

**RIBOSWITCH-BASED SENSORS OF SMALL
MOLECULES IN BACTERIA**

**THE ENGINEERING OF RIBOSWITCH-BASED SENSORS OF
SMALL MOLECULES IN BACTERIA AND THEIR APPLICATION
IN THE STUDY OF VITAMIN B12 BIOLOGY**

By

CASEY C. FOWLER

B.Sc., Queen's University, Kingston Ontario, 2004

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

© by Casey C. Fowler, August 2011

DOCTOR OF PHILOSOPHY (2011)
(Biochemistry & Biomedical Sciences)

McMaster University
Hamilton, Ontario

TITLE: The Engineering of Riboswitch-Based Sensors of Small Molecules in Bacteria
and Their Application in the Study of Vitamin B12 Biology

AUTHOR: Casey C. Fowler, B.Sc. (Queen's University)

SUPERVISOR: Dr. Yingfu Li

NUMBER OF PAGES: xiii, 191

Abstract

Small molecule metabolites have important and diverse roles in every major cellular function. To study the activities of metabolites and the biological processes in which they are involved, it is important to be able to detect their levels within cells. Technologies that measure the concentrations of small molecules within the context of living, growing cells are highly advantageous but are challenging to produce. In this thesis, a novel class of intracellular small molecule sensors is produced, characterized and applied to address novel and relevant research questions. These sensors detect a specific target molecule within bacterial cells using RNA regulatory elements known as riboswitches and one of many possible reporter proteins. In addition to a project that yielded new methodology to create custom riboswitches, two projects that assess the capabilities of sensors that detect an active form of vitamin B12 are described. These projects present an abundance of data that provide novel insights into the transport and metabolism of vitamin B12 in *E. coli* cells. Overall, the results presented indicate that riboswitch-based sensors represent valuable and unique tools for the study of microbial biology. The thesis is concluded with a discussion that describes design strategies and several exciting potential applications for future riboswitch sensors.

Acknowledgements

First and foremost I would like to thank my supervisors Dr. Yingfu Li and Dr. Eric Brown. Their contributions to the research presented in this thesis and to my development as a scientist have been monumental. My interactions with both of these fantastic mentors will continue to benefit me for the remainder of my career.

I would also like to recognize the contributions of my other supervisory committee member, Dr. David Andrews. His vigorous and carefully considered questions and suggestions kept me on my toes during committee meetings and were very helpful in addressing problems that arose over the course of my research. I would also like to acknowledge Dr. Murray Junop for the many long discussions we've had regarding the work described in Chapter 5 of this thesis. With his guidance I have learned a great deal about structural biology.

The Li lab has provided a pleasant and productive research environment throughout my time at McMaster. I would like to recognize Simon Mcmanus, Naveen Kumar, Wendy Mok and Kacper Kuryllo for the useful discussions, advice and assistance they have provided in the lab.

On a more personal level, I would like to thank Naomi as well as my family and friends for making my life outside of the lab so enjoyable throughout my time in graduate school.

To my parents, who have set and kept me on a happy path in life.

Table of Contents

Abstract	iii
Acknowledgements	iv
List of Tables	ix
List of Figures	x
List of Abbreviations	xii
Chapter 1. A General Introduction	1
1.1 Overview of the problem to be addressed.....	1
1.2 The isolation, properties and applications of aptamers	9
1.3 Riboswitches: metabolite-sensing RNA regulatory elements.....	18
1.4 The engineering and applications of bacterial riboswitches.....	35
1.5 Vitamin B12 and its synthesis, uptake and utilization in <i>E. coli</i> cells.....	44
1.6 Specific research objectives and thesis organization.....	52
Chapter 2. A FACS-Based Approach to Engineering Artificial Riboswitches	57
2.1 Author's preface.....	57
2.2 Abstract.....	59
2.3 Introduction.....	59
2.4 Materials and Methods.....	63
2.4.1 Growth media, strains, plasmids and molecular cloning.....	63
2.4.2 Fluorescence activated cell sorting and colony isolation.....	64
2.4.3 Fluorescence assays.....	65

2.5 Results	66
2.5.1 Selection scheme and library design.....	66
2.5.2 FACS to isolate candidate riboswitch sequences.....	67
2.5.3 Assaying candidate riboswitches for theophylline response.....	70
2.5.4 Mechanistic characterization of RS11A.....	72
2.6 Discussion	75
2.7 References	78

Chapter 3. Using a Riboswitch Sensor to Examine Coenzyme B12 Metabolism and Transport in *E. coli*80

3.1 Author's preface	80
3.2 Abstract	81
3.3 Introduction	82
3.4 Materials and Methods	85
3.4.1 Plasmids and cell strains.....	85
3.4.2 Molecular cloning.....	86
3.4.3 Growth conditions and reporter assays.....	86
3.4.4 Ethanolamine utilization assays.....	88
3.5 Results	89
3.5.1 Construction of riboswitch-based AdoCbl sensors.....	89
3.5.2 Riboswitch-based sensors can be used to monitor AdoCbl metabolism.....	93
3.5.3 Riboswitch-based sensors can detect the availability of B12 lower axial ligands.....	98
3.5.4 Riboswitch-based sensors can be used to monitor AdoCbl transport.....	100
3.6 Discussion	103

3.7	References.....	107
Chapter 4.	Characterization of Intermolecular Interactions Involving the Periplasmic Binding Protein BtuF Using Riboswitch-Based Sensors.....	113
4.1	Author’s preface.....	113
4.2	Abstract.....	114
4.3	Introduction.....	115
4.4	Materials and Methods.....	119
4.4.1	Strains, plasmids and molecular cloning.....	119
4.4.2	Low-throughput reporter assays.....	120
4.4.3	Reporter assays for saturation mutagenesis libraries.....	122
4.4.4	Ethanolamine utilization assays.....	123
4.5	Results.....	123
4.5.1	An assay system to detect the activity of BtuF mutants.....	123
4.5.2	Assessing the activity of BtuC interaction mutants.....	127
4.5.3	Assessing the activity of mutations to the BtuF ligand binding pocket.....	130
4.5.4	Dose response of BtuF mutants.....	135
4.5.5	Substrate-specific transport defects of BtuF mutants.....	138
4.5.6	The effects of BtuF mutations on coenzyme B12-dependent growth.....	140
4.6	Discussion.....	144
4.7	References.....	149
Chapter 5.	Discussion and future directions.....	152

Appendix	166
Chapter 2. Supplemental materials.....	166
Chapter 3. Supplemental materials.....	169
Chapter 4. Supplemental materials.....	181
References	182

List of Tables

Chapter 1

Table 1-1 A sampling of aptamers of special interest.....	14
Table 1-2 Riboswitch classes discovered to date.....	29

Supplemental

Table S3-1 Description and sequence of oligonucleotide primers.....	171
--	-----

List of Figures

Chapter 1

- Figure 1-1** Schematic of general designs for protein-based genetically encoded metabolite sensors.....6
- Figure 1-2** A typical SELEX selection scheme for the isolation of RNA aptamers.....11
- Figure 1-3** Schematic representations of the three general regulatory mechanisms employed by bacterial riboswitches.....21
- Figure 1-4** Secondary structures of predicted bound and unbound conformations of a lysine riboswitch.....26
- Figure 1-5** Schematic representations of methods used to isolate synthetic riboswitches.....38
- Figure 1-6** Chemical structure of relevant vitamin B12 molecules.....46

Chapter 2

- Figure 2-1** Selection scheme and library design.....62
- Figure 2-2** FACS results.....68
- Figure 2-3** Performance of some selected riboswitches.....71
- Figure 2-4** Probing the regulation mechanism of the RS11A riboswitch.....73

Chapter 3

- Figure 3-1** Depiction of riboswitch-based intracellular sensors.....90
- Figure 3-2** Intracellular detection of AdoCbl by riboswitch-based sensors.....92
- Figure 3-3** Using riboswitch-based sensors to probe AdoCbl metabolism.....94
- Figure 3-4** Growth of metabolic deletion strains using ethanolamine as the sole nitrogen source.....97

Figure 3-5 Monitoring the availability of lower axial ligands using riboswitch-based sensors.....99

Figure 3-6 Using Riboswitch-Based Sensors to Probe AdoCbl Transport.....102

Chapter 4

Figure 4-1 The transport activity of chromosomal and plasmid-encoded BtuF.....126

Figure 4-2 The transport activity of E72 and E202 mutants.....129

Figure 4-3 The transport activity of BtuF ligand binding mutants.....132

Figure 4-4 The transport activity of BtuF variants as a function of CNCbl concentration.....137

Figure 4-5 The transport activity of BtuF variants using different substrates.....139

Figure 4-6 AdoCbl-dependent growth of BtuF mutants.....142

Supplemental

Figure S2-1 Cloning schematic.....166

Figure S2-2 In Line Probing.....166

Figure S3-1 Cloning scheme for riboswitch sensor constructs showing the relevant plasmid features.....169

Figure S3-2 Raw reporter activity levels in each of the tested strains under unrepressed conditions.....170

Figure S3-3 Activity of the mutant riboswitch construct in each of the strains tested.....170

Figure S4-1 Sample plate layout for the saturation mutagenesis library.....181

List of Abbreviations

ABC	ATP-binding cassette
AdoCbl	Coenzyme B12 (5' deoxyadenosylcobalamin)
ATP	adenosine 5'-triphosphate
bp	base-pair
Cbi	cobinamide
CNCbl	vitamin B12 (cyanocobalamin)
DMB	5,6-dimethylbenzimidazole
DNA	2'-deoxyribonucleic acid
ELISA	enzyme-linked immunosorbent assay
EV	empty vector
FACS	fluorescence activated cell sorting
FRET	fluorescence resonance energy transfer
GFP	green fluorescent protein
GlcN6P	glucosamine-6-phosphate
mRNA	messenger RNA
ONPG	2-Nitrophenylgalactoside
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBP	periplasmic binding protein
RFP	red fluorescent protein
RNA	ribonucleic acid
RNAP	RNA polymerase

rRNA	ribosomal RNA
SAM	S-adenosylmethionine
SELEX	systematic evolution of ligands by exponential enrichment
tRNA	transfer RNA
TPP	thiamine pyrophosphate
UTR	untranslated region
WT	wild-type
X-gal	5-bromo-4-chloro-3-indolyl- β -d-galactopyranoside

Chapter 1

A General Introduction

1.1 Overview of the problem to be addressed

The central dogma of biology speaks of a cell's DNA, RNA and protein aspects of its biology. To study these molecules, it has been necessary to generate technologies to identify, detect and quantify them. DNA sequencing technologies, which continue to improve at an astounding rate, have placed the complete genetic content of organisms at arm's reach (Shendure and Ji, 2008). With this information also comes the identity of the RNA and protein molecules that can potentially be produced. An assortment of technologies have been developed to detect and quantify the expression of individual RNAs and proteins that have been essential to studying their cellular functions. Methods such as northern blots, PCR-based methods, microarrays and, most recently, high-throughput RNA sequencing technologies (Ozsolak and Milos, 2011) have provided reliable and standardized means of detecting the expression and relative levels of individual cellular RNAs. For proteins, western blots, two dimensional polyacrylamide gel electrophoresis (PAGE), mass spectrometry and a wide range of antibody-based assays are amongst the common methodologies used to detect their levels in cells.

Not mentioned in the central dogma of biology, but equally crucial to the life of all organisms, is the collection of small molecules that a cell must acquire from the environment or produce metabolically. The myriad of tasks for which cells require small

molecules includes providing and storing energy, intra- and inter-cellular signalling, acting as enzymatic cofactors and providing building blocks for the construction of macromolecules. A great deal of effort has also gone into developing technologies to detect cellular metabolites, however this is an inherently more difficult undertaking than the detection of macromolecules. Proteins and nucleic acids are large molecules that are made up repeating subunits. This uniform constitution leaves these molecules with common physical properties that facilitates the development of standardized separation, detection and quantification methods. Another significant advantage for detecting proteins and nucleic acids is the availability of purified enzymes that can be used to synthesize or chemically modify these polymers *in vitro*. Finally, nucleic acids and proteins benefit from having a reliable source of recognition elements; antibodies can often be raised to recognize a specific protein of interest and nucleic acids are intrinsically recognized by DNA or RNA molecules with a complementary sequence. The lack of such a reservoir for the detection of metabolites, as well as their small size and diverse chemical composition, can make quantifying a particular small molecule within a population of cells a substantial challenge.

The traditional approach to measuring metabolite levels involves lysing a population of cells, followed by separation and analytical procedures to isolate and quantify molecules of interest. The nature of the separation and analytical steps can vary greatly. Commonly, a combination of extraction and chromatographic steps are applied to clear away cellular debris and to isolate molecules of interest or to segregate them into less complex mixtures (Bajad et al., 2006; Maharjan and Ferenci, 2003; Robards, 1995).

Individual molecules can subsequently be identified and quantified on the basis of their physical properties by applying techniques such as mass spectrometry or spectroscopic methods (Dettmer et al., 2007; Tang et al., 2004; Wittmann et al., 2004). Protocols following this basic scheme continue to dominate the field of small molecule detection and have many benefits. For many common molecules, detailed publications are available that describe suitable quantification protocols. Additionally, lysis-based methods generally provide quantitative data that, applying estimates of the number of cells used, can be used to approximate the absolute cellular concentration of compounds. Another benefit is that multiple metabolites within a single sample can often be detected. This premise has been taken to its extreme in recent years with the emergence of the field of metabolomics (Goodacre et al., 2004; Reaves and Rabinowitz, 2010). Impressive steps have been made in mass spectrometry-based techniques to simultaneously identify and quantify large numbers of cellular metabolites from cell extracts (Han et al., 2009; Mapelli et al., 2008). While these methods are clearly the best choice for experiments that seek a broad understanding of the metabolic status of the cell, they are not optimized to provide the most accurate possible information on an individual molecule, particularly for less abundant metabolites.

While very useful and informative, lysis-based methods of metabolite detection also suffer from many drawbacks. The cellular environment is extremely complex, containing thousands of distinct molecules. Procedures to isolate a single molecule can therefore require many steps and can be labour-intensive, expensive and time consuming. More involved procedures can also limit the number of samples that can practically be

processed in parallel. The instrumentation required for many of these procedures is very expensive and often requires a great deal of expertise; for many researchers a collaborator must therefore be found in order to carry out metabolite measurements. Furthermore, molecules can be lost during lysis and separation steps due to aggregation, degradation or their association with macromolecules or cellular debris. This can lead to inaccuracies and misleading results and, for chemically labile molecules, lysis-based detection procedures are often not possible.

An alternative to lysis-based procedures to detect metabolite levels is to devise ways of performing the detection within intact cells. While there are obvious challenges associated with introducing a suitable detection system into cells, this format overcomes many of the limitations described above for *in vitro* methods. Many varieties of *in vitro* biosensors have been created that couple the binding of a ligand by a recognition element to the production of a detectable signal (Cooper, 2004). The prospect of moving such a system into living cells is an appealing one. In order to accomplish this, however, two difficult challenges must be overcome: delivering the sensors into cells in a consistent and minimally invasive manner and maintaining reliable detection within the complex cellular environment. A number of strategies have been investigated that can be used to deliver nano-scale cargo into cells including microinjection, electroporation or chemical modifications with cell penetrating peptides or lipids (Chou et al., 2010; Delehanty et al., 2009). These approaches are often technically challenging, invasive, inefficient and may not be suitable for use in larger populations of cells.

To circumvent the difficult problem of permeating cells, small molecule sensors that are expressed from DNA vectors within the cell have been developed. These genetically-encoded sensors consist of engineered proteins that produce a fluorescent or luminescent signal in response to the binding of a specific target metabolite (Deuschle et al., 2005; Morris, 2009). This strategy is relatively non-invasive, as it requires only the expression of a nontoxic protein. Examples of some prominent strategies employed for the construction of genetically-encoded metabolite sensors are shown schematically in Figure 1-1. In each instance, the engineered sensor protein contains a well characterized, naturally evolved receptor domain that is responsible for recognition of the target metabolite. In many designs, the receptor is fused to fluorescent or luminescent reporter proteins that have been split into two inactive parts (Ozawa, 2006). Binding of the target metabolite causes a predictable structural change in the receptor domain that brings together the split reporter protein, restoring its activity and therefore producing a signal (Taneoka et al., 2009). Alternatively, two fluorescent proteins with compatible spectra are used in place of the split reporter. In these designs, binding of the metabolite brings together the fluorophores to allow for fluorescence resonance energy transfer (FRET) (Fehr et al., 2003; Honda et al., 2001; Zhang et al., 2002). Quantification of the reporter signal or the FRET therefore provides information on the cellular concentration of target metabolite.

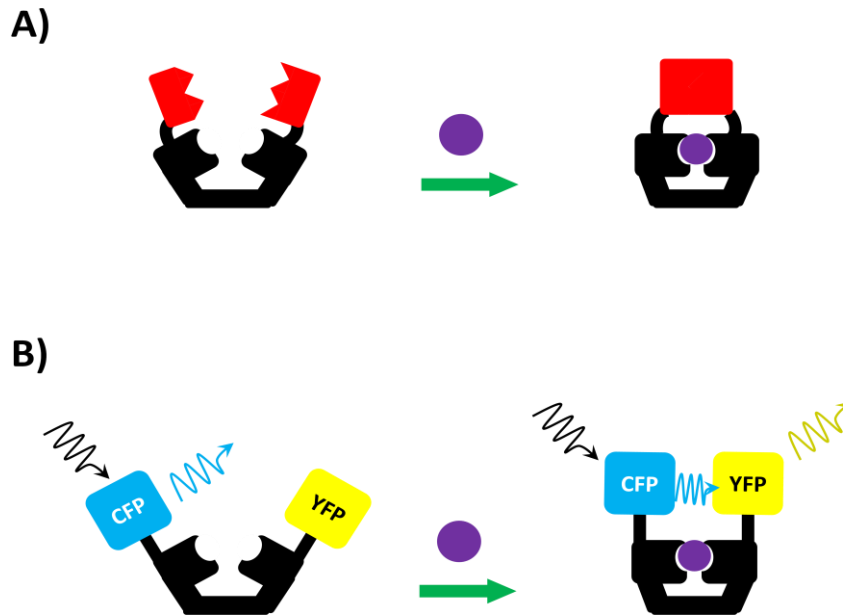


Figure 1-1 Schematic of general designs for protein-based genetically encoded metabolite sensors. General strategies that have previously been employed to detect metabolites within cells using genetically encoded, engineered reporter proteins. (A) A ‘split’ reporter protein (red), in which the active protein is split into two inactive domains, is reassembled in response to the binding of the target metabolite (purple circle). A well characterized binding protein (black) that undergoes predictable structural changes is used for molecular recognition. (B) A similar strategy that uses FRET rather than changes in reporter signal intensity to communicate ligand concentrations. Two fluorescent proteins that are compatible for FRET are tethered to the binding protein. In the unbound state, the distance between the fluorescent proteins is too great to allow energy transfer to occur. Ligand binding promotes structural rearrangements that bring the two fluorescent proteins closer together, producing a FRET signal.

