-21.
202. Cech, Thomas R. and Joan A. Steitz, *The Noncoding RNA Revolution - Trashing Old Rules to Forge New Ones.* Cell, 2014. **157**(1): p. 77-94.
203. Roth, A. and R.R. Breaker, *The structural and functional diversity of metabolite-binding riboswitches.* Annu Rev Biochem, 2009. **78**: p. 305-34.
204. Winkler, W.C. and R.R. Breaker, *Regulation of bacterial gene expression by riboswitches.* Annu Rev Microbiol, 2005. **59**: p. 487-517.
205. Quereda, J.J. and P. Cossart, *Regulating Bacterial Virulence with RNA.* Annu Rev Microbiol, 2017. **71**: p. 263-280.
206. Wrist, A., W. Sun, and R.M. Summers, *The Theophylline Aptamer: 25 Years as an Important Tool in Cellular Engineering Research.* ACS Synth Biol, 2020. **9**(4): p. 682-697.
207. Hallberg, Z.F., et al., *Engineering and In Vivo Applications of Riboswitches.* Annu Rev Biochem, 2017. **86**: p. 515-539.
208. Bridwell-Rabb, J. and C.L. Drennan, *Vitamin B12 in the spotlight again.* Current Opinion in Chemical Biology, 2017. **37**: p. 63-70.
209. Licht, S., G.J. Gerfen, and J. Stubbe, *Thiyl radicals in ribonucleotide reductases.* Science, 1996. **271**(5248): p. 477-81.
210. Berg, I.A., et al., *A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea.* Science, 2007. **318**(5857): p. 1782-6.
211. Ljungdahl, L.G., *The autotrophic pathway of acetate synthesis in acetogenic bacteria.* Annu Rev Microbiol, 1986. **40**: p. 415-50.
212. Stupperich, E. and R. Konle, *Corrinoid-Dependent Methyl Transfer Reactions Are Involved in Methanol and 3,4-Dimethoxybenzoate Metabolism by Sporomusa ovata.* Appl Environ Microbiol, 1993. **59**(9): p. 3110-6.
213. Forage, R.G. and M.A. Foster, *Glycerol fermentation in Klebsiella pneumoniae: functions of the coenzyme B12-dependent glycerol and diol dehydratases.* J Bacteriol, 1982. **149**(2): p. 413-9.
214. Erb, T.J., et al., *Ethylmalonyl-CoA mutase from Rhodobacter sphaeroides defines a new subclade of coenzyme B12-dependent acyl-CoA mutases.* J Biol Chem, 2008. **283**(47): p. 32283-93.
215. Korotkova, N., et al., *Glyoxylate regeneration pathway in the methylotroph Methylobacterium extorquens AM1.* J Bacteriol, 2002. **184**(6): p. 1750-8.
216. Ferguson, D.J., Jr. and J.A. Krzycki, *Reconstitution of trimethylamine-dependent coenzyme M methylation with the trimethylamine corrinoid protein and the isozymes of methyltransferase II from Methanosarcina barkeri.* J Bacteriol, 1997. **179**(3): p. 846-52.
217. Knox, H.L., et al., *Structural basis for non-radical catalysis by TsrM, a radical SAM methylase.* Nat Chem Biol, 2021. **17**(4): p. 485-491.
218. Wang, P.H., et al., *Retroconversion of estrogens into androgens by bacteria via a cobalamin-mediated methylation.* Proc Natl Acad Sci U S A, 2020. **117**(3): p. 1395-1403.

219. Romine, M.F., et al., *Elucidation of roles for vitamin B12 in regulation of folate, ubiquinone, and methionine metabolism.* Proc Natl Acad Sci U S A, 2017. **114**(7): p. E1205-e1214.
220. Bridwell-Rabb, J., et al., *A B12-dependent radical SAM enzyme involved in oxetanocin A biosynthesis.* Nature, 2017. **544**(7650): p. 322-326.
221. Allen, K.D. and S.C. Wang, *Initial characterization of Fom3 from Streptomyces wedmorensis: The methyltransferase in fosfomycin biosynthesis.* Arch Biochem Biophys, 2014. **543**: p. 67-73.
222. Pierre, S., et al., *Thiostrepton tryptophan methyltransferase expands the chemistry of radical SAM enzymes.* Nat Chem Biol, 2012. **8**(12): p. 957-9.
223. Marous, D.R., et al., *Consecutive radical S-adenosylmethionine methylations form the ethyl side chain in thienamycin biosynthesis.* Proc Natl Acad Sci U S A, 2015. **112**(33): p. 10354-8.