These elegantly designed sensors have many beneficial qualities, but are also subject to some limitations. The fluorescent sensors are very useful when applied in conjunction with imaging technologies and can be applied to examine single cells with

good time resolution. However, the dynamic range and signal amplitude of these sensors are generally very limited (Deuschle et al., 2005). Luminescent sensors have shown superior signalling properties for the detection of target concentrations in cell populations. A drawback to luminescent sensors, which have been more commonly applied to studying protein-protein interactions using a bimolecular design, is that the reporter enzyme requires appropriate levels of ATP, magnesium and the substrate luciferin. Because the detection is based on a transient interaction, these assays must take place within the cellular environment and cannot follow conventional *in vitro* protocols that use cell lysates. Conditions that affect cellular magnesium or ATP concentrations or that affect the diffusion of luciferin into cells can therefore skew the results (Takeuchi et al., 2010). Perhaps the most significant drawback of each of these designs is that they require a thoroughly characterized receptor that undergoes a significant conformational change upon ligand binding. For the majority of interesting metabolites such a receptor is not known and it is not clear how sensors following this scheme could be constructed.

The work presented in this thesis investigates the possibility of using genetically encoded sensors where the detection is performed by RNA rather than proteins. These probes are intended to sense the cellular concentrations of a target metabolite within a population of bacterial cells. Like the protein-based sensors, the signal from reporter proteins is quantified to provide information on metabolite concentrations. In our RNA-based strategy, however, the target molecule is detected by an RNA receptor, which regulates the expression of the reporter protein in response to the binding event. Changes in metabolite levels therefore result in changes in the number of reporter proteins

produced, which can be quantified using fast, simple and inexpensive assays. The design is flexible with respect of the nature of the reporter protein used and is amenable to high-throughput approaches. The prospect of engineering such sensors to respond to a metabolite for which a natural receptor does not exist is also investigated in this thesis. Adaptability to new and interesting metabolites is an enormous potential advantage of these probes that helps overcome the principal limitation of protein-based designs.

The following sections of this introduction are intended to provide background information on the concepts underlying the construction and activity of these probes and the molecules that they are used to detect. These sensors depend on the recognition capabilities of RNA receptors known as aptamers, which are described in section 1.2 of this introduction. To produce a measurable signal, aptamers are used as regulatory elements, dubbed riboswitches, that control the expression of reporter genes. Naturally-occurring riboswitches are discussed in section 1.3 and efforts to produce synthetic riboswitches are discussed in section 1.4. The target metabolite of the riboswitch sensors created in this thesis is an active form of vitamin B12. Appropriate background on vitamin B12 biology is provided in section 1.5 of this introduction. Finally, the specific hypotheses and research goals of this work, as well as the organization of this thesis, are outlined in section 1.6.

1.2 The isolation, properties and applications of aptamers

Interacting selectively with other molecules is an important aspect of any protein's function. Certain classes of proteins, such as antibodies and receptors, are particularly celebrated for their ability to specifically recognize target ligands using carefully tailored binding pockets. Nucleic acids, by contrast, have traditionally been viewed as inert, passive molecules. This image has changed greatly over the past few decades thanks to the discovery of a remarkable array of functionally diverse nucleic acid molecules. With respect to molecular recognition, it is now clear that DNA and RNA molecules are capable of forming complex structures with highly specific ligand binding pockets, much like their protein receptor counterparts (Breaker, 1997; Ellington, 1994; Tucker and Breaker, 2005). These molecules, dubbed aptamers, have been isolated in laboratory experiments to bind an assortment of biologically-relevant targets. This section provides a brief review of artificial aptamers, focusing on the techniques used to isolate them, their characteristics and their potential applications. Naturally evolved aptamers, which have been discovered as components of genetic regulatory systems, are described separately in section 1.4.

The natural evolution of the first aptamers is thought to precede the evolution multi-cellular life forms (Breaker, 2010). Our introduction to the concept of aptamers, however, did not come until the early 1990s when the first man-made aptamers were isolated (Ellington and Szostak, 1990; Tuerk and Gold, 1990). In the twenty years since this time, the basic approach used to generate synthetic aptamers has not changed. This

methodology, often referred to as '*in vitro* selection' or 'SELEX' (Systematic Evolution of Ligands by EXponential enrichment), is currently used by many laboratories to isolate nucleic acids with a wide range of interesting properties. The basic goal of SELEX is to isolate rare molecules that carry out a desired function from a large, random-sequence library of oligonucleotides (Joyce, 1994). To accomplish this, SELEX employs iterative cycles (or 'rounds'), each of which consists of an activity-based selection step, followed by PCR amplification to regenerate a large library. The same basic SELEX principle has been routinely used to identify both aptamers and nucleic acids with catalytic capabilities. For aptamer-seeking experiments, each cycle of SELEX is intended to further enrich the surviving collection of molecules for oligonucleotides that specifically bind a target molecule with high affinity. In a successful selection, this pool will consist of a large fraction of active molecules after several cycles, at which point individual oligonucleotides can be cloned and studied. A typical experimental scheme for an RNA aptamer-seeking SELEX experiment is shown in Figure 1-2.

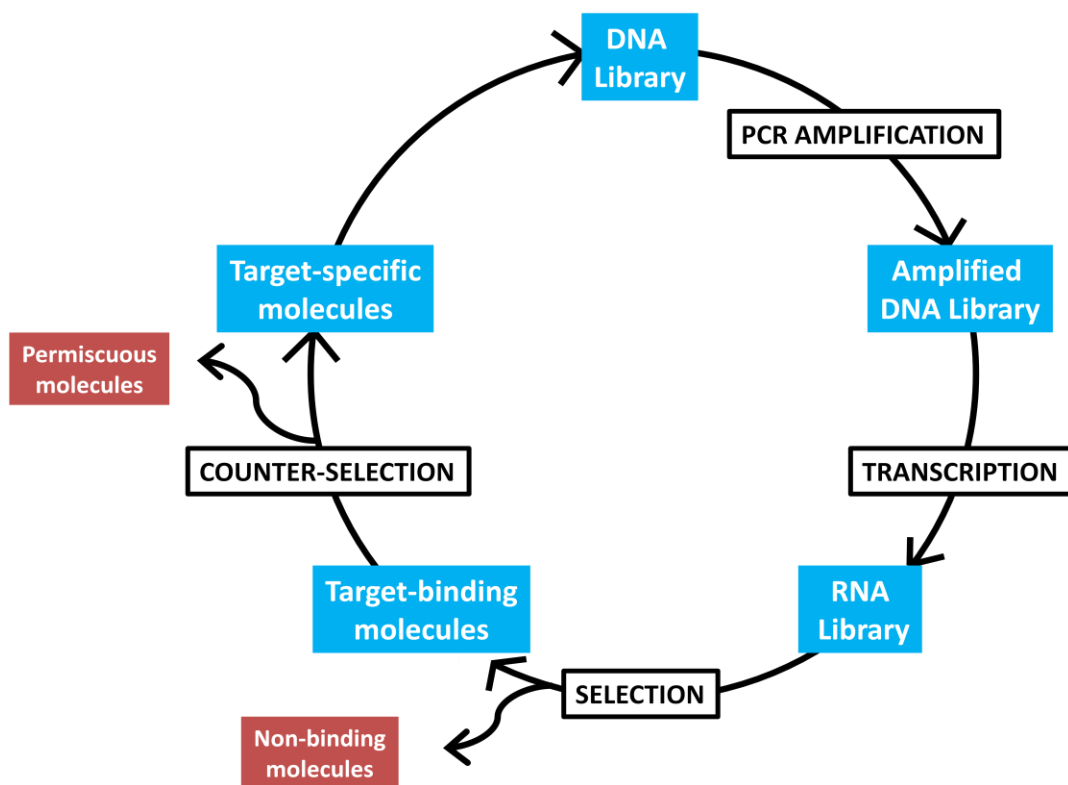


Figure 1-2 A typical SELEX scheme for the isolation of RNA aptamers. A large, diverse library of random-sequence DNA oligonucleotides is amplified, transcribed into RNA and is then passed through selection and counter-selection steps to identify specific, target binding RNA species. Selection steps identify target binding oligonucleotides on the basis of a physical interaction and counter-selection step remove promiscuous binders on the basis of a physical interaction with non-target molecules. Iterative cycles of this selection process are repeated; each cycle is intended to enrich the library for desirable sequences. After several cycles the library is cloned and sequenced and individual RNA molecules are characterized.

The most important aspect of any SELEX experiment is the selection step. To identify high quality aptamers, the selection step must be stringent enough to ensure that the bulk of the non-binding molecules are eliminated, yet not so strict as to risk losing potentially interesting sequences that might be present at very low copy. An assortment

of different methods have been used for this step, each of which physically separates oligonucleotides interacting with the target from the pool of unbound molecules (Gopinath, 2007). One commonly used format involves immobilizing the target ligand on solid supports such as columns or magnetic beads. Following incubation with the library, the solid support is washed to remove unbound oligonucleotides. Desirable molecules are then eluted from the support and passed on to the amplification step of the selection cycle. Counter-selection steps, in which promiscuous binders are removed on the basis of their interaction with a ligand other than the intended target, are commonly applied to ensure specificity.

An enormous advantage of SELEX is its flexibility with respect to methodology, binding conditions and library design. In addition to the format used for the selection and counter-selection steps, other types of modifications can also be introduced to the standard *in vitro* selection scheme (Bittker et al., 2002). Mutagenic PCR conditions, introduced into the amplification step, can be used to create library diversity. This approach is usually reserved for late rounds of selection and is intended to introduce subtle mutations to active molecules to improve their binding characteristics. Although fixed regions must be maintained for PCR amplification, the possibilities for library design are endless. Because the starting library is synthesized chemically, each residue may be designed to be completely random, fixed or bias such that certain nucleotides are more highly represented. Libraries can therefore be designed to be predisposed to contain desirable structural or functional motifs (Davis and Szostak, 2002).

Another common SELEX tactic is the incorporation of unnatural nucleotides into the oligonucleotide library (Keefe and Cload, 2008). Unnatural nucleotides can introduce additional chemical diversity to the library and can also produce molecules that are more stable and less susceptible to nuclease degradation. This latter point is particularly important for aptamers being developed for therapeutic purposes, since the half life of standard oligonucleotides *in vivo* is extremely limited (Nimjee et al., 2005). The 2' position of the ribose is a commonly substituted moiety for RNA aptamers, where the hydroxyl group is often exchanged for a fluoro or an O-methyl group among other possibilities. Modified nucleotides with chemical alterations to the phosphate or to the nucleobases have also been investigated (Keefe and Cload, 2008).

Over the past two decades many aptamers with highly attractive properties have been successfully isolated. The targets of these aptamers generally fall into two categories, proteins and small molecules. More recently, researchers have also begun to seek aptamers that recognize specific cell types or strains (Cao et al., 2009; Phillips et al., 2008). Because of the simple polymeric nature and limited chemical diversity of aptamers, ligands with particular features tend to be more or less compatible with aptamer binding. Ligands that make good aptamer targets include aromatic compounds, which can participate in stacking interactions, and positively charged molecules, which are compatible with the negatively charged nature of nucleic acids (Wilson and Szostak, 1999). By contrast, negatively charged or highly hydrophobic molecules present a greater challenge. Table 1-1 depicts an assortment of some aptamers of special interest. This table is not intended to be an all-encompassing list, but instead to illustrate the

diversity of biologically interesting ligands that have been successfully targeted and to present some key features of these aptamers.

Target	Target Type	K _d	Comment
Cibacron Blue 3G-A	Small molecule	100 μM	Amongst the first RNA aptamers isolated
Reactive green 19	Small molecule	30 μM	Amongst the first DNA aptamers isolated
Theophylline	Small molecule	300 nM	An aptamer that is highly specific against structural analogues
L-arginine	Small molecule	300 nM	An aptamer that is highly enantioselective
Neomycin	Small molecule	100 nM	Aptamer provided insight into activity of an RNA-binding antibiotic
ATP	Small molecule	10 μM	Widely adopted as a model small-molecule binding aptamer for bioanalytical applications
Keratinocyte growth factor	Protein	300 pM	2'-substituted RNA aptamer with very high affinity
Thrombin	Protein	50 nM	“Toggle SELEX” used to create aptamers that recognize the target from multiple species
Vascular endothelial growth factor (VEGF)	Protein	50 nM	2'-fluoro substituted RNA aptamers that show <i>in vivo</i> inhibition of VEGF; one of the aptamers was eventually developed as a drug known as “Macugen” (also known as “pegaptanib”) for ocular vascular disease
MUC1	Protein	100 nM	A DNA aptamer for a cancer marker
Thrombin	Protein	20 nM	Widely adopted as a model protein binding aptamer for bioanalytical applications
RET-receptor kinase	Protein	40 nM	An aptamer that blocks signalling of its cognate target in cell-based assays
Human influenza virus	Virus	200 pM	An aptamer that distinguishes between closely related viral strains
African trypanosomes	Protozoa	70 nM	An aptamer that recognizes a region on the surface of a dangerous blood parasite
Mesenchymal stem cells	Mammalian cell	N.D.	Developed for use in selectively purifying mesenchymal stem cells from bone marrow

Table 1-1 A sampling of aptamers of special interest. A list of some notable aptamers is provided along with their targets, their affinity for their target and a brief comment on their relevance. N.D., not determined.

The two principal determinants of an aptamer's quality are its affinity and specificity for the intended ligand. As indicated by the table above, *in vitro* selection has regularly produced oligonucleotides with low- or sub-micromolar binding constants. While the requisite binding affinity for an aptamer depends on its intended application, this affinity range is typically compatible with the therapeutic and biosensing applications for which aptamers are best known (discussed below). With respect to specificity, aptamers are also highly competitive with their protein receptor counterparts. The most famous example of aptamer specificity comes from the TCT8-4 aptamer, which binds its target theophylline with 10,000-fold higher affinity than caffeine, which differs by a single methyl group (Jenison et al., 1994). This is not an isolated example, however. Aptamers that strongly discriminate against homologous proteins (Gopinath et al., 2006), stereoisomers of a compound (Geiger et al., 1996) and on the basis of the phosphorylation state of a protein (Seiwert et al., 2000) have been reported. A detailed analysis of affinity and specificity is an important aspect of the identification of any aptamer. It is important to be wary of published accounts of aptamers that have not been thoroughly characterized in these respects.

The attractive binding properties exhibited by aptamers are due to their ability to adopt intricate three dimensional configurations with customized binding pockets. Several biochemical methods are available to probe the structural nature of SELEX-derived aptamers. Mutational analysis, methylation interference, hydroxyl radical footprinting, in-line probing and circular dichroism are common techniques that have been used to provide valuable structural information. For a more complete and definitive

understanding, however, high resolution structural techniques are required. Several NMR or X-ray crystallographic structures of artificial aptamers in complex with their protein or small molecule ligands have now been solved (Hermann and Patel, 2000; Patel and Suri, 2000; Patel et al., 1997). While the specifics of ligand binding vary from aptamer to aptamer, some common themes have been observed. Foremost amongst these is the observation that aptamers tend to be relatively less structured in their unbound form, but form highly ordered structures upon exposure to their target. In many cases, the ligand is thought to play an integral role in aptamer folding and the resulting structures feature the ligand deeply immersed within the aptamer (Patel et al., 1997). This is in contrast to proteins receptors, which tend to have pre-formed binding pockets at their surface. Aptamers exploit the same general arsenal of intermolecular forces to recognize their targets as proteins. It has been noted, however, that relative to protein receptors aptamers rely less on hydrogen bonding and acid-base interactions, and more on stacking interactions and a deep engulfment of the ligand (Hermann and Patel, 2000).

Other than experiments aimed to improve or test *in vitro* selection methodology, aptamers are usually isolated with a specific downstream application in mind. Since aptamer SELEX is essentially an elaborate affinity purification method, it seems intuitive that high quality aptamers can in turn be used to purify their target. Successful examples of this application include the purification of a specific enantiomer from a racemic mixture (Michaud et al., 2003), protein purification from cell lysates (Romig et al., 1999) and the isolation of a specific cell-type from a heterogeneous population (Guo et al., 2006).

One of the most exciting developments in the field of aptamers has been the isolation of molecules that bind drug targets to provide a therapeutic benefit. Although small molecules have dominated the therapeutics market, nucleic acids represent a new and interesting material for the development of novel drugs (Bunka et al., 2010; Lee et al., 2006). The first aptamer drug appeared on the market in 2005 after successfully completing clinical trials for the treatment of age related macular degeneration. The drug, dubbed “Macugen”, binds the human vascular endothelial growth factor and inhibits its ability to promote unwanted growth of new blood vessels in the eyes (Zhou and Wang, 2006). At least eight other aptamers are currently in clinical trials and many other therapeutically relevant aptamers are also in the late stages of development (Esposito C.L., 2011). Molecules currently in clinical trials include anti-coagulants, cancer treatments and diabetes treatments (Bunka et al., 2010).

Aptamers also have a celebrated ability to be applied as sensors for the molecules they bind. Because nucleic acids can be chemically modified to include a signal-producing moiety or tethered to any number of surfaces, a remarkable array of designs have been used for aptamer-based *in vitro* biosensors (Navani and Li, 2006; Sefah et al., 2009). Fluorescence is amongst the most popular signalling methods used for nucleic acid biosensors (Juskowiak, 2011). Aptamers featuring covalently attached fluorophores can be applied to detect their ligand via fluorescence anisotropy for larger molecules (Fang et al., 2001). Alternatively, designs that exploit fluorescence quenching or FRET can be achieved using fluorescence quenchers or additional fluorophores (Navani and Li, 2006). Aptamers tethered to quantum dots, which possess impressive fluorescence

properties, have also been employed in similar fluorescence designs (Levy et al., 2005). In addition to fluorescent signals, other successful aptamer sensor designs include modified versions of standard enzyme-linked immunosorbent assays (ELISAs) (Drolet et al., 1996), acoustic (Gronewold et al., 2005) or electrochemical sensors (Bang et al., 2005) and nanoparticle-conjugated aptamer sensors (Zhao et al., 2008).

Another exciting facet of the field of aptamers is the prospect of employing them as tools for synthetic biology. Of particular interest is the possibility of using aptamers to control gene expression in response to specific chemical cues. Naturally evolved regulatory systems that exploit aptamers in this manner have been discovered and vigorously researched over the last decade. The molecular recognition and regulatory activities of these elements are the foundation for the metabolite sensors that headline this thesis. The following section provides an introduction to naturally evolved regulatory elements that exploit aptamers to detect small molecules.

1.3 Riboswitches: metabolite-sensing RNA regulatory elements

Prokaryotic cells make use of a wide range of mechanisms for controlling gene expression in response to environmental cues. It is now clear that RNA regulatory elements play an enormous role in these processes. Dozens of characterized *trans*-encoded small RNAs (sRNAs) and *cis*-encoded antisense sRNAs have been described in prokaryotic organisms (Beisel and Storz, 2010; Thomason and Storz, 2010). Deep

sequencing and genomic mining efforts have identified many more of these elements, which await experimental confirmation (Sharma and Vogel, 2009). sRNAs identify target mRNA sequences at least partly through standard Watson-Crick base-pairing. Upon binding a transcript, sRNAs modulate the expression of its genetic products at the level of translational initiation, RNA stability or in some cases both (Beisel and Storz, 2010; Thomason and Storz, 2010). Generally, sRNAs are themselves regulated at the level of transcription and their role in the regulatory process is less about making regulatory decisions and more about the execution of those decisions.

Another renowned class of prokaryotic regulatory RNAs, dubbed riboswitches, specializes in both making and executing regulatory decisions (Nudler and Mironov, 2004; Winkler and Breaker, 2005). Riboswitches are typically located within the 5' untranslated region (UTR) of mRNAs where they regulate the downstream gene or operon in *cis*. Regulation is achieved using a naturally evolved RNA aptamer that is embedded within the transcript, which senses the concentration of a specific metabolite. Depending on whether or not the aptamer is bound to its ligand, the 5'UTR adopts one of two distinct conformational states. The basis of riboswitch regulation is that one of these conformations inhibits the expression of the downstream gene or genes, while the other does not. Over the responsive range of the aptamer, increasing concentrations of the target metabolite result in a larger fraction of riboswitches in the bound state. Depending on the nature of the riboswitch, higher ligand concentrations will therefore result in either an increase or a decrease in the expression of associated genes.

The regulatory mechanisms of prokaryotic riboswitches fall into three main categories: transcriptional, translational and catalytic (Barrick and Breaker, 2007) (see Figure 1-3). Translational riboswitches (Figure 1-3A), the predominant class in Gram negative species, act at the level of translational initiation (Winkler and Breaker, 2003). Riboswitches that work by this mechanism adopt a restrictive conformational state in which the ribosome binding site (RBS) of the regulated gene is highly structured. The RBS is said to be “sequestered” in this state, which prevents ribosome binding and thus translation. This mechanism is similar to that employed by many trans-acting sRNAs. Relative to the other two riboswitch regulatory mechanisms, translational riboswitches have some interesting properties. First, translational riboswitches are less well suited to control operons of genes, since individual genes in operons normally have their own RBS. Second, regulatory decisions made by translational riboswitches are potentially reversible, since the dynamic nature of ligand binding could permit conformational states to switch over the transcript’s lifetime (Smith et al., 2010). This is not the case for transcriptional and catalytic riboswitches, whose regulatory decision for a given transcript is final.

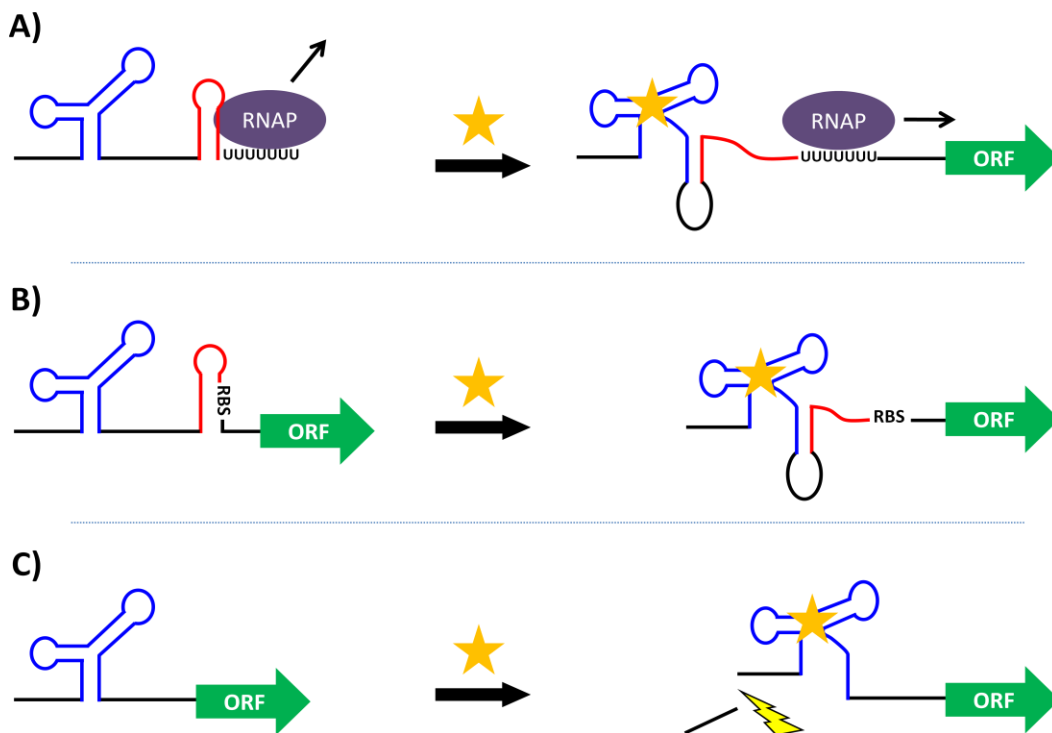


Figure 1-3 Schematic representations of the three general regulatory mechanisms employed by bacterial riboswitches. (A) Transcriptional regulation. The binding of the target metabolite (star) affects the formation of an intrinsic transcriptional terminator (stem-loop structure followed by a string of uridine residues). While RNA polymerase (RNAP) is paused at the uridine-rich sequence, the formation of the stem loop structure results in RNAP dissociation from the template and the regulated genes are not transcribed. (B) Translational regulation. The binding of the cognate metabolite affects the formation of a stem-loop structure that blocks ribosome access and thus translational initiation. RBS, ribosome binding site. (C) Catalytic mechanism in which the target metabolite regulates the activity of a self-cleaving ribozyme. Cleavage can regulate the expression of downstream genes in different ways and the metabolite can act as either a coenzyme or an allosteric regulator of cleavage, depending on the riboswitch (see text). Ligand-induced repression is shown for panels (A) and (B), however synonymous mechanisms in which ligand binding activates expression are also common.

Transcriptional riboswitches (Figure 1-3B) are the predominant category found in Gram positive bacteria (Winkler and Breaker, 2003). The source of regulation for

transcriptional riboswitches is a rho-independent transcriptional terminator. Rho-independent transcriptional terminators are comprised of a simple stem-loop structure which is followed by a string of uridine residues. During transcription, the uridine-rich stretch acts as a pause signal for RNA polymerase. This pause allows for the formation of the terminator stem-loop structure, which in turn dislodges RNA polymerase from the template, preventing the transcription of downstream sequences (Platt, 1986). Ligand binding in a transcriptional riboswitch results in either the formation or disruption of the terminator stem-loop structure. The permissive structural conformation of the riboswitch therefore allows for the transcription of downstream genes and the inhibitive structure does not. The activity of transcriptional riboswitches is, at least in some cases, kinetically driven (Lemay et al., 2011; Wickiser et al., 2005). This is a result of the rate of transcription forcing the regulatory decision to take place before the ligand-aptamer interaction has reached thermodynamic equilibrium. An extension of this concept is that, for these riboswitches, the apparent K_d values determined *in vitro* are not necessarily representative of the responsive range of metabolite concentrations within living cells.

The third and final category of riboswitches are ribozymes that catalyze a self-cleavage reaction at a defined linkage in their backbone (Ferre-D'Amare, 2010; Lee et al., 2010)(Figure 1-3C). Only two known classes of ribozyme riboswitches have been identified, making this category far less common than the others. The first ribozyme riboswitch to be reported responds to glucosamine-6-phosphate (GlcN6P) (Winkler et al., 2004). Unlike all other riboswitches discovered to date, the mechanism of this riboswitch does not rely on structural changes in response to ligand binding. Instead, GlcN6P acts as

a cofactor in the catalysis of the cleavage event (Hampel and Tinsley, 2006). The cleavage reaction produces a downstream product - the product that encodes the regulated genes - with a 5' hydroxyl group. This unusual 5' moiety is recognized by RNase J, which leads to transcript degradation and thus a decrease in the expression of the regulated genes (Collins et al., 2007). The other ribozyme riboswitch class, which responds to the second messenger cyclic di-GMP, was discovered quite recently (Lee et al., 2010). In contrast to the GlcN6P riboswitch, cyclic di-GMP does not participate in the catalytic mechanism, but instead acts as an allosteric regulator of ribozyme activity (Lee et al., 2010). The link between cleavage and regulation is not as well studied for this class of riboswitches. The initial hypothesis proposes that cleavage produces a transcript with increased RBS availability, leading to an increase in the translation of the downstream gene.

The term riboswitch has been reserved for RNA regulatory elements that directly bind small molecule metabolites. The distinction between riboswitches and related RNA elements that act in a similar manner is a blurry one at best. Long before the term riboswitch was first coined in the early 2000s, many *cis*-acting regulatory RNAs had been uncovered that directly bind specific molecules to elicit expressional changes. Protein binding RNA regulatory elements with similar mechanisms to those described for riboswitches are common in prokaryotes (Geissmann et al., 2009; Stulke, 2002). While the protein element has traditionally been viewed as the receptor and the RNA as its ligand in these systems, the emergence of aptamers makes this viewpoint debatable.

Instead, RNA-protein interactions such as these should be considered to be the result of co-evolution of the two elements to yield a beneficial interaction.

Another group of riboswitch-like regulatory elements whose discovery predates that of riboswitches are the tRNA-binding motifs found in 5' UTRs of numerous bacterial mRNAs (Green et al., 2010). Like riboswitches, these elements specifically bind a target molecule, which in turn dictates whether or not transcriptional terminator or translation-inhibiting helices form. Both charged and uncharged tRNAs bind to these motifs, however only uncharged tRNAs bring about expressional changes (Grundy et al., 1994). Charged tRNAs therefore compete with uncharged tRNAs for the binding site, allowing the RNAs to detect the ratio of charged to uncharged forms of a specific tRNA. The general role of these tRNA sensors is to activate expression of aminoacyl-tRNA synthetases as well as amino acid biosynthetic and uptake genes when the ratio of the uncharged tRNA is high. Despite the mechanistic similarities, these elements are not considered to be riboswitches by some because their ligand is intrinsically involved in gene expression. This position has been debated by others (Green et al., 2010), although in the end this argument is one of semantics.

The finding that riboswitches represent a very common and important regulatory mechanism for prokaryotes was spearheaded by Ronald Breaker's group at Yale University. The identification of the first riboswitches by this group was the result of mining the literature for instances where such a mechanism seemed likely (Mironov et al., 2002; Nahvi et al., 2002; Winkler et al., 2002a; Winkler et al., 2002b). Indeed, for

some riboswitch classes it had been previously proposed that regulation was the result of RNA-metabolite binding, however no direct biochemical evidence had been produced to support this claim (Lundrigan et al., 1991; Roth et al., 1996). To clarify these situations, the Breaker group demonstrated a direct and specific interaction between the RNA aptamers and their ligands. Moreover, they provided evidence that binding-induced structural rearrangements affected transcription or translation by the mechanisms described above (Nahvi et al., 2002; Winkler et al., 2002a; Winkler et al., 2002b). The identification of many such elements, all working by a similar mechanism, made it clear that riboswitches are not a rare exception, but rather one of the most abundant bacterial regulatory mechanisms (Barrick and Breaker, 2007).

A technique known as in-line probing has been very prominent in the riboswitch field (Regulski and Breaker, 2008). It has been instrumental in detecting RNA-metabolite interactions, determining apparent K_d values for these interactions and in providing clues about the bound and unbound conformational states of riboswitches. In-line probing takes advantage of spontaneous degradation of RNA due to internal transesterifications. These reactions occur more frequently at RNA linkages that are unrestrained, and are less common at sections of the RNA chain involved in secondary or tertiary structures (Soukup and Breaker, 1999). Spontaneous cleavage of radioactive RNA, monitored by PAGE, can therefore provide information about which segments of an RNA molecule are structurally affected by a stimulus. When an RNA aptamer binds a small molecule, concentration dependent changes can be observed in the pattern and intensity of certain spontaneous cleavage products. Metabolites that do not bind the RNA

produce no structural changes and thus the cleavage pattern is not affected. Furthermore, clues about the bound and unbound RNA structures can be gathered by inspecting the specific linkages that become more or less structured upon ligand binding. Using this information in conjunction with phylogenetically conserved structural features, it is often possible to predict putative regulatory mechanisms. An example of such a prediction is shown in Figure 1-4.

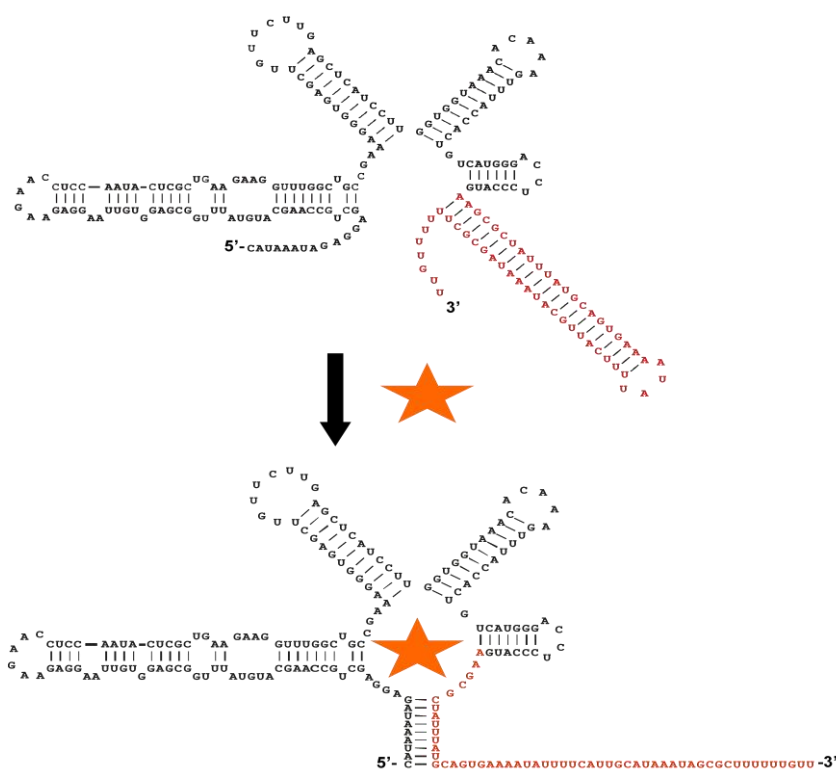


Figure 1-4 Secondary structures of predicted bound and unbound conformations of a lysine riboswitch. Using bioinformatic data on conserved structural features and results from structure-probing experiments such as in-line probing, putative regulatory mechanisms can often be proposed. The above structures show the bound and unbound structures of a predicted lysine-binding riboswitch from the organism *Thermoanaerobacter tengcongensis*. In the unbound structure (top) an intrinsic transcriptional terminator is predicted to form (red). Ligand binding disrupts this structure and therefore activates the expression of downstream genes. Predictions of the

secondary structures for this riboswitch and several others were provided in a list format in a previously published study (Sudarsan et al., 2003).

While the first handful of riboswitch classes were found with the assistance of published literature, bioinformatic approaches have since taken over as the principal means by which riboswitches are discovered (Weinberg et al., 2007; Weinberg et al., 2010). These approaches take advantage of the fact that the aptamer domains of riboswitches are highly conserved phylogenetically. While this conservation is not always detectable at the level of the primary sequence, computer programs that seek a combination of conserved sequence and secondary structure have been highly successful in identifying candidate riboswitches. As with literature mining approaches, candidate riboswitches must then be thoroughly assessed experimentally. In many instances the genomic context can be used to predict the identity of regulatory ligand, which can then be tested using *in vitro* binding assays and other follow up experiments. In other cases the identity of the putative regulatory molecule is not decipherable from the function of neighboring genes (Weinberg et al., 2010). Despite the fact that some such RNAs have all the hallmarks of a riboswitch, methods to identify target ligands for these riboswitch candidates have not yet been established (Meyer et al., 2011). While it is likely that many of these sequences do encode legitimate riboswitches, some might serve a different function or might recognize a protein rather than a metabolite.

Unique riboswitch classes are identified on the basis of the conserved aptamer domain, where each class employs an aptamer with unique sequence and structural

features. The total number of riboswitch classes that have been uncovered to date depends on the definition of riboswitch that is used as well as the criteria used to separate the classes. Counting only riboswitches that respond to organic metabolites and metal ions, but not mechanistically similar RNAs that respond to temperature, proteins or tRNAs, 23 riboswitch classes have been identified and experimentally validated (Table 1-2). It should be noted that this number counts a unique class for each ligand in the few cases where similar folds are employed to recognize distinct but related molecules. The list of riboswitch target molecules contains a mix of core metabolites, such as enzymatic cofactors and amino acids. Noteworthy exceptions to this are riboswitches that respond to the second messenger cyclic di-GMP and to the divalent metal ion magnesium. The list of riboswitch classes also contains a large number of molecules with nucleobase or nucleobase-like moieties (Dambach and Winkler, 2009). This is presumably due to the fact that many important metabolites fit this description and that these molecules are very suitable for aptamer recognition. It has been proposed that the current list of established riboswitches is biased to contain the most abundant classes, and that those that are less prevalent remain undetected (Breaker, 2010). This argument is based on the fact that most riboswitch classes have been found using bioinformatic algorithms that sift through genomic sequences. Since the ability of these programs to identify a new class is highly dependent on the number of representatives, riboswitches found in a low number of sequenced prokaryotic genomes are less likely to be detected.

Class	Affinity	Approximate aptamer size (nt)	Notes
Adenine	300 nM	70	Single base mutation switches specificity to guanine
AdoCbl	300 nM	200	Class used in chapters 3 & 4 of this thesis
c-di-GMP-I	1 nM	110	First riboswitch found that responds to a second messenger
c-di-GMP-II	200 pM	90	Allosteric group I ribozyme
Deoxyguanosine	80 nM	70	Discriminates vs. guanosine by ~ 3 orders of magnitude
FMN	5 nM	120	Strong discrimination vs. FAD & riboflavin
GlcN6P	200 μ M	170	First catalytic riboswitch uncovered
Glutamine	150 μ M	70	Only found in cyanobacteria & marine metagenomes
Glycine	30 μ M	110	Cooperative binding of two adjacent aptamers
Guanine	5 nM	70	Single base mutation switches specificity to adenine
Lysine (L-box)	1 μ M	175	Putative target of certain antimicrobial lysine analogs
Mg ²⁺ I (mgtA)	N.D.	100	First metal-sensing riboswitch uncovered
Mg ²⁺ II (M-Box)	200 μ M	70	Evidence suggests 6 distinct magnesium binding sites
Moco	N.D.	130	Metabolite not stable enough for testing <i>in vitro</i>
PreQ1-I	50 nM	40	Smallest conserved aptamer domain uncovered thus far
PreQ1-II	100 nM	85	Relatively rare class
SAH	20 nM	65	Discriminates vs. SAM by > 3 orders of magnitude
SAM-I	4 nM	100	Most widely distributed SAM riboswitch
SAM-II	1 μ M	60	Only found in actinobacteria and proteobacteria
SAM-III	N.D.	80	Only found in lactobacillales
SAM-IV	15 μ M	60	Found primarily in marine bacteria
THF	70 nM	100	Relatively rare class
TPP	100 nM	120	Present in all domains of life

Table 1-2 Riboswitch classes discovered to date. A list of the riboswitch classes that have been experimentally validated to date. The binding affinity, the size of the aptamer domain and a short note is provided for each class. It should be noted that binding affinities are approximate and are often dependent on the specific construct tested. Similarly, the size of the aptamer can vary considerably from one riboswitch to the next. N.D., not determined.