224. Takano, H., et al., *Role and Function of LitR, an Adenosyl B12-Bound Light-Sensitive Regulator of Bacillus megaterium QM B1551, in Regulation of Carotenoid Production.* J Bacteriol, 2015. **197**(14): p. 2301-15.
225. Takano, H., et al., *Involvement of CarA/LitR and CRP/FNR family transcriptional regulators in light-induced carotenoid production in Thermus thermophilus.* J Bacteriol, 2011. **193**(10): p. 2451-9.
226. Jost, M., et al., *Structural basis for gene regulation by a B12-dependent photoreceptor.* Nature, 2015. **526**(7574): p. 536-41.
227. Pavlova, N., D. Kaloudas, and R. Penchovsky, *Riboswitch distribution, structure, and function in bacteria.* Gene, 2019. **708**: p. 38-48.
228. Battersby, A.R., *Tetrapyrroles: the pigments of life.* Nat Prod Rep, 2000. **17**(6): p. 507-26.
229. Chan, C.H. and J.C. Escalante-Semerena, *ArsAB, a novel enzyme from Sporomusa ovata activates phenolic bases for adenosylcobamide biosynthesis.* Mol Microbiol, 2012. **81**(4): p. 952-67.
230. Newmister, S.A., et al., *Structural insights into the function of the nicotinate mononucleotide:phenol/p-cresol phosphoribosyltransferase (ArsAB) enzyme from Sporomusa ovata.* Biochemistry, 2012. **51**(43): p. 8571-82.
231. Sokolovskaya, O.M., et al., *Cofactor Selectivity in Methylmalonyl Coenzyme A Mutase, a Model Cobamide-Dependent Enzyme.* mBio, 2019. **10**(5): p. e01303-19.
232. Cooper, M.B., et al., *Cross-exchange of B-vitamins underpins a mutualistic interaction between Ostreococcus tauri and Dinoroseobacter shibae.* ISME J, 2019. **13**(2): p. 334-345.
233. Croft, M.T., et al., *Algae acquire vitamin B12 through a symbiotic relationship with bacteria.* Nature, 2005. **438**(7064): p. 90-3.
234. Ma, A.T., J. Beld, and B. Brahamsha, *An amoebal grazer of cyanobacteria requires cobalamin produced by heterotrophic bacteria.* Appl. Environ. Microbiol., 2017. **83**(10): p. e00035-17.
235. Ma, A.T., B. Tyrell, and J. Beld, *Specificity of cobamide remodeling, uptake and utilization in Vibrio cholerae.* Molecular Microbiology, 2020. **113**(1): p. 89-102.
236. Polaski, J.T., et al., *Cobalamin riboswitches exhibit a broad range of ability to discriminate between methylcobalamin and adenosylcobalamin.* J Biol Chem, 2017. **292**(28): p. 11650-11658.

237. Polaski, J.T., et al., *A functional genetic screen reveals sequence preferences within a key tertiary interaction in cobalamin riboswitches required for ligand selectivity.* Nucleic Acids Res, 2018. **46**(17): p. 9094-9105.
238. Cadieux, N., et al., *Identification of the periplasmic cobalamin-binding protein BtuF of Escherichia coli.* J Bacteriol, 2002. **184**(3): p. 706-17.
239. Van Bibber, M., et al., *A new class of cobalamin transport mutants (btuF) provides genetic evidence for a periplasmic binding protein in Salmonella typhimurium.* J Bacteriol, 1999. **181**(17): p. 5539-41.
240. Forouhar, F., et al., *Functional insights from structural genomics.* J Struct Funct Genomics, 2007. **8**(2-3): p. 37-44.
241. Lundrigan, M.D. and R.J. Kadner, *Altered cobalamin metabolism in Escherichia coli btuR mutants affects btuB gene regulation.* J Bacteriol, 1989. **171**(1): p. 154-61.
242. Sudarsan, N., et al., *Tandem riboswitch architectures exhibit complex gene control functions.* Science, 2006. **314**(5797): p. 300-4.
243. Warner, D.F., et al., *A riboswitch regulates expression of the coenzyme B12-independent methionine synthase in Mycobacterium tuberculosis: implications for differential methionine synthase function in strains H37Rv and CDC1551.* J Bacteriol, 2007. **189**(9): p. 3655-9.
244. Chan, C.W. and A. Mondragón, *Crystal structure of an atypical cobalamin riboswitch reveals RNA structural adaptability as basis for promiscuous ligand binding.* Nucleic Acids Res, 2020. **48**(13): p. 7569-7583.