The vast majority of riboswitches uncovered thus far are found in bacterial species. For reasons that are not entirely clear, certain branches or organisms within the bacteria domain appear to rely much more heavily on riboswitch regulation than others (Barrick and Breaker, 2007). This is exemplified in the common model organisms *B. subtilis* and *E. coli*; *B. subtilis* employs ~30 riboswitches to regulate more than seventy genes, whereas the *E. coli* genome contains less than a third as many riboswitches that regulate fewer than twenty total genes (Barrick and Breaker, 2007; Rfam-database, 2011). Certain riboswitch classes are also much more highly represented and widely distributed than others. The two most abundant riboswitch classes, which recognize thiamine pyrophosphate (TPP) and coenzyme B12 (adenosylcobalamin or AdoCbl) respectively, are present more than three times per genome on average over the ~700 fully sequenced bacterial genomes and are found in a very wide range of distantly related species (Breaker, 2010). By contrast, within this same database there are less than 100 total instances of PreQ-II and THF classes, which are found in only a few branches of bacteria domain (Breaker, 2010). It is very common for a single riboswitch class to be present in a single genome multiple times. One notable example of this is the *Geobacter uraniumreducens* genome, which contains thirty separate cyclic di-GMP-responsive riboswitches controlling a diverse array of genes (Sudarsan et al., 2008).

In addition to their widespread importance for bacterial gene regulation, riboswitches have been found in eukaryotic, archaeal and bacteriophage genomes as well (Barrick and Breaker, 2007; Sudarsan et al., 2008). Although only the TPP riboswitch class has been found in eukaryotes thus far, it is very likely that additional examples

exist. Eukaryotic genomes are considerably larger and contain a great deal more noncoding RNA than those of prokaryotes. This makes identifying conserved structural motifs in noncoding regions considerably more difficult and goes beyond the abilities of the computer programs currently used to locate prokaryotic riboswitches (Weinberg et al., 2010). The regulatory mechanisms of the known eukaryotic riboswitches do not conform to any of the prokaryotic mechanisms, but instead rely on ligand-induced changes in RNA splicing (Wachter, 2010). However, the same premise that underlies prokaryotic riboswitch regulation is maintained; the structure of the transcript switches between two mutually exclusive conformations in response to ligand binding. To control RNA splicing, key processing signals are structurally sequestered in one conformation and made available in the other. Ultimately, alternatively spliced transcripts affect expression either at the level of RNA stability (Wachter et al., 2007) or by either creating an upstream open reading frame (ORF) that interferes with the expression of the primary ORF (Cheah et al., 2007). The TPP riboswitch class is present in all three domains of life, which is in strong support of an ancient origin for riboswitch regulation. In fact, prudent arguments have been made that suggest that riboswitches could be remnants of the hypothetical RNA world that is proposed to have preceded life in its current form (Breaker, 2010; Vitreschak et al., 2004). While such arguments are difficult – if not impossible – to prove, it remains an intriguing hypothesis that begins to answer questions concerning how RNA-based life forms could sense and respond to their environments.

While often used as stand-alone genetic control elements, riboswitches have also been observed as part of more complex regulatory schemes (Tucker and Breaker, 2005).

A simple riboswitch consisting of a single aptamer that forms a one-to-one complex with its ligand is restricted with respect to dose response. The simple ligand-receptor binding scenario expected in such instances implies that for a regulated gene to go from maximal expression to maximal repression, ligand concentrations would need to change by ~100-fold (Breaker, 2010). While this arrangement might be practical in some instances, a sharper regulatory response is beneficial in others. Furthermore, many genes benefit from regulation in response to multiple environmental cues. More complex arrangements of riboswitches have been found that begin to address these limitations. The *gcvT* glycine-responsive riboswitch from *B. subtilis* is comprised of two adjacent aptamers that bind glycine in a cooperative fashion to elicit a sharper regulatory response to smaller changes in ligand concentrations (Mandal et al., 2004). This arrangement appears to be common amongst glycine riboswitches. Several examples of tandem riboswitches that function independently have also been described, including both riboswitches that respond to the same target and those with distinct regulatory cues (Sudarsan et al., 2006; Welz and Breaker, 2007). Tandem arrangements of distinct riboswitches allow for responses to multiple environmental cues. Repeats of a single riboswitch class provide a somewhat sharper response to changing ligand concentrations and can also increase the dynamic range of the response (Coppins et al., 2007; Welz and Breaker, 2007). Additionally, riboswitches have been observed to regulate genes in conjunction with protein-binding RNA regulatory elements (Fox et al., 2009). The promoters that drive the expression of many riboswitch-regulated genes are also presumably subject to regulation, although this premise has not yet been thoroughly investigated.

The prevailing regulatory function for riboswitches is to provide simple feedback regulation. When its ligand accumulates beyond a certain level, it is recognized by the riboswitch resulting in the down-regulation of genes involved in its synthesis or uptake (Dambach and Winkler, 2009). A variation on this theme that is somewhat less common is ligand binding leading to the activation of genes responsible for the target molecule's efflux or degradation. Riboswitches that respond to certain enzymatic cofactors have also been found to regulate enzymes or pathways for which they are required (Regulski et al., 2008; Sudarsan et al., 2006). The cyclic di-GMP riboswitches that have been uncovered provide an example of more exotic regulation. Cyclic di-GMP is an important second messenger in bacteria that has drawn a great deal of interest due to its central role in regulating genetic programs that dictate the 'lifestyle' of a bacterium (Hengge, 2009; Yan and Chen, 2010). Cyclic di-GMP responsive riboswitches regulate genes involved in a number of important physiological processes such as virulence, motility, DNA uptake and biofilm formation.

Our understanding of the activity and specificity of riboswitches at the molecular level has been greatly advanced by high resolution structures that have been solved for the aptamer domain of several riboswitch classes (Montange and Batey, 2008). Some common features of these structures provide explanations for previously observed biochemical data. First, riboswitch aptamers usually contact nearly all of its ligand's functional groups, providing an explanation for the extremely high level of specificity observed for all riboswitch classes (Serganov, 2009). A second common observation is that aptamer sequences that are predicted to interact with gene expression controlling

elements in the unbound state tend to be crucial elements of the ligand binding pocket (Serganov, 2009). This is likely an important feature to ensure efficient ligand-induced structure switching. Relative to their artificially-derived counterparts, riboswitch aptamers tend to have increased affinity and specificity for their ligands. This can also be elucidated from the structural data, as riboswitch aptamers form larger and more elaborate architectures that allow for more extensive and precise contact with their ligand (Serganov, 2009).

Riboswitches are of great scientific interest for reasons that go beyond their status as unique and widespread regulatory elements. One source of interest regarding riboswitches is the possibility that they can be used as targets for the development of antimicrobial compounds (Blount and Breaker, 2006; Mulhbachter et al., 2010). The premise of this tactic is that analogs of riboswitch ligands that maintain riboswitch binding, but are not metabolically useful, could be used to cause misregulation of critical genes and pathways. Strong support for this strategy has been provided by certain established metabolite analogs that have antimicrobial activity. These include the thiamine analog pyrithiamine, the lysine analogs L-aminoethylcysteine and DL-4-oxalysine and the riboflavin analog roseoflavin. All of these antimicrobial compounds have been shown to tightly bind a riboswitch target and trigger expressional changes in riboswitch-regulated genes (Blount et al., 2007; Lee et al., 2009; Sudarsan et al., 2005). Importantly, strains selected to be resistant to these compounds have mutations that map to the riboswitch region and that disrupt the binding of the antimicrobial compound. It is therefore proposed that the mechanism of toxicity for these molecules, which was

previously not understood, might be to disrupt riboswitch-regulation. Recently, evidence was presented that lysyl-tRNA synthetase may be the primary target of the lysine analog L-aminoethylcysteine rather than the proposed lysine riboswitches (Ataide et al., 2007). This indicates that in some or all of these cases riboswitches might not be the only target of these compounds. Regardless, riboswitch-targeted analogs may yet prove to be a fruitful ground for the discovery of antimicrobial molecules.

Riboswitches are also of great interest due to their tremendous promise as tools for synthetic biology (Topp and Gallivan, 2010). Additionally, the central premise of this thesis is that riboswitches have a great deal of potential to be used as intracellular sensors of their cognate metabolite. The availability of riboswitches with suitable properties greatly influences their flexibility to be used in these capacities. Methods to modify or create novel riboswitches are therefore invaluable for these applications by expanding upon the arsenal of naturally evolved riboswitches. The following section provides a discussion of the engineering and potential applications of riboswitches.

1.4 The engineering and applications of bacterial riboswitches

Controlling gene expression using small molecule inducers or repressors is amongst the most common themes in design of recombinant genetic systems (Boyle and Silver, 2009). In its simplest form, inducible expression permits a single gene to be turned on or off in a simple and reversible manner. More elaborate scenarios such as

graded responses, genes regulated by multiple molecules and larger regulatory circuits are increasingly being employed to program bacteria to exhibit more complex behaviors (de Las Heras et al., 2010). Riboswitches have many attributes that make them very attractive options as regulatory elements for such applications (Topp and Gallivan, 2010). Additionally, the molecular recognition properties of riboswitches make them good candidates for construction of biological sensors. The established potential to modify riboswitches to create riboswitches that respond to a selected metabolite greatly expands their utility. This section discusses the methods that have been used to construct custom riboswitches as well as the applications of natural and artificial bacterial riboswitches.

Creating a riboswitch that responds to a particular molecule requires both an RNA aptamer that specifically binds the target and a means of linking this binding to a change in gene expression. Building a novel riboswitch completely from scratch, either by design or through cell-based screening or selections, is an appealing concept. However, the complexity required for a robust aptamer domain precludes design, and selection requires sampling a sequence space that is likely to exceed the library sizes that can currently be practically investigated within a living organism. The riboswitches that have been engineered to date have therefore employed established RNA aptamers, either generated *in vitro* or borrowed from natural riboswitches. From this starting point, a variety of design, screening and selection strategies can be employed to identify elements where aptamer binding results in genetic regulation.

Armed with a structural understanding of an aptamer, the prospect is presented of rationally designing riboswitches that mimic common natural mechanisms. This idea was exploited to produce one of the first artificial prokaryotic riboswitches in 2004 (Suess et al., 2004). Structural information concerning the well-studied theophylline-binding aptamer was used to design a riboswitch that worked at the level of translational initiation. The design was such that the RBS of the downstream gene was sequestered by Watson-Crick base pairing in the absence of theophylline. Theophylline binding at the adjacently positioned aptamer resulted in a shift in the inhibitory helix that allowed for greater ribosome access and an ~8-fold increase in expression.

Although the above study demonstrates the possibility of rationally designing a riboswitch, a more popular approach has been to use screening and selection strategies to identify active sequences from a library. Rational design still plays a large role in these methods in guiding the construction of a partially randomized library of riboswitch candidates. A number of creative ways to isolate candidates with the best regulatory activity have been employed (Topp and Gallivan, 2010). Common to each of these methods is the critical function of the gene placed under the control of the putative riboswitch library. In an active riboswitch, the expression of this gene will switch between “on” and “off” depending upon the concentration of the target metabolite. Selection and screening methods exploit this by alternating between growth conditions featuring high and low metabolite concentrations, and isolating clones where the gene displays the desired regulatory pattern. Examples of some specific screening methodologies that have been employed are shown schematically in Figure 1-5.

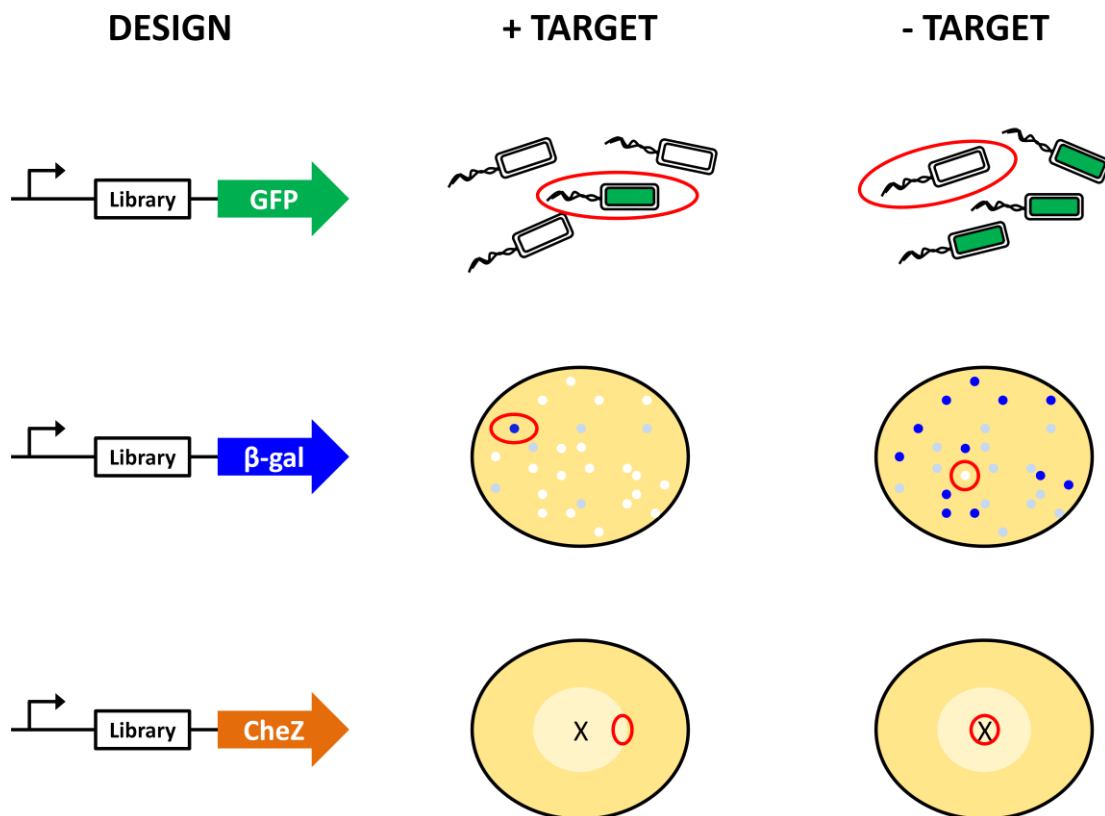


Figure 1-5 Schematic representations of methods used to isolate synthetic riboswitches. Three different methods to isolate active riboswitches from libraries of candidate sequences are shown. The library construction for each method is shown on the left. The phenotypes being sought by the selection process are shown in the center (+target) and on the right (-target). The red circles identify the clones that are selected at that selection step. The top panel shows a FACS-based method, where flow cytometry is used to quickly sort individual cells to identify those with a desirable level of fluorescence. The middle panel shows a scheme that relies on identifying expression levels for individual colonies on agar plates using β -galactosidase cleavage of X-gal, which produces a blue colored colony. The bottom panel shows a motility-based of selection in which the riboswitch library controls the expression of *cheZ*. The net movement of the bacteria plated on soft agar is proportional to the levels of CheZ. The X in this panel indicates the location at which the bacteria are initially stabbed into the agar.

One possible selection strategy borrows the classic genetics approach of providing a selective growth advantage to clones that demonstrate the desired activity. The Yokobayashi group at University of California Davis has developed and successfully exploited such a method using the *tetA* tetracycline resistance gene (Nomura and Yokobayashi, 2007a). While expression of *tetA*, which encodes a tetracycline/H⁺ antiporter, provides resistance to the antibiotic tetracycline, it also results in sensitivity to certain chemicals including nickel. This provides the means for a single gene to be used for selection for both the ON and OFF states. For example, to select for a riboswitches that are activated by the target metabolite, tetracycline is added in to the growth medium when high levels of the target molecule are used, and a suitable concentration of nickel is added for the selection step when target concentrations are low. Using this premise the Yokobayashi group has isolated a number of interesting synthetic riboswitches. This includes a rewired natural riboswitch that was altered to be activated by its ligand rather than repressed (Nomura and Yokobayashi, 2007b), as well as synthetic riboswitches that contain two aptamers and respond to two distinct molecules (Sharma et al., 2008).

Another approach that has been taken to identify ligand-responsive expression within riboswitch-seeking libraries has been to use reporter proteins. Reporter proteins produce a readily detectable signal that can be used to quantify their expression. Screening individual clones for ligand-responsive expression can therefore be achieved by quantifying reporter output under both high and low target molecule conditions. In order to screen large libraries, these techniques must employ high throughput methods that evaluate reporter expression in individual clones. Such an approach taken for the

first project of this thesis, for which green fluorescent protein (GFP) expression coupled with fluorescence activated cell sorting (FACS) was used to explore a library of candidate riboswitches for active sequences (Fowler et al., 2008) (described in detail in sections 1.6 & 2). Less than a year after this work was published, the Gallivan group at Emory University published a study that used a very similar FACS-based approach (Lynch and Gallivan, 2009). In this study a translational riboswitch that was activated >90-fold in the presence of excess target molecule was isolated. This remains the most efficient artificial prokaryotic riboswitch created to date.

The Gallivan group, a leader in the field of synthetic prokaryotic riboswitches, has employed other riboswitch selection strategies as well. Prior to exploring FACS, they developed a reporter-based method which used the familiar reporter protein β -galactosidase to seek theophylline-induced riboswitches (Lynch et al., 2007). β -galactosidase, encoded by the *E. coli* gene *lacZ*, catalyzes the breakdown of lactose to form D-galactose and D-glucose. Alternate substrates have been developed in which the glucose moiety has been replaced to make substrates that produce a color or become fluorescent or luminescent upon cleavage (Jain and Magrath, 1991; Zhang et al., 1991). 2-Nitrophenylgalactoside (ONPG), which produces a soluble yellow dye upon cleavage and 5-bromo-4-chloro-3-indolyl- β -d-galactopyranoside (X-gal), which produces an insoluble blue dye upon cleavage - two commonly used β -galactosidase substrates - were both used in this study. Theophylline-induced β -galactosidase expression was identified by examining individual clones grown on agar plates containing X-gal, followed by quantitative assays done in liquid culture using ONPG. Colony picking robots and liquid

handling instrumentation were used to increase the throughput of this screen. The end result of this study was the identification of riboswitches that were induced ~30-fold in the presence of excess theophylline.