245. Peselis, A. and A. Serganov, *Structural insights into ligand binding and gene expression control by an adenosylcobalamin riboswitch.* Nat Struct Mol Biol, 2013. **19**(11): p. 1182-4.
246. Marques, H.M., X. Zou, and K.L. Brown, *The solution structure of adenosylcobalamin and adenosylcobinamide determined by nOe-restrained molecular dynamics simulations.* Journal of Molecular Structure, 2000. **520**(1): p. 75-95.
247. Sussman, D. and C. Wilson, *A water channel in the core of the vitamin B12 RNA aptamer.* Structure, 2000. **8**(7): p. 719-27.
248. Mancia, F., et al., *How coenzyme B12 radicals are generated: the crystal structure of methylmalonyl-coenzyme A mutase at 2 A resolution.* Structure, 1996. **4**(3): p. 339-50.
249. Froese, D.S., et al., *Structures of the human GTPase MMAA and vitamin B12-dependent methylmalonyl-CoA mutase and insight into their complex formation.* J Biol Chem, 2010. **285**(49): p. 38204-13.
250. Drennan, C.L., et al., *How a protein binds B12: A 3.0 A X-ray structure of B12-binding domains of methionine synthase.* Science, 1994. **266**(5191): p. 1669-74.
251. Payne, K.A., et al., *Reductive dehalogenase structure suggests a mechanism for B12-dependent dehalogenation.* Nature, 2015. **517**(7535): p. 513-516.
252. Jost, M., et al., *Structural Basis for Substrate Specificity in Adenosylcobalamin-dependent Isobutyryl-CoA Mutase and Related Acyl-CoA Mutases.* J Biol Chem, 2015. **290**(45): p. 26882-26898.
253. Tanioka, Y., et al., *Methyladeninylcobamide functions as the cofactor of methionine synthase in a Cyanobacterium, Spirulina platensis NIES-39.* FEBS Lett, 2010. **584**(14): p. 3223-6.
254. Gallo, S., et al., *The change of corrin-amides to carboxylates leads to altered structures of the B12-responding btuB riboswitch.* Chem Commun (Camb), 2011. **47**(1): p. 403-5.

255. Ma, B., et al., *Conformational Ensemble of TteAdoCbl Riboswitch Provides Stable Structural Elements for Conformation Selection and Population Shift in Cobalamin Recognition.* J Phys Chem B, 2021. **125**(10): p. 2589-2596.
256. Trausch, J.J. and R.T. Batey, *A disconnect between high-affinity binding and efficient regulation by antifolates and purines in the tetrahydrofolate riboswitch.* Chem Biol, 2014. **21**(2): p. 205-16.
257. Fedosov, S.N., et al., *Mechanisms of discrimination between cobalamins and their natural analogues during their binding to the specific B12-transporting proteins.* Biochemistry, 2007. **46**(21): p. 6446-58.
258. Chan, C.H., et al., *Dissecting cobamide diversity through structural and functional analyses of the base-activating CobT enzyme of Salmonella enterica.* Biochim Biophys Acta, 2014. **1840**(1): p. 464-75.
259. Cheong, C.G., J.C. Escalante-Semerena, and I. Rayment, *Structural investigation of the biosynthesis of alternative lower ligands for cobamides by nicotinate mononucleotide: 5,6-dimethylbenzimidazole phosphoribosyltransferase from Salmonella enterica.* J Biol Chem, 2001. **276**(40): p. 37612-20.
260. Cheong, C.G., J.C. Escalante-Semerena, and I. Rayment, *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 A resolution.* Biochemistry, 1999. **38**(49): p. 16125-35.
261. Trzebiatowski, J.R. and J.C. Escalante-Semerena, *Purification and characterization of CobT, the nicotinate-mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase enzyme from Salmonella typhimurium LT2.* J Biol Chem, 1997. **272**(28): p. 17662-7.
262. Trzebiatowski, J.R., G.A. O'Toole, and J.C. Escalante-Semerena, *The cobT gene of Salmonella typhimurium encodes the NaMN: 5,6-dimethylbenzimidazole phosphoribosyltransferase responsible for the synthesis of N1-(5-phospho-alpha-D-ribosyl)-5,6-dimethylbenzimidazole, an intermediate in the synthesis of the nucleotide loop of cobalamin.* J Bacteriol, 1994. **176**(12): p. 3568-75.
263. Villa, J.K., et al., *Synthetic Biology of Small RNAs and Riboswitches.* Microbiol Spectr, 2018. **6**(3).
264. Galizi, R. and A. Jaramillo, *Engineering CRISPR guide RNA riboswitches for in vivo applications.* Curr Opin Biotechnol, 2019. **55**: p. 103-113.