Perhaps the most creative of the strategies that have been used to screen riboswitch libraries is an approach that uses cell motility to identify desirable clones (Topp and Gallivan, 2008). Chemotaxis, the movement of an organism along chemical concentration gradients, is a highly studied and relatively well characterized process in *E. coli* (Blair, 1995). Depending on the direction of rotation of their flagella, *E. coli* will either tumble in place or swim in a straight line. The protein CheZ plays an important role in regulating the transition between these states by dephosphorylating CheY-P. In its phosphorylated state, CheY binds to the flagellar motor and adjusts its rotation, causing cells to tumble. Higher levels of CheZ lead to increased dephosphorylation and inactivation of CheY and thus cells will tend to swim. By placing *cheZ* under the control of a library of potential theophylline-responsive riboswitches, active sequences were identified on this basis of their ability to migrate in response theophylline on semi-solid agar plates (Topp and Gallivan, 2008).

There are many common aspects of the studies described above, such as the use of *E. coli* as the host organism and theophylline as the target molecule. As one of the most highly studied organisms, *E. coli* has obvious benefits for developing riboswitch engineering methods. Foremost amongst these benefits are established techniques for high throughput cloning and transformation of stably carried plasmids. A recent study

used *E. coli*-isolated riboswitches that worked via the canonical translational mechanism and modified these riboswitches rationally to function in a wide range of distantly related species (Topp et al., 2010). This was accomplished by adjusting the sequence of both the RBS and the riboswitch elements with which it interacts to better suit the ribosome binding characteristics and 16S rRNA sequences of other species. A series of five plasmids were created that enabled efficient theophylline-responsive expression, at least one of which was suitable for use in the eight diverse bacterial species that were tested, including both Gram negative and Gram positive organisms. The repeated use of theophylline has also been highly convenient for method development purposes. The TCT-8-4 theophylline-binding aptamer is very well characterized, binds with high affinity and is very specific (Jenison et al., 1994; Zimmermann et al., 1997). Additionally, theophylline is an inexpensive, FDA-approved drug that is not a native metabolite in bacteria but can permeate cells when added to the media. Although theophylline has been a common choice, synthetic riboswitches that respond to other metabolites have been created (Nomura and Yokobayashi, 2007b; Sinha et al., 2010). As the synthetic riboswitch field continues to progress, future selections should increasingly focus on isolating riboswitches for organisms and targets that suit a specific downstream application.

The transition to application-driven synthetic riboswitch projects has begun, as a few interesting such studies have emerged over the last few years. In a very impressive endeavour, a synthetic riboswitch was recently engineered that provided an *E. coli* strain with the ability to seek out the hazardous herbicide atrazine (Sinha et al., 2010). This

project began with a SELEX selection for an atrazine-specific RNA aptamer, which was subsequently converted to a riboswitch using the motility-based selection strategy discussed above. By placing *cheZ* under the control of the resulting riboswitch, the bacteria migrated up chemical gradients of atrazine. Finally, the strain was also armed with an enzyme that converts atrazine to hydroxyatrazine, an analog that is not considered to be dangerous to human health. The resulting strain was therefore able to both seek out and destroy atrazine in a laboratory setting. A synthetic riboswitch was also recently employed in a more traditional role for a regulatory element. A theophylline-responsive riboswitch was used to control the chromosomal copy of the essential *E. coli* gene *csrA* in order to probe aspects of its function (Jin et al., 2009). Using this strain it was discovered that CsrA is an important regulator of autoaggregation in *E. coli*, a physiological activity that has been implicated in biofilm formation and virulence.

These reports demonstrate the potential utility of riboswitches, a concept that promises to continue to grow in the coming years. As small, direct, efficient and specific regulatory elements, they are attractive choice for recombinant use in a laboratory setting or in the development of engineered bacterial strains for industrial, environmental or biomedical applications. The prospect of combining riboswitches, either in tandem or with other regulatory elements, could allow for very tight regulation or genes regulated by multiple inducer or repressor molecules, enabling the design of more complex regulatory circuits. In addition to their potential as custom regulatory elements, their molecular recognition capabilities make riboswitches a natural choice for the

development of biological sensors. The aptamer domain of a riboswitch was recently used to construct an *in vitro* biosensor for TPP using a structure-switching, fluorescence-signalling design (Lau et al., 2010). The central hypothesis of this thesis is that riboswitches can also be employed as intracellular sensors of small molecules. AdoCbl-responsive riboswitches were used to explore this possibility, work described in chapters 3 and 4 of this thesis. Accordingly, an introduction to relevant aspects of AdoCbl-related biology is presented in the following section.

1.5 Vitamin B12 and its synthesis, uptake and utilization in *E. coli* cells

AdoCbl was an appealing choice as the target metabolite to test the concept of intracellular riboswitch-based sensors for several reasons. The AdoCbl class of riboswitches is amongst the most prevalent and widespread in bacteria (Breaker, 2010) and a characterized representative was available for *E. coli*, our organism of choice (Gallo et al., 2008; Nahvi et al., 2004; Nahvi et al., 2002). The *E. coli* AdoCbl riboswitch had previously been shown to be highly efficient, repressing expression by factors of ~100-fold in the presence of excess ligand (Lundrigan and Kadner, 1989; Nahvi et al., 2002). Binding mutants had been previously characterized, which would serve as useful controls (Nahvi et al., 2002). AdoCbl is also ideal as a target for proof of principle studies since it is not part of core metabolism and can be added or taken away at will. Although it is not essential for survival, AdoCbl has an interesting and somewhat mysterious function in *E. coli* biology. Finally, a large number of metabolic and transport

proteins are employed for AdoCbl acquisition, the activity of which could potentially be examined using the sensors. This section of the introduction presents background information on AdoCbl. The basics of its chemistry are described as well as its role in *E. coli* biology and the means by which it is produced and procured from the environment by *E. coli* cells.

The term ‘vitamin B12’ is often used as a generic means of referring to vitamin B12-like molecules, including both active forms and precursors. These large, chemically complex molecules are accentuated by a central cobalt atom, which is the crux of vitamin B12’s activity as an enzymatic cofactor. There are two predominant active forms of vitamin B12: methylcobalamin (MeCbl) and adenosylcobalamin (AdoCbl). The compound ‘vitamin B12’, or cyanocobalamin (CNCbl), is a functionally inert molecule that is prepared for human consumption. The chemical structure of these molecules, shown in Figure 1-6, is identical other than the moiety that provides the upper axial ligand for cobalt. In addition to its upper axial ligand, cobalt is coordinated by five additional interactions, four of which are in the planar porphyrin-like structure known as the corrin ring. The fifth interaction is provided by the lower axial ligand, which for standard cobalamins is 5,6-dimethylbenzimidazole (DMB). DMB is covalently attached to the corrin ring through a short appendage known as the nucleotide loop. Natural, functional analogs have been documented that contain lower axial ligands other than DMB (Hoffmann et al., 2000; Stupperich et al., 1988). For enteric bacteria, the most notable such analogs, known as *pseudo*-B12, feature adenine in place of DMB. Although DMB is used preferentially, in its absence *pseudo*-B12 can be produced using the same

enzymes and retains activity as a cofactor for at least some of its documented reactions (Anderson et al., 2008).

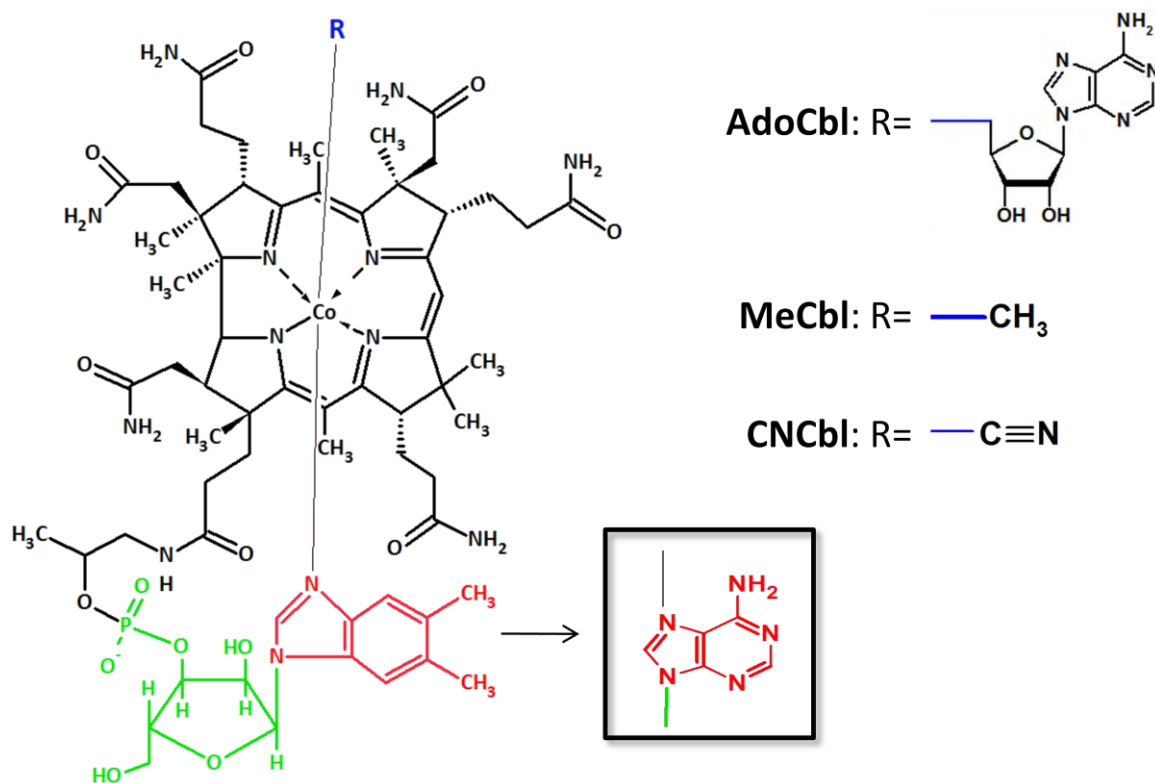


Figure 1-6 Chemical structure of relevant vitamin B12 molecules. The structures of three common vitamin B12 molecules, CNCbl, AdoCbl and MeCbl are identical other than cobalt's upper axial ligand (R). Cobalt's lower axial ligand is typically DMB (red in the main structure). *Pseudo*-B12 molecules use adenine as their lower axial ligand rather than DMB (red molecule in black box). The precursor Cbi lacks the green and red portions of the molecule. Its upper and lower axial ligands are both cyano groups in the commercially prepared form of the compound.

Vitamin B12 features a covalent cobalt-carbon bond, one of the few known examples of a natural carbon-metal bond. The enzyme-mediated breaking of this weak

bond, which can occur either homolytically or heterolytically, instigates the reactions for which vitamin B12 is a cofactor (Banerjee and Ragsdale, 2003). Reaction's that employ AdoCbl are often isomerization reactions that involve swapping the positions of a hydrogen atom and a neighbouring chemical moiety between adjacent carbon atoms of a substrate. These reactions often conclude with the elimination of the relocated chemical moiety (Roth et al., 1996). In isomerization reactions, AdoCbl's cobalt—carbon bond is broken homolytically, resulting in the formation of an 5'-deoxyadenosyl radical (Halpern, 1985). The cofactor radical abstracts the substrate hydrogen atom leaving a substrate radical that rearranges to generate a product radical. A hydrogen from the cofactor adenosyl is then supplied to produce the final product and a cofactor radical that must be recycled.

In enteric bacteria, the substantial commitment of genomic space to vitamin B12-related genes has been puzzling to researchers. *S. enterica* strains generally carry ~60 genes dedicated to the synthesis, transport and utilization of vitamin B12 (Roth et al., 1996). Despite this, the exclusion of vitamin B12 does not produce an obvious phenotype under most laboratory conditions. The only known role of MeCbl in these organisms is in the final step of methionine synthesis (Weissbach and Brot, 1991). Given the presence of a redundant enzyme that can complete this step without MeCbl, it seems likely that AdoCbl is the driving force behind the retention of vitamin B12. The primary AdoCbl-requiring processes in these organisms are the breakdown of alcohols such as ethanolamine and propanediol into useful sources of carbon, nitrogen and energy (Roth et al., 1996). The importance of these emergency energy sources within the native

environments encountered by enteric bacteria has been documented. Of special interest is the apparent significance of ethanolamine utilization and propanediol utilization for the virulence of various pathogenic species of bacteria (Bertin et al., 2011; Garsin, 2010; Joseph et al., 2006; Klumpp and Fuchs, 2007; Lawhon et al., 2003; Maadani et al., 2007).

E. coli species have considerably fewer vitamin B12-related genes than their *S. enterica* relatives (Lawrence and Roth, 1995). This difference is primarily due to *E. coli*'s lack of the propanediol utilization and corrin ring biosynthetic genes. *E. coli* does, however, maintain an extensive vitamin B12 uptake system as well as metabolic genes that enable it to convert incomplete corrinoid molecules to active cofactors (discussed below). In addition to a redundant MeCbl-dependent methionine synthase, *E. coli* contains three enzymes for which AdoCbl has been implicated as a cofactor.

Epoxyqueuosine reductase is an enzyme predicted to be involved in the synthesis of the non-essential, hypermodified tRNA base queuosine. Conflicting results have been observed with respect to whether or not this enzyme requires AdoCbl as a cofactor (Frey et al., 1988; Lawrence and Roth, 1996). Very recently, these observations were explained when *E. coli*'s epoxyqueuosine reductase was identified, purified and assayed (Miles et al., 2011). *In vitro*, AdoCbl was observed not to be essential for the activity of this enzyme, but it did stimulate the reaction by ~30%. The mechanistic and physiological relevance of this stimulation is not currently clear. A second *E. coli* enzyme that exploits AdoCbl is methylmalonyl-CoA mutase. This enzyme converts succinyl-CoA to methylmalonyl-CoA in a pathway that serves to produce propionate from succinate (Froese et al., 2009; Haller et al., 2000). The role of this pathway in *E.*

coli metabolism is not understood. Very low levels of expression have been observed for the four gene operon that makes up this pathway and its deletion has no effect on cell growth under any conditions tested to date (Froese et al., 2009; Haller et al., 2000).

The third and most celebrated of *E. coli*'s AdoCbl-utilizing reactions is catalyzed by the enzyme ethanolamine ammonia lyase. AdoCbl is essential for this reaction, which is the first step in the metabolic breakdown of ethanolamine (Scarlett and Turner, 1976). The prevalence of phosphatidylethanolamine in the membranes of all organisms provides a ready source of ethanolamine in the environment. The breakdown of ethanolamine in *E. coli* cells produces ammonia, which can be exploited as a nitrogen source, and acetyl-CoA, which provides a carbon and energy source. A hypothetical explanation for the evolutionary preservation of *E. coli*'s ~20 vitamin B12 related genes is that ethanolamine is an important metabolite in certain harsh environments to which *E. coli* is routinely exposed (Lawrence and Roth, 1995). Within these environments, corrinoid molecules are generally available, resulting in the loss of the corrin ring biosynthetic system. A reasonable extension of this hypothesis is that propanediol is an important metabolite in certain environments that are more relevant for *S. enterica* than for *E. coli* (Lawrence and Roth, 1996). In these environments, corrinoid molecules are not generally available, and thus there is a selective pressure to maintain the ~20 genes required to produce the corrin ring *de novo*. This idea is supported by genomic information, where a strong correlation has been observed between the presence of corrin ring biosynthetic genes and propanediol utilization genes (Lawrence and Roth, 1996).

As mentioned above, *E. coli* lacks the bulk of the genes required to produce vitamin B12 *de novo*. The five known biosynthetic genes that are present provide *E. coli* with the ability to complete AdoCbl synthesis by assembling the upper and lower axial ligands (Lawrence and Roth, 1995). The minimal metabolic precursor required for this synthesis is cobinamide (Cbi), although for elevated levels of synthesis cobalt's lower axial ligand, DMB, must also be provided exogenously. The enigmatic source of DMB for enteric bacteria has been a source of much confusion and debate. A trickle of DMB appears to be available as a consequence of the breakdown of flavenoids (Keck et al., 1998). More significant levels may be bioavailable in certain environments, although the source of this environmental DMB is not clear. In the absence of suitable levels of DMB, enteric bacteria appear to use adenine in its place, resulting in the production and utilization of *pseudo*-B12 (Anderson et al., 2008). A figure outlining the synthesis of AdoCbl from cobinamide and DMB is an integral part of Chapter 3 and is not included in this introduction to avoid duplication. This diagram is found in chapter 3 as Figure 3-3A.

With an incomplete set of metabolic genes, it is clear that *E. coli* depends on the uptake of environmental vitamin B12 to support its associated cellular functions. This presents a challenge, since vitamin B12 is a very scarce molecule that is produced and used at low levels. Accordingly, *E. coli* possesses an elaborate and highly effective vitamin B12 uptake system. Because it greatly exceeds the size limits to diffuse through standard outer membrane porins, vitamin B12 must be actively transported across the outer membrane. This process requires a high affinity receptor, BtuB, as well as the TonB/ExbB/ExbD complex, which exploits the proton motive force to energize a range

of parallel outer membrane transporters (Braun, 1995). Each TonB outer membrane receptor that has been characterized to date follows a common structural, and probable mechanistic, scheme (Noinaj et al., 2010). Each receptor consists of a large barrel-shaped structure through which the ligand passes, a periplasmic TonB interacting domain and a “plug” domain that blocks passage through the barrel until the ligand is recognized. The most famous of these outer membrane transporters import iron siderophores, secreted molecules that chelate iron with extremely high affinity (Miethke and Marahiel, 2007). The extreme scarcity of bioavailable iron and its established role in virulence makes the secretion and uptake of iron siderophores an area of significant research interest. In addition to providing passage to large metabolites across the outer membrane, TonB-mediated transport also allows for the concentration of rare and precious molecules in the periplasm. Subsequent active transport into the cytoplasm provides a second opportunity to further concentrate these metabolites.

Vitamin B12 transport across the cytoplasmic membrane is accomplished by an ATP-binding cassette (ABC) transporter. ABC transporters are amongst the most widely distributed and heavily studied protein families (Holland and Blight, 1999). In bacteria, ABC transporters are commonly used for both uptake and efflux of metabolites across the cytoplasmic membrane (Davidson et al., 2008). Uptake systems in Gram negative bacteria, such as the one employed by *E. coli* to import vitamin B12, have the following composition: two domains or proteins dedicated to ATP hydrolysis, two integral membrane domains or proteins that form a channel through which the substrate can pass and a transiently-associated periplasmic protein that binds and delivers the substrate to

the transporter. For the *E. coli* vitamin B12 ABC transporter, two molecules of BtuD act as the cytoplasmic ATPase, two molecules of BtuC comprise the transmembrane channel forming component and BtuF acts as the periplasmic binding protein (Locher and Borths, 2004). Upon association with substrate-loaded BtuF, ATP hydrolysis by BtuD is thought to drive conformational changes in BtuC that result in the directional transport of vitamin B12. As with metabolism, a diagram of vitamin B12 transport is shown in Chapter 3 (Figure 3-6).

1.6 Specific research objectives & thesis organization

The central hypothesis of this thesis is that riboswitches can be used as the basis of intracellular, small molecule sensors in bacteria. Furthermore, this biosensing paradigm provides a simple, reliable, flexible, inexpensive and high-throughput compatible method for metabolite detection, thus providing a novel and valuable technology. Finally, it is proposed that riboswitch sensors can be applied to help answer specific and relevant research questions that would otherwise be difficult to address. Three related research projects are presented in chapters 2-4 that aim to address this hypothesis and that work toward developing and establishing riboswitch-based sensing technology. All of the work presented uses the model Gram negative bacterium *E. coli*, although the results obtained are likely very significant for the design and application of similar biosensors in other bacterial species.

The first project presented aims to establish a general and efficient strategy to create synthetic riboswitches. While this project is not as strongly connected to the central hypothesis as the other two, the ability to create and modify riboswitches to respond to a selected metabolite in a desired manner is extremely valuable for riboswitch-based biosensing. The amenability of riboswitches sensors to be produced or modified with customized properties is one of their prime advantages over protein-based methods (see section 1.1). While several strategies have now been established to screen libraries for active riboswitch sequences (discussed in section 1.4), no such technologies had been described when this project was instigated. The specific goal of this project was to develop methodology to screen large, partially-randomized libraries that are rationally designed to boast riboswitch activity. Such a library was constructed and positioned to regulate the expression of a co-transcribed GFP open reading frame (ORF). Two steps of FACS followed by colony screening was then employed in each round of selection to isolate active riboswitch sequences. This work is described in chapter 2 of this thesis.

The second project undertaken, described in chapter 3, tests the heart of the hypothesis of this thesis. The biosensing capabilities of vectors featuring an *E. coli* AdoCbl-responsive riboswitch that is used to regulate various reporter proteins are carefully examined. Importantly, the applicability of these sensors to be used to study vitamin B12-related biology is also tested. Specifically, a number of questions relating to AdoCbl metabolism and transport are addressed using the riboswitch sensors to report on cellular AdoCbl concentrations.

Three different reporter proteins were tested in the second project: β -galactosidase, firefly luciferase and the DsRed-Express red fluorescent protein (RFP). Each of these reporters has different advantages depending upon the intended application. β -galactosidase, described in section 1.4, is a very extensively used and tested reporter protein that is traditionally used in conjunction with colorimetric substrates. This reporter can be used in conjunction with qualitative blue-white screening and quantitative β -galactosidase assays are straightforward and employ a substrate that is relatively inexpensive. Firefly luciferase is one of most widely used luminescent reporter proteins. It catalyzes the oxidation of its substrate luciferin to form a highly unstable, electron excited compound oxyluciferin, which emits light as it returns to ground state (Wood et al., 1989). Firefly luciferase is an extremely sensitive reporter method, however the assay is relatively expensive and the short-lived nature of the signal produced following substrate addition can cause technical challenges. DsRed-Express RFP is a member of a large and growing list of engineered fluorescent proteins that have been derived from naturally-evolved fluorescent proteins (Shaner et al., 2005; Zhang et al., 2002). DsRed-Express was selected because it had previously been employed for experiments in *E. coli* cells and because it produces a bright signal with emission/excitation spectra that are in a range with low levels of autofluorescence (Bevis and Glick, 2002). Assays to quantify fluorescent protein expression do not require substrate addition or cell lysis, making them inexpensive and very amenable to high-throughput approaches. Background fluorescence from cells and the non-enzymatic nature of fluorescent reporters can lead to decreased sensitivity, however. Fluorescent proteins have been engineered to produce

variants with features such as altered spectra, improved photostability, increased brightness, faster maturation times, shorter lifetimes and altered oligomerization states (Giepmans et al., 2006; Shaner et al., 2005). There is therefore a great deal of flexibility in the properties of available fluorescent reporter proteins.

The final project, described in chapter 4, aims to thoroughly address the applicability of riboswitch sensors in investigating relevant research problems. Specifically, a detailed study of the intermolecular interactions involving the periplasmic vitamin B12 transport protein BtuF is conducted using the AdoCbl riboswitch sensors as the primary experimental methodology. The focus of this chapter is therefore shifted away somewhat from the topic of riboswitch sensors and instead the main emphasis is on uncovering new and interesting aspects of the mechanisms underlying vitamin B12 transport. This chapter showcases the ability of riboswitch sensors to accurately detect graded responses, as well as their capacity to be used in higher throughput experiments.

The first two research chapters (2 &3) have been previously published in scientific journals (Fowler et al., 2008, 2010). Chapter 4 is being prepared for submission to a journal and will be submitted prior to the defence of this thesis. These chapters have been re-formatted according to guidelines, but the content is essentially the same as in the original publication. This thesis is therefore presented in the style of a “sandwich thesis”, such that the present introduction and a concluding discussion section (Chapter 5) frame published research chapters. The first subsection of each research chapter is an “Author’s preface” that places the research into the context of the thesis,

describes my contributions to the work and gives a full citation of the original research article. In general, I took a leading role on the conception and design of the projects described as well as on the interpretation of the results and the writing of the manuscripts. I carried out all of the experimental work that is presented. In accordance with sandwich thesis guidelines, the references cited in the research chapters are maintained in their original format and are included as the final section of the chapter. References for the introduction and discussion chapters (chapters 1 and 5) are kept separate and appear at the end of this document.

Chapter 2

A FACS-Based Approach to Engineering Artificial Riboswitches

2.1 AUTHOR'S PREFACE

The first research project described in this thesis aims to develop a general strategy that can be used to produce custom synthetic riboswitches or to modify existing riboswitches. This work was initiated in 2005, less than three years after the first naturally evolved riboswitches were uncovered by the Breaker group. Following these publications, it quickly became clear to many synthetic biologists that the riboswitch paradigm had a great deal of potential for the engineering of custom made regulatory elements. At the time of this project's conception, a few synthetic riboswitches had been produced that were rationally designed or produced using trial and error. From years of research into the construction of functional nucleic acid molecules *in vitro*, it was evident that large scale screening and selection strategies hold many advantages over rational design. The goal of this paper was to develop such a strategy for the selection of active riboswitches within living bacterial cells.

The ability to create or modify riboswitches is tremendously valuable for riboswitch-based sensing. Methods that facilitate such engineering projects are essential to the development of riboswitch sensors that detect molecules for which naturally evolved riboswitches have not been described. Furthermore, these methods could be

adopted to alter the responsive range of a given riboswitch or to optimize a riboswitch to better function within a foreign organism. The FACS-based methodology illustrated in this chapter represents a proficient technique for any of these potential applications.

The following chapter is a modified version of a previously published research article that describes this project in detail. I am the first author listed for this publication and I took a leading role for the conception of the project, designing experiments, interpreting results and in writing the manuscript. Significant assistance and advice in these regards was provided by Dr. Yingfu Li and Dr. Eric Brown, both of whom are also listed as authors of this paper. Other than the direct manipulation of the flow cytometer, I conducted all of the experimental work described. The citation for the original publication follows:

Fowler, C.C., Brown, E.D. and Li, Y. (2008) A FACS-Based Approach to Engineering Artificial Riboswitches. *Chembiochem*, 9, 1906-1911.

2.2 ABSTRACT

Riboswitches are regulatory elements composed entirely of RNA that control the expression of associated genes by directly binding a ligand of interest. Here, we report the development of a method to isolate synthetic riboswitches from a pool of candidates within living *E. coli* cells. Riboswitch candidates were cloned into the 5' untranslated region of GFP and then sorted using two steps of fluorescence activated cell sorting (FACS). In this proof-of-principle endeavour we isolated a novel riboswitch that is significantly and specifically induced by the small molecule theophylline. The method demonstrated here is very general and could be used in conjunction with a wide range of library designs and inducer molecules as well as to modify natural riboswitches. Ultimately, custom-made riboswitches have a great deal of potential for biosensing as well as in the construction of programmable bacteria.

2.3 INTRODUCTION

Nucleic acids have established themselves as valuable building blocks for the construction of synthetic elements with a wide range of uses in biological research.^[1] Indeed, nucleic acid enzymes (ribozymes and deoxyribozymes) and receptors (aptamers) have emerged as major players in the era of synthetic biology, and can be applied in such disciplines as biosensing,^[2, 3] affinity purification,^[4, 5] nanodevices^[6, 7] and therapeutic development.^[8, 9] The success in isolating these elements can be largely

attributed to the methods used to select for molecules that carry out a specific function from extremely large, random sequence libraries.^[10,11] While a great deal of effort and progress has been made toward advancing these protocols *in vitro*, synonymous methods for selections within a cellular environment are far less developed. This shortcoming is becoming more relevant as our understanding of the diverse array of functions that nucleic acids—particularly RNA—have within cells continues to grow at an astounding pace.^[12, 13] The ability to produce these interesting new RNA molecules with custom designed characteristics would be invaluable to biological research.

One of the most intriguing classes of newly discovered RNAs is a group of genetic control elements called riboswitches.^[14, 15] They reside within the untranslated regions (UTRs) of mRNA where they sense the metabolic status of the cell through directly binding a relevant ligand.^[14–22] The binding event results in structural changes in the RNA, which ultimately leads to modulation of the level of expression of neighbouring gene or genes. The bulk of natural riboswitches that have been uncovered have been found in bacteria where they regulate the expression of a remarkable number of genes in response to a growing number of different metabolic cues.^[23] From a molecular engineering point of view, riboswitches represent a highly desirable design for artificial gene expression systems. They have a very simple composition that consists of an RNA receptor domain (aptamer domain) and a neighbouring expression platform with a simple stem–loop architecture. Riboswitches are relatively small, *cis*-acting elements, usually encoded by 300 bases or less. These factors simplify the design, optimisation and troubleshooting associated with their creation. Riboswitches have shown the potential to

be extremely efficient as well, and demonstrate both remarkable specificity and impressive induction levels in response to their cognate ligand.^[21, 22]

Artificial riboswitches could be designed for a wide range of applications. One such possibility is the engineering of programmable cells. This could include the production of cells that have an array of genes under the control of different inducer molecules; this would open the door to experiments for working out the interplay of numerous genes or pathways. Programmed bacteria also have great potential as tools for environmental cleanup and drug delivery. Topp and Gallivan recently made progress toward this goal by employing an artificial riboswitch to create a strain of *E. coli* cells that migrate specifically toward the small molecule theophylline.^[24] We believe that riboswitches also hold great potential as intracellular sensors for the detection and quantification of small molecules. While these fields are still largely unexploited, the continuing development of methods to create and modify riboswitches should make such systems possible.^[25, 26]

The goal of this work was to develop a fast, efficient and general intracellular selection method for isolating riboswitches from a library of sequences. In order to sort through a large number of sequences quickly and easily, we opted to use fluorescent proteins and fluorescence activated cell sorting (FACS). FACS can be used to sort individual cells based on their fluorescence intensity at rates that exceed 10^4 cells per second. Previously, FACS has been employed to identify proteins with desirable properties from libraries of random mutants.^[27–30] Here, we apply similar thinking to

2.4 MATERIALS AND METHODS

2.4.1 Growth media, strains, plasmids and molecular cloning

All cells were grown in Luria–Bertani (LB) broth supplemented with ampicillin (50 mg/mL). Library transformations were done by using *E. coli* DH5 α max efficiency competent cells (Invitrogen); this strain was also used for FACS and 96-well fluorescence assays. All other experiments were carried out by using Nova blues (Novagen) *E. coli* cells. The pRB374 plasmid was used as the backbone for all constructs described. Sequence and detailed cloning information are provided in the appendix, chapter 2 supplemental materials, Figure S2-1 and materials and methods.

All molecular cloning, including restriction digests, ligations, transformations, DNA phosphorylation and second-strand synthesis were carried out according to well established, standard protocols by following supplier instructions. Cloning of transcriptional terminators and library sequences was done by using chemically synthesized DNA that encoded the desired sequence (Mobix lab, McMaster University; Integrated DNA Technologies, Coralville, IA, USA). In some cases, two complementary DNA oligonucleotides with suitable overhangs for cloning into the appropriate restriction sites were annealed, phosphorylated by using T4 polynucleotide kinase (PNK) and ligated into the digested plasmid. For other constructs, DNA oligonucleotides (with the desired sequence flanked on either side by suitable restriction sites) were made into a double-stranded form by using a short primer and Klenow DNA polymerase, followed by

restriction digestion and ligations. The sequences of the 6E12 and RS11A linker regions were 5'-ACTATATGAAGAGG-3' and 5'-ATTGAAGGAAGAGG-3', respectively.

2.4.2 Fluorescence activated cell sorting and colony isolation

Prior to cell sorting, *E. coli* cells were washed several times with PBS and resuspended in PBS at a concentration of approximately 10^7 cells/ml. FACS was carried out by using a FACSVantage SE™ flow cytometry system (BD Biosciences). A 488 nm laser was used for excitation, and fluorescence was detected by using a 530 ± 15 nm band pass filter. Sorting rates that ranged from 800–12,000 events per second were used, depending on the requirements of the particular experiment.

For the first round of selection, transformed cells were grown at 37°C for ~6 h in media supplemented with theophylline (2 mM; Sigma) prior to being sorted. Positive cells were collected and stored (diluted in several volumes of growth media) at 4°C for ~6 h. They were then grown at 37°C in media that lacked theophylline for ~10 h prior to the second phase of FACS. The cells were diluted and plated on agar to yield ~100 cells per plate. Individual colonies were picked manually, grown, overnight, at 37°C and stored as frozen stocks in LB supplemented with glycerol (15 %). The second round of sorting was carried out in the same manner as described above, except that cells were grown at 37°C for only 5 h (rather than 10 h) prior to the second phase of FACS. This was done to minimize selective growth advantages and other anomalies associated with long periods of bacterial cell growth.

2.4.3 Fluorescence assays

High-throughput cell assays were carried out in 96-well, clear polypropylene plates (Corning). Duplicate plates were prepared by pinning, and allowed to grow for approximately 24 h at 37°C, one plate with theophylline and an identical plate without. Fluorescence for both sets of plates was measured by using a Tecan Safire fluorometer spectrophotometer with excitation/emission set at 488/517 nm. The values were normalized by subsequently measuring absorbance at 600 nm (OD₆₀₀) for both sets of plates. Colonies were chosen for examination based on the ratio of the corrected fluorescence values in the presence and absence of theophylline.

Low-throughput assays were carried out by picking individual colonies from plates of cells that harboured the plasmid of interest. These colonies were grown, overnight, at 37°C, and diluted 1:1000 in fresh media supplemented with the indicated concentration of inducer. After 7 h of growth at 37°C the OD₆₀₀ was measured for each sample and approximately 3 ml of culture was centrifuged (the centrifuged volume was varied to correct for differences in OD₆₀₀) and the pellet was resuspended in PBS. The resuspension process led to the cultures being concentrated ~eightfold (from ~3 ml to 400 ml). This was done to optimize the signal strength to provide the most accurate possible data. A fraction of the resuspended cells (100 ml) was added to half-area, solid black 96-well plates (Corning) and measured for fluorescence as described above. Unless otherwise stated, the concentration of theophylline used for all experiments was 2 mM.

2.5 RESULTS

2.5.1 Selection scheme and library design

As a pilot project to demonstrate our selection scheme we chose to seek riboswitches that respond to the small molecule theophylline. We built a library of “potential riboswitches” that consisted of an RNA aptamer that binds theophylline connected to an intrinsic transcriptional terminator by a partially random linker region (Figure 2-1B). Transcriptional terminators are RNA elements that consist of a simple stem–loop followed by a string of uracil residues.^[31, 32] When RNA polymerase encounters such an element it is released from the template and sequences downstream are not transcribed. We sought linker regions that were able to transduce the binding of theophylline to a disruption of the neighbouring transcriptional terminator stem; this would result in activating transcription of the downstream green fluorescent protein (GFP) gene. This riboswitch mechanism (transcriptional termination) appears to be the most common, and has not been exploited by artificial riboswitches that have been created to date.

Before beginning the selection process we first chose each of the elements for the construction of our riboswitch library. The TCT8-4 theophylline aptamer,^[33] which has previously been used for a wide range of proof-of-principle experiments, was chosen for a variety of reasons: it is very well characterized, it exhibits remarkable specificity for its ligand and it is functional in a cellular environment.^[34–37] To choose a transcriptional terminator we cloned a panel of candidates upstream of GFP within our selection plasmid

and carried out fluorescence assays. The terminator from the *B. subtilis* MetI riboswitch^[22] was chosen as it appeared to be the most efficient within our system (data not shown). It should be noted that we only assayed terminators from riboswitches as it has been postulated that they contain pause signals that allow for efficient ligand binding and riboswitch function.^[37, 38] The final element of our library—the linker region—was biased to resemble the antiterminator from the MetI riboswitch to provide our library with an enhanced ability to disrupt the terminator (Figure 2-1B, see legend).

2.5.2 FACS to isolate candidate riboswitch sequences

Since the goal of this project was to develop a selection method that could isolate rare active sequences from a library, we first cloned and sequenced eight random, unique library sequences to determine if they were theophylline responsive. These plasmids were assayed in cell-based fluorescence assays (see Section 2.4) and none showed a significant theophylline response. We thus set out to test our selection method by cloning a large number of sequences into competent *E. coli* cells. By plating dilutions of this transformation it was determined that approximately 100,000 positive clones were successfully transformed. Cells were grown in the presence of theophylline for approximately 6 h and sorted to keep only those with a relatively high fluorescence (Figure 2-2A). Choosing an induction time required striking a balance between allowing the induced GFP time to fold and minimizing growth time to avoid losing library diversity. We were particularly concerned about this because cells that express lower levels of GFP would presumably have a slight growth advantage.