265. Zelder, F., M. Sonnay, and L. Prieto, *Antivitamins for Medicinal Applications.* Chembiochem, 2015. **16**(9): p. 1264-78.
266. Mukherjee, S. and S. Sengupta, *Riboswitch Scanner: an efficient pHMM-based web-server to detect riboswitches in genomic sequences.* Bioinformatics, 2016. **32**(5): p. 776-8.
267. Singh, P., et al., *Riboswitch detection using profile hidden Markov models.* BMC Bioinformatics, 2009. **10**: p. 325.
268. Mathews, D.H., et al., *Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure.* Proc Natl Acad Sci U S A, 2004. **101**(19): p. 7287-92.
269. Reuter, J.S. and D.H. Mathews, *RNAstructure: software for RNA secondary structure prediction and analysis.* BMC Bioinformatics, 2010. **11**: p. 129.

270. Waterhouse, A.M., et al., *Jalview Version 2--a multiple sequence alignment editor and analysis workbench.* Bioinformatics, 2009. **25**(9): p. 1189-91.
271. Gibson, D.G., et al., *Enzymatic assembly of DNA molecules up to several hundred kilobases.* Nature methods, 2009. **6**(5): p. 343-345.
272. Guiziou, S., et al., *A part toolbox to tune genetic expression in Bacillus subtilis.* Nucleic Acids Res, 2016. **44**(15): p. 7495-508.
273. Zhang, X.Z. and Y. Zhang, *Simple, fast and high-efficiency transformation system for directed evolution of cellulase in Bacillus subtilis.* Microb Biotechnol, 2011. **4**(1): p. 98-105.
274. Koo, B.M., et al., *Construction and Analysis of Two Genome-Scale Deletion Libraries for Bacillus subtilis.* Cell Syst, 2017. **4**(3): p. 291-305.e7.
275. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Research, 2000. **28**(1): p. 235-242.
276. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis.* J Comput Chem, 2004. **25**(13): p. 1605-12.
277. Anagnostopoulos, C. and J. Spizizen, *Requirements for transformation in Bacillus subtilis.* J Bacteriol, 1961. **81**(5): p. 741-6.
278. Sherlock, M.E., et al., *Tandem riboswitches form a natural Boolean logic gate to control purine metabolism in bacteria.* Elife, 2018. **7**.
279. Peselis, A. and A. Serganov, *ykkC riboswitches employ an add-on helix to adjust specificity for polyanionic ligands.* Nat Chem Biol, 2018. **14**(9): p. 887-894.
280. Roßmanith, J. and F. Narberhaus, *Modular arrangement of regulatory RNA elements.* RNA Biol, 2017. **14**(3): p. 287-292.
281. Wachsmuth, M., et al., *Design criteria for synthetic riboswitches acting on transcription.* RNA Biol, 2015. **12**(2): p. 221-31.
282. Yaneva, N., et al., *Bacterial acyl-CoA mutase specifically catalyzes coenzyme B12-dependent isomerization of 2-hydroxyisobutyryl-CoA and (S)-3-hydroxybutyryl-CoA.* J Biol Chem, 2012. **287**(19): p. 15502-11.
283. Takami, H., et al., *Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis.* Nucleic Acids Res, 2000. **28**(21): p. 4317-31.
284. Gray, M.J., N.K. Tavares, and J.C. Escalante-Semerena, *The genome of Rhodobacter sphaeroides strain 2.4.1 encodes functional cobinamide salvaging systems of archaeal and bacterial origins.* Mol Microbiol, 2008. **70**(4): p. 824-36.
285. Taga, M.E. and G.C. Walker, *Sinorhizobium meliloti requires a cobalamin-dependent ribonucleotide reductase for symbiosis with its plant host.* Mol Plant Microbe Interact, 2010. **23**(12): p. 1643-54.
286. O'Toole, G.A., M.R. Rondon, and J.C. Escalante-Semerena, *Analysis of mutants of Salmonella typhimurium defective in the synthesis of the nucleotide loop of cobalamin.* J Bacteriol, 1993. **175**(11): p. 3317-26.
287. Coppins, R.L., K.B. Hall, and E.A. Groisman, *The intricate world of riboswitches.* Curr Opin Microbiol, 2007. **10**(2): p. 176-81.
288. Andreasson, J.O.L., et al., *Comprehensive sequence-to-function mapping of cofactor-dependent RNA catalysis in the glmS ribozyme.* Nat Commun, 2020. **11**(1): p. 1663.