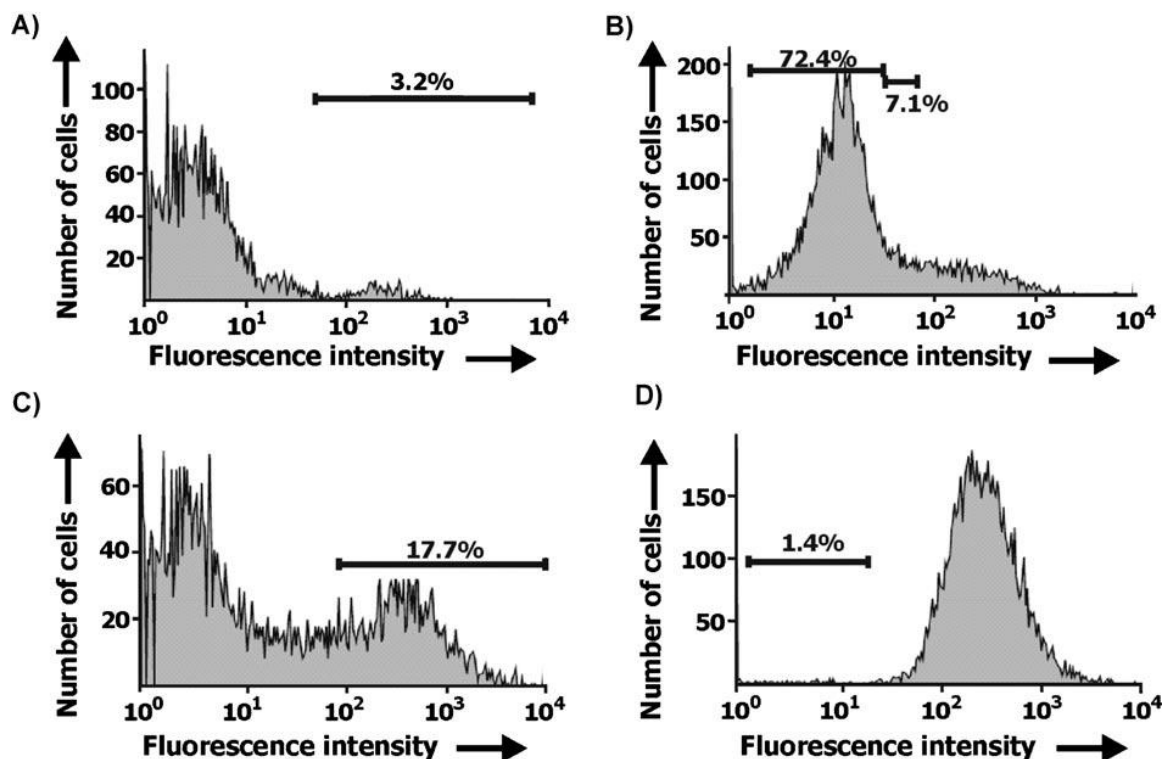


Figure 2-2 FACS results. (A) selection round 1, with theophylline; (B) selection round 1, without theophylline; (C) selection round 2, with theophylline; (D) selection round 2, without theophylline.

The remaining cells, which had survived the first sort, were grown in the absence of theophylline, and those with a relatively low fluorescence were kept (Figure 2-2B). For our first round of selection we opted to use rather lenient gating in order to avoid losing cells that contained weakly active riboswitches. The astonishingly high proportion of cells with a low fluorescence in the second FACS sort indicated that our pool either contained a very large number of active sequences, or that the first round of FACS was

too lenient and thus a large number of weakly fluorescent cells were able to pass through the first sort.

To examine the cells that had survived the selection, individual clones were isolated by plating on agar plates. Single colonies were then grown in separate wells of 96-well plates and kept as frozen stocks. From these stocks duplicate plates were grown, one with theophylline and one without, and were assayed for fluorescence. The majority of the clones exhibited small changes in fluorescence in response to theophylline; this indicates that perhaps our FACS gating was too lenient. Several of the clones showed a very clear theophylline response and they were selected for further examination.

The clones examined showed a two- to fourfold increase in GFP expression in response to theophylline. Generally, there were no obvious trends observed that related the sequence of these constructs to their activity. We wanted to see if a second round of this selection process would yield a construct with a more significant theophylline response. The plasmid 6E12 exhibited the highest fold induction from round one and so it was chosen to be the basis of a new library for a second round of selection. At each of the 14 positions in the linker region, the new library had a 76% chance of retaining the 6E12 sequence, and an 8% chance of being mutated to each of the other three possible nucleotides.

The second selection was carried out in a similar manner to the first. The FACS spectrum for the first sort looked similar to that observed in the first round (Figure 2-2C), however a greater fraction of clones exhibited a high level of fluorescence. The spectrum

for the second sort was much more consistent with what we predicted than in the first round (Figure 2-2D), and only a small percentage of cells passed our more stringent criteria. Follow-up 96-well plate fluorescence assays yielded a much higher proportion of theophylline-responsive clones than the first round; this indicates that the more stringent FACS criteria were beneficial. Also, several riboswitches with fold-induction values that exceeded the 6E12 riboswitch were uncovered; the most efficient of these was RS11A. The sequence of RS11A in the original random portion (blue nucleotides in Figure 2-1B) was ACTATATGAAGAGG and the remainder of the sequence (including the aptamer region, Figure 2-1B) did not incur any unexpected mutations.

2.5.3 Assaying candidate riboswitches for theophylline response

Fluorescence assays both in the presence and absence of theophylline for 6E12, RS11A and RS11A with a point mutant that disrupted theophylline binding (C27A) are shown in Figure 2-3A. These data show that RS11A was nearly twice as efficient as its predecessor 6E12, and it owed its improved induction to a lower level of uninduced expression. The C27A mutant of RS11A showed no theophylline response; this indicates that aptamer binding was indeed responsible for the induction of GFP expression and implies a riboswitch-based mechanism of regulation. Figure 2-3B demonstrates the selectivity of RS11A for theophylline over caffeine—a close structural relative. It also demonstrates that the riboswitch was responsive to theophylline over a wide concentration range, up to 2 mM. Concentrations greater than 2 mM were not tested as they were detrimental to cell growth. It should be noted that previous studies have shown

that the intracellular concentration of theophylline in cultured *E. coli* is far lower than the extracellular concentration, presumably due to efflux.^[39] This explains the high media concentration of inducer required to elicit a maximal response. The same phenomenon has been previously observed by other groups.^[35, 36]

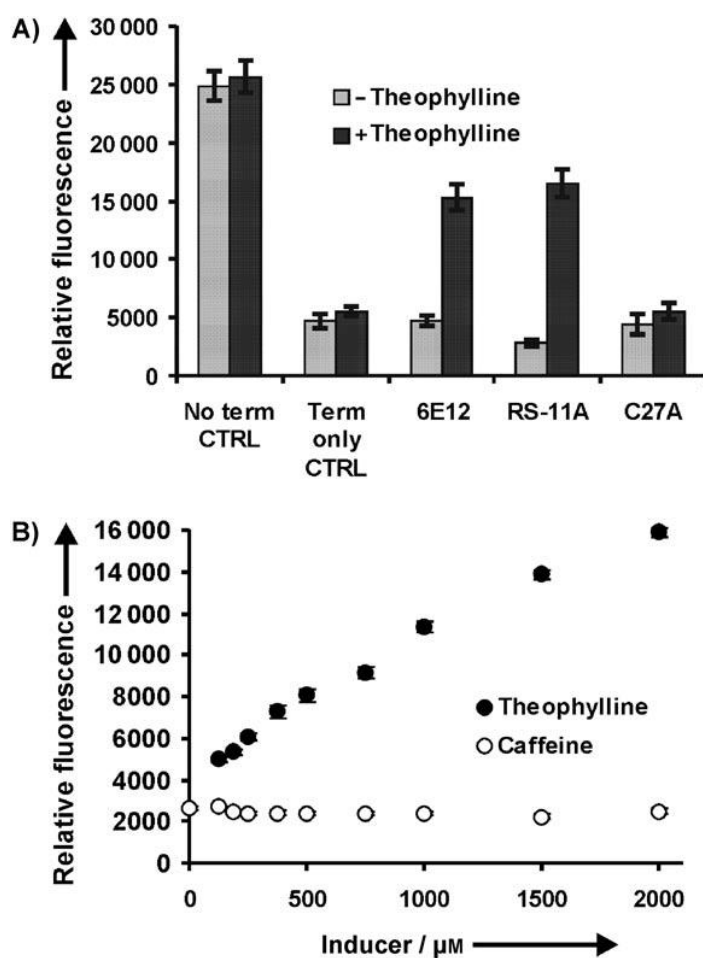


Figure 2-3 Performance of some selected riboswitches. (A) Fluorescence assays of the top riboswitch candidates isolated from each round of selection. No term CTRL: construct that lacked the entire aptamer–linker–terminator region; Term only CTRL: construct with the MetI transcriptional terminator, but no aptamer and linker region. (B) Fluorescence of RS11A in response to changes in the concentrations of theophylline and caffeine.

2.5.4 Mechanistic characterization of RS11A

We next sought to examine whether or not RS11A was acting at the level of transcription, as was intended by our library design. As a first step in this process we identified a potential antiterminator stem in RS11A that could compete with the formation of the terminator (Figure 2-4A). It is interesting to note that the putative antiterminator stem of RS11A was shifted slightly from the wild-type MetI antiterminator, upon which the original library was based. To determine whether RS11A used this terminator/antiterminator mechanism, we mutated the polyU tract that follows the terminator (Figure 2-4B). While these residues should be crucial for terminator function, they should not have a significant effect on RNA folding. As a control we mutated the polyU tract of the MetI terminator in a construct that lacked the aptamer and the linker region of RS11A. This mutation significantly disrupted terminator function. To our surprise, however, RS11A was completely unaffected by the polyU mutation. This implies that the antiterminator efficiently prevented the formation of the terminator both in the presence and absence of theophylline, and that RS11A does not exert its effect through transcriptional termination (Figure 2-4C).

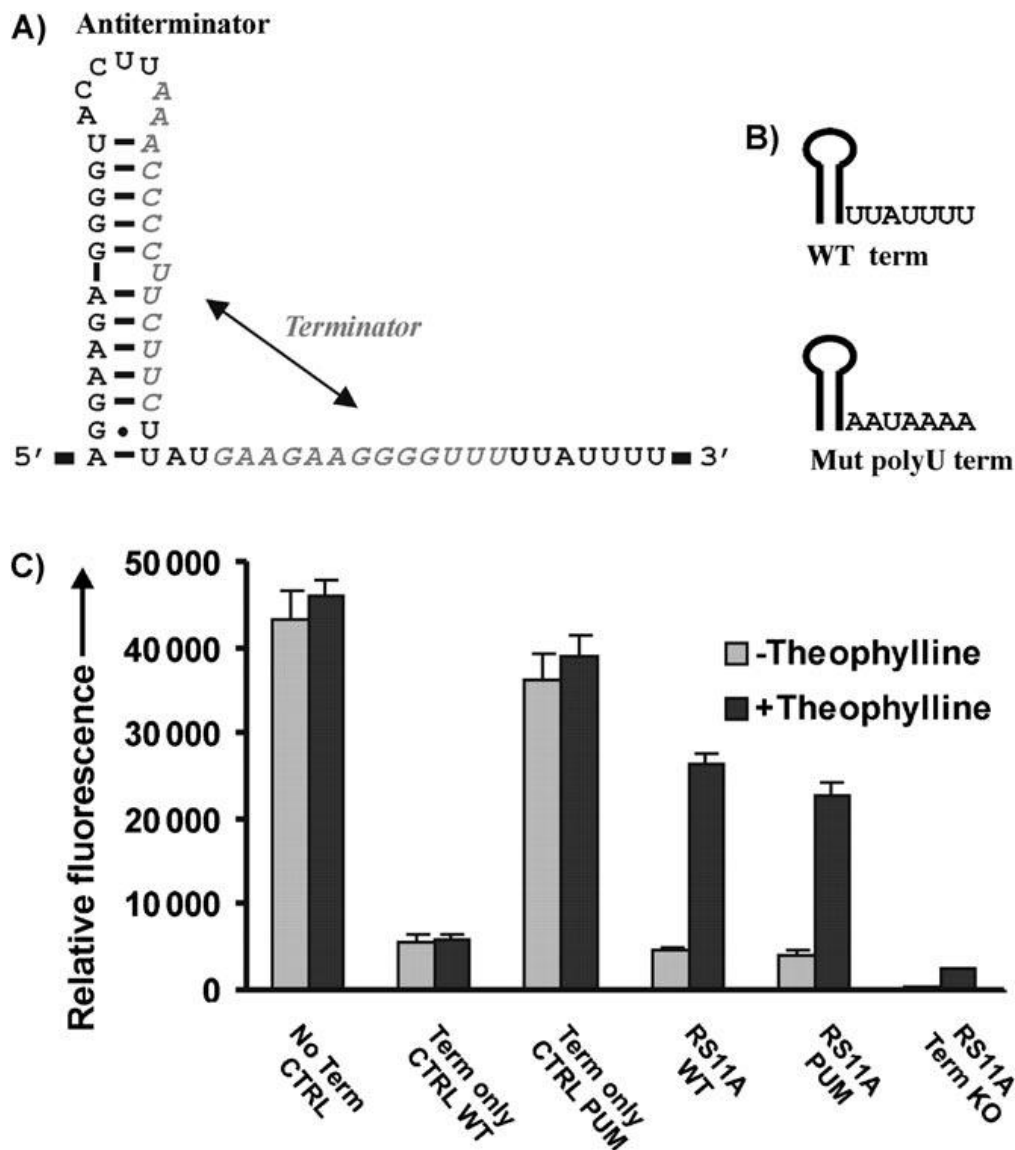


Figure 2-4 Probing the regulation mechanism of the RS11A riboswitch. (A) Predicted antiterminator stem of RS11A. Residues shown in grey and italic letters represent the stem of the MetI terminator. (B) Schematic representation of the wild-type and polyU mutant transcriptional terminators used to examine the RS11A mechanism. (C) Fluorescence assay that compares various riboswitch constructs and controls and their response to theophylline; the controls are described in the legend of Figure 2-3. WT: construct with a wild-type MetI terminator; PUM: construct that had a MetI terminator with a mutated polyU tract; RS11A Term KO: an RS11A construct that completely lacked the MetI transcriptional terminator.

To further examine the role of the terminator, we made a construct that completely removed the terminator, but maintained the selected RS11A linker sequence. It was difficult to anticipate the effects of such a deletion, as it involved the removal a large chunk of sequence from the middle of the 5'-UTR. Interestingly, this construct retained its switching ability (sevenfold induction), but the level of expression (both induced and uninduced) fell dramatically. This implies that the removed sequence might not be involved in riboswitch activity, but that the distance of the riboswitch from translational signals is important for expression levels. The next obvious possibility for the RS11A mechanism was ribosome binding site sequestration/antisequestration^[14]—the other predominant mechanism exploited by naturally occurring riboswitches. By examining the sequence manually and by using structure-prediction programs, a few possibilities for such structures were noted, but subsequent experiments dismissed these options.

To further probe the mechanism we performed in-line probing on the RS11A 5'-UTR, a method that probes for structural changes in RNA in response to ligand binding. Although some changes were noted in the aptamer and linker regions, the data did not provide a clear indication of mechanism (see appendix, chapter 2 supplemental materials, Figure S2-2 and additional information that follows). Several groups have previously shown ligand-responsive expression in response to cloning an aptamer immediately upstream of a gene.^[35, 40–42] While the precise mechanism has not always been clear, it is evident that in the correct context the structural changes caused by aptamer–ligand binding can have a significant effect on neighbouring gene expression. For RS11A, we

have selected for a sequence, and presumably an RNA structure, that places the aptamer in a suitable context to allow for theophylline-activated gene expression. This could perhaps occur through the ribosome's ability to access the different structures or at the level of RNA stability, although this is entirely speculative.

2.6 DISCUSSION

In summary, we have demonstrated a novel means of creating synthetic riboswitches that relies on FACS and fluorescence assays to isolate active riboswitches from a library. We believe that this general scheme offers some distinct advantages in engineering riboswitches in *E. coli*. Firstly, by selecting and screening for the best sequences, there is less reliance on clever engineering. Such designs are not usually generally applicable and can require a structural understanding of the aptamer that is not always available. FACS analysis allows for millions of individual cells to be sorted in less than an hour, which makes this strategy amenable to relatively high throughput selections. FACS also allows the user to choose gate settings that make it possible to fine-tune the stringency for each round of selection, and ultimately to target whatever expression levels they hope to achieve. It should be noted, however, that this method is limited with respect to the size of library that can be sorted. While high-efficiency cloning strategies can allow for much larger libraries to be produced than that used in this study, libraries that approach the size commonly used for *in vitro* experiments ($\sim 10^{15}$) are not obtainable and could not be sorted with this method. While sorting does allow for

much more flexibility than systems that are strictly rationally designed, this method still requires a certain degree of engineering.

This proof-of-principle effort targeted a transcriptional-termination mechanism, although our data indicate that this is not what was obtained. We believe this to be a function of the library we used, as it is entirely possible that such a mechanism was not achievable within the confines of our design. The unexpected nature of what we isolated reinforces the advantages of selection over strict engineering for elements such as this and shows the power of the FACS method employed here. Application of this method to sort libraries that more efficiently target a given mechanism would presumably yield more efficient riboswitches. This could be achieved by more thoroughly designing a transcriptional-termination mechanism that includes carefully designed aptamer and terminator stems by using random elements to fine-tune the system. Alternatively, translational initiation is likely much simpler to target as only a few bases of sequence (the ribosome binding site) need to be structurally affected by ligand binding. Recently published work that uses such a library design demonstrated impressive induction levels and further supports this idea.^[25] One advantage of riboswitches that work by transcriptional termination, however, is that multiple copies of these elements could potentially be used to obtain very tightly regulated systems. The concept of multiple riboswitches that control the expression of a single gene has been exploited in natural systems.^[21, 43]

In addition to de novo creation of riboswitches, we feel that this method has the potential to be used for modifying those that have evolved naturally. Evolution has provided these elements with exquisite ligand-recognition characteristics and gene-regulation capabilities. This makes natural riboswitches a great starting point for the development of tools for synthetic biology. However, in many cases it is desirable to modify these riboswitches to exhibit different ligand specificity, to be responsive over a different ligand concentration range or to exhibit more or less tightly regulated expression. The speed and high-throughput capacity of the method demonstrated here make it an ideal choice for such endeavours.

2.7 REFERENCES

- [1] D. H. Bunka, P. G. Stockley, *Nat. Rev. Microbiol.* 2006, 4, 588–596.
- [2] J. Liu, Y. Lu, *Methods Mol. Biol.* 2006, 335, 275–288.
- [3] S. Hahn, S. Mergenthaler, B. Zimmermann, W. Holzgreve, *Bioelectrochemistry* 2005, 67, 151–154.
- [4] D. Hutanu, V. T. Remcho, *Adv. Chromatogr.* 2007, 45, 173–196.
- [5] S. Tombelli, M. Minunni, M. Mascini, *Biosens. Bioelectron.* 2005, 20, 2424–2434.
- [6] U. Feldkamp, C. M. Niemeyer, *Angew. Chem.* 2006, 118, 1888–1910;
Angew. Chem. Int. Ed. 2006, 45, 1856–1876.
- [7] P. Guo, J. Nanosci. *Nanotechnol.* 2005, 5, 1964–1982.
- [8] J. F. Lee, G. M. Stovall, A. D. Ellington, *Curr. Opin. Chem. Biol.* 2006, 10, 282–289.
- [9] C. R. Ireson, L. R. Kelland, *Mol. Cancer. Ther.* 2006, 5, 2957–2962.
- [10] A. D. Ellington, J. W. Szostak, *Nature* 1990, 346, 818–822.
- [11] C. Tuerk, L. Gold, *Science* 1990, 249, 505–510.
- [12] A. Serganov, D. J. Patel, *Nat. Rev. Genet.* 2007, 8, 776–790.
- [13] J. S. Mattick, I. V. Makunin, *Hum. Mol. Genet.* 2006, 15 Spec. No. 1, R17–29.
- [14] W. Winkler, A. Nahvi, R. R. Breaker, *Nature* 2002, 419, 952–956.
- [15] A. Nahvi, N. Sudarsan, M. S. Ebert, X. Zou, K. L. Brown, R. R. Breaker, *Chem. Biol.* 2002, 9, 1043.
- [16] N. Sudarsan, J. K. Wickiser, S. Nakamura, M. S. Ebert, R. R. Breaker, *Genes Dev.* 2003, 17, 2688–2697.
- [17] W. C. Winkler, S. Cohen-Chalamish, R. R. Breaker, *Proc. Natl. Acad. Sci. USA* 2002, 99, 15908–15913.
- [18] A. Roth, W. C. Winkler, E. E. Regulski, B. W. Lee, J. Lim, I. Jona, J. E. Barrick, A. Ritwik, J. N. Kim, R. Welz, D. Iwata-Reuyl, R. R. Breaker, *Nat. Struct. Mol. Biol.* 2007, 14, 308–317.
- [19] C. E. Dann III, C. A. Wakeman, C. L. Sieling, S. C. Baker, I. Irnov, W. C. Winkler, *Cell* 2007, 130, 878–892.
- [20] M. Mandal, R. R. Breaker, *Nat. Struct. Mol. Biol.* 2004, 11, 29–35.
- [21] M. Mandal, M. Lee, J. E. Barrick, Z. Weinberg, G. M. Emilsson, W. L. Ruzzo, R. R. Breaker, *Science* 2004, 306, 275–279.
- [22] W. C. Winkler, A. Nahvi, N. Sudarsan, J. E. Barrick, R. R. Breaker, *Nat.*

- Struct. Biol.* 2003, 10, 701–707.
- [23] B. J. Tucker, R. R. Breaker, *Curr. Opin. Struct. Biol.* 2005, 15, 342–348.
- [24] S. Topp, J. P. Gallivan, *J. Am. Chem. Soc.* 2007, 129, 6807–6811.
- [25] S. A. Lynch, S. K. Desai, H. K. Sajja, J. P. Gallivan, *Chem. Biol.* 2007, 14, 173–184.
- [26] Y. Nomura, Y. Yokobayashi, *J. Am. Chem. Soc.* 2007, 129, 13814–13815.
- [27] S. van den Berg, P. A. Löfdahl, T. Härd, H. Berglund, *J. Biotechnol.* 2006, 121, 291–298.
- [28] B. P. Cormack, R. H. Valdivia, S. Falkow, *Gene* 1996, 173, 33–38.
- [29] A. Aharoni, K. Thieme, C. P. Chiu, S. Buchini, L. L. Lairson, H. Chen, N. C. Strynadka, W. W. Wakarchuk, S. G. Withers, *Nat. Methods* 2006, 3, 609–614.
- [30] S. W. Santoro, L. Wang, B. Herberich, D. S. King, P. G. Schultz, *Nat. Biotechnol.* 2002, 20, 1044–1048.
- [31] P. Gollnick, P. Babitzke, *Biochim. Biophys. Acta Gene Struct. Expression* 2002, 1577, 240–250.
- [32] E. Nudler, M. E. Gottesman, *Genes Cells* 2002, 7, 755–768.
- [33] R. D. Jenison, S. C. Gill, A. Pardi, B. Polisky, *Science* 1994, 263, 1425–1429.
- [34] G. R. Zimmermann, R. D. Jenison, C. L. Wick, J. P. Simorre, A. Pardi, *Nat. Struct. Biol.* 1997, 4, 644–649.
- [35] S. K. Desai, J. P. Gallivan, *J. Am. Chem. Soc.* 2004, 126, 13247–13254.
- [36] B. Suess, B. Fink, C. Berens, R. Stentz, W. Hillen, *Nucleic Acids Res.* 2004, 32, 1610–1614.
- [37] J. K. Wickiser, M. T. Cheah, R. R. Breaker, D. M. Crothers, *Biochemistry* 2005, 44, 13404–13414.
- [38] J. K. Wickiser, W. C. Winkler, R. R. Breaker, D. M. Crothers, *Mol. Cell* 2005, 18, 49–60.
- [39] A. L. Koch, *J. Biol. Chem.* 1956, 219, 181–188.
- [40] G. Werstuck, M. R. Green, *Science* 1998, 282, 296–298.
- [41] I. Harvey, P. Garneau, J. Pelletier, *RNA* 2002, 8, 452–463.
- [42] B. Suess, S. Hanson, C. Berens, B. Fink, R. Schroeder, W. Hillen, *Nucleic Acids Res.* 2003, 31, 1853–1858.
- [43] N. Sudarsan, M. C. Hammond, K. F. Block, R. Welz, J. E. Barrick, A. Roth, R. R. Breaker, *Science* 2006, 314, 300–304.

Chapter 3

Using a Riboswitch Sensor to Examine Coenzyme B12 Metabolism and Transport

In *E. coli*

3.1 AUTHOR'S PREFACE

The research project presented in this chapter addresses the heart of the hypothesis underlying this thesis. In this project, a series of genetically-encoded riboswitch sensors are constructed and they are assessed for their detection capabilities and their capacity to be used to study biological processes. The sensors constructed detect AdoCbl levels within *E. coli* cells using one of three different reporter systems. In addition to describing the first thoroughly examined riboswitch-based sensor, this study also produces interesting and novel data regarding the metabolism and transport of AdoCbl in *E. coli*.

As with chapter 2, this chapter is a modified version of a previously published research article. I am the first author listed for this publication and I took a leading role for the conception of the project, designing experiments, interpreting results and in writing the manuscript. Dr. Yingfu Li and Dr. Eric Brown, both of whom are also listed as authors of this paper, provided me with important suggestions and guidance and helped edit the final manuscript. I conducted all of the experimental work described. The citation for the original publication follows:

Citation:

Fowler, C.C., Brown, E.D. and Li, Y. (2010) Using a Riboswitch Sensor to Examine Coenzyme B(12) Metabolism and Transport in *E. coli*. *Chem Biol*, 17, 756-765.

3.2 ABSTRACT

Small molecules play crucial roles in every major cellular process. Despite this, detecting their levels within cells remains a significant challenge. Here, we describe intracellular sensors of coenzyme B₁₂ that make use of the exquisite molecular detection capabilities of a naturally occurring riboswitch. These probes sensitively detect their target using colorimetric, fluorescent or luminescent reporters. To assess their utility in the study of biological systems, the sensors were applied to examine the synthesis and the import of coenzyme B₁₂. The sensors were able to monitor the effects of genetic deletions, recombinant expression of foreign genes, and varied growth conditions on both of these processes. These results indicate that riboswitch-based sensors can provide valuable information on intracellular small molecule concentrations that can be employed in the study of related cellular processes.

3.3 INTRODUCTION

Cells continually synthesize, degrade, import and export a vast array of small molecule metabolites. These processes are crucial both to adapting to environmental changes and to carrying out routine cellular functions. Despite the widespread importance of small molecules to cellular physiology, monitoring changes in their intracellular concentrations remains a significant challenge. Traditionally, measuring cellular metabolite levels has been carried out *in vitro*, following cell lysis. While these approaches remain valuable for certain applications, they have many drawbacks as well. Efficient separation and identification of the target metabolite can require multiple steps and specialized equipment and knowledge that are not readily accessible to all researchers. Furthermore, the true intracellular concentration of the target molecule can be skewed by contamination from the growth medium or by degradation, aggregation or adhesion to cellular debris following cell lysis. Intracellular detection of small molecules is an attractive alternative that has seen much progress in recent years (Binkowski et al., 2009; Medintz, 2006) . One successful approach has made use of protein fusions between molecular receptor domains and reporter proteins such as GFP or firefly luciferase. Here, ligand binding to the receptor activates an otherwise inert reporter protein, which then produces a readily detectable signal (Fan et al., 2008; Paulmurugan and Gambhir, 2006) . Alternatively, two fluorescent proteins attached at specific sites of the receptor undergo changes in FRET upon metabolite binding (Fehr et al., 2002; Miyawaki et al., 1997; Romoser et al., 1997). To date, however, this methodology has been limited to detecting ligands with natural, well characterized, high affinity protein receptors.

Like protein, RNA has also been used effectively to detect biological agents (Cho et al., 2009) . Using large oligonucleotide libraries and a selection technique known as SELEX (Ellington and Szostak, 1990; Tuerk and Gold, 1990), dozens of specific, high affinity nucleic acid receptors (dubbed aptamers) have been isolated. Artificially obtained aptamers have been successfully modified to act as *in vitro* sensors of target molecules using a variety of creative strategies (Navani and Li, 2006). Aptamers have also been found naturally, as sensory domains in RNA regulatory elements known as riboswitches (Mironov et al., 2002; Nahvi et al., 2002) . Riboswitches are most commonly found in the 5' untranslated regions of prokaryotic mRNAs, where they regulate the expression of downstream genes through structural changes undergone in response to the binding by a specific target molecule. Typically, ligand binding at the aptamer domain either alters ribosome access to the transcript or affects the formation of intrinsic transcriptional terminators [reviewed in (Barrick and Breaker, 2007)].

Intrinsically, riboswitches act as relatively simple and direct metabolite sensors that are capable both of molecular recognition and of transducing this information by altering protein expression. As depicted in Figure 3-1, we hypothesized that a riboswitch controlling the expression of a reporter gene could be applied as a reliable, relatively non-invasive intracellular metabolite sensor. In the present study, we construct such a sensor to monitor adenosylcobalamin (AdoCbl) concentrations in the model Gram negative bacterium *E. coli*. The ultimate goal for any technology such as this is its application in addressing biologically relevant questions. Accordingly, this study also exploits the

ability of these sensors to monitor target concentrations as a means to investigate the processes of AdoCbl import and metabolism.

AdoCbl, also known as coenzyme B₁₂, is a cofactor required for the catalysis of a number of isomerization reactions (Roth et al., 1996). AdoCbl is not important for growth under standard laboratory conditions for *E. coli* and related species of bacteria. Its primary importance appears to be in the metabolism of compounds that are important sources of carbon, nitrogen and energy in specific environments. B₁₂-dependent enzymes for metabolizing glycerol, propanediol and ethanolamine are prevalent in enteric bacteria (Abeles and Lee, 1961; Bradbeer, 1965; Toraya et al., 1979). *E. coli* lacks these enzymes for both glycerol and propanediol and thus ethanolamine utilization appears to be the paramount B₁₂-requiring process (Lawrence and Roth, 1996). A growing number of studies point to an important role for ethanolamine utilization for microbial *in vivo* growth and pathogenesis (Bourgogne et al., 2006; Huang et al., 2007; Joseph et al., 2006; Klumpp and Fuchs, 2007; Korbel et al., 2005; Lawhon et al., 2003; Maadani et al., 2007), providing an added layer of interest in understanding the metabolic and import processes responsible for providing cells with AdoCbl. The sensors described in this study represent valuable tools for the study of these processes as they allow for sensitive detection of cellular AdoCbl levels using fast and convenient assays.

3.4 METHODS AND MATERIALS

3.4.1 Plasmids and Cell Strains

All plasmids used were derived from the pBAD series of vectors (Guzman et al., 1995) and modified as described below and in Supplemental Experimental Procedures. Molecular cloning was carried out using *E. coli* NovaBlue cells (Novagen). Unless otherwise noted, all experiments shown were conducted using the *E. coli* strain BW25113 (*rrnB*_{T14} *ΔlacZ*_{WJ16} *hsdR514* *ΔaraBAD*_{AH33} *ΔrhaBAD*_{LD78}) (Datsenko and Wanner, 2000). The single gene deletion strains were from the Keio Collection and are derived from the BW25113 parental strain (Baba et al., 2006). The *ΔcobT* and *ΔbtuB* strains used were not taken from this library. The *ΔbtuB* strain from the Keio collection was recently found to be subject to a duplication event such that a copy of *btuB* was retained (Yamamoto et al., 2009), presumably due to overlap with the essential *murI* gene. The *ΔcobT* strain in the copy of this library used was also found to contain a WT copy of the gene, perhaps due to an error replicating the library. Both of these strains were therefore created for the purpose of this study using the same procedure (Datsenko and Wanner, 2000), except that a *cm*^R cassette was used rather than *kan*^R. Deletions were confirmed by PCR. Primers used to create deletion strains are given in the appendix, chapter 3 supplemental materials, Table S3-1.

3.4.2 Molecular cloning

The details of plasmid construction are provided in the appendix, chapter 3 supplemental materials, chapter 3 supplemental materials and methods. Primer sequences are provided in the appendix, chapter 3 supplemental materials, Table S3-1. Briefly, DNA fragments for both the *btuB* riboswitch and for the individual reporter genes were obtained using PCR from a suitable template. Riboswitch-reporter gene fusions were obtained by crossover PCR of the individual fragments. The resulting fusions were cloned into a modified version of the pBAD18 plasmid (lacking all arabinose-responsive elements) using restriction digestion and ligation. The full sequence of each of these sensing plasmids is provided in the appendix, chapter 3 supplemental materials. Synonymous constructs were also made using the pBAD30 plasmid, a lower copy version of pBAD18. The (A₁₄₉A₁₅₀>TT) mutations were made in a similar manner, using crossover PCR with mutations built into the crossover primers. The *bluB* gene from *M. smegmatis* was cloned by PCR-amplification of the gene from genomic DNA, followed by digestion and ligation into the pBAD18-Kan plasmid. All constructs were confirmed by sequencing (Mobix lab, McMaster University).

3.4.3 Growth conditions and reporter assays.

For all assays described, cells were grown in rich, chemically defined medium (described in the appendix, chapter 3 supplemental materials, supplemental materials and methods) supplemented with the indicated concentrations of AdoCbl, Cbi (Cbi supplied

as cobinamide dicyanide) and DMB (all purchased from Sigma) at 37°C with shaking. The sensor plasmid was transformed into the indicated strain by electroporation and plated on LB agar plates supplemented with 50 µg/mL ampicillin. Individual colonies were picked and grown overnight. Saturated cultures were split out into fresh media (1/1000 dilution) containing the indicated vitamin B₁₂ compounds and grown for 6 hours (mid-late log phase). Cells were then assayed using the procedures described below. All data shown represent the average of at least three biological replicates, confirmed by at least two independent experiments. Fold repression values were obtained by dividing the calculated reporter activity for the unrepressed conditions by that calculated for the repressed conditions for each sample. All reporter activity was normalized for cell density using OD₆₀₀ measurements taken using a VERSAmax spectrophotometer (Molecular Devices).

For RFP assays, 1.5 ml of cells was pelleted and resuspended in 320 µl of PBS. For each biological replicate, 100 µl of resuspension was added to three wells of a black, half area 96 well plate (Corning 3686) and fluorescence was read at 555/583 nm excitation/emission using a Safire fluorometer (Tecan). Measurements from the three wells were averaged to give the value for a given biological replicate. Firefly luciferase assays were conducted using the Luciferase Assay System (Promega) and performed essentially according to the supplier instructions provided for bacterial cells. Luciferase activity was measured using the Lumat LB 9507 luminometer (Berthold Technologies). β-galactosidase assays were performed based on the method developed by Miller (Miller,

1972) modified essentially as described (Zhang and Bremer, 1995). OD_{420} measurements were taken using a Cary 100 spectrophotometer (Variant, Inc).

Assays featuring BluB (Figure 4B) were carried out using the pBAD30 version of the β -galactosidase sensor in order to allow for plasmid compatibility. The uninduced and induced samples were taken from the same overnight cultures. For the induced samples, L-arabinose (Sigma) was added to the media to a final concentration of 0.2%.

3.4.4 Ethanolamine utilization assays.

For all strains tested, overnight cultures were grown in rich, chemically defined media. 4.5 μ l of the saturated overnight cultures was then added to 4.5 ml of growth media in which ethanolamine was the sole nitrogen source. This medium consisted of glycerol (0.2%), ethanolamine (0.1%), KH_2PO_4 (3 mg/ml), K_2HPO_4 (7 mg/ml), $MgSO_4$ (2 mM) $CaCl_2$ (100 μ M) and the indicated B_{12} – related molecules at a concentration of 500 nM. The pH of the media was adjusted to ~7.2 using HCl. 200 μ l of these cultures was removed at the indicated time points and the OD_{600} was measured using a VERSAmax spectrophotometer (Molecular Devices).

3.5 RESULTS

3.5.1 Construction of riboswitch-based AdoCbl sensors.

The AdoCbl sensors presented here exploit the naturally occurring AdoCbl-responsive riboswitch that regulates the *E. coli btuB* gene (Nahvi et al., 2002). It has been shown to be very specific for AdoCbl, strongly discriminating against a range of structural analogs, including the other active form of vitamin B₁₂, methylcobalamin (Nahvi et al., 2002). The complete mechanism of genetic control is not fully understood for this riboswitch; however several studies indicate that the primary mechanism involves AdoCbl-induced sequestration of the ribosome binding site (shown schematically in Figure 1) (Franklund and Kadner, 1997; Lundrigan et al., 1991; Nou and Kadner, 1998, 2000). Evidence also suggests that the beginning of the *btuB* coding region is important for a second form of expressional control that occurs at the level of transcription (Franklund and Kadner, 1997; Lundrigan et al., 1991; Nou and Kadner, 1998). We have used this entire riboswitch, including the complete 5'-UTR and the first 70 amino acids of BtuB, as the sensing element for our AdoCbl sensors.

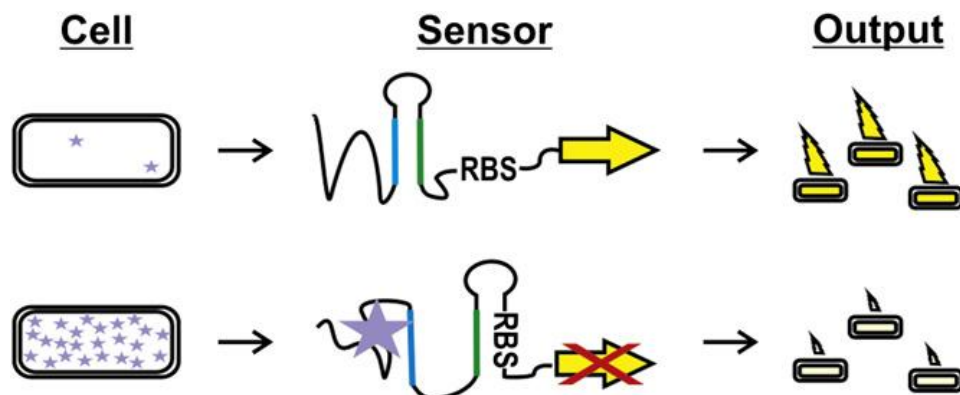


Figure 3-1 Depiction of Riboswitch-Based Intracellular Sensors. In the simplified example shown, cells containing a low concentration of the target metabolite (purple star) have a higher proportion of riboswitches in the unbound state. The RNA conformation in this state features an interaction between the green and blue segments and the ribosome binding site (RBS) is unstructured. This allows for efficient translation of the reporter gene (yellow arrow) downstream and results in a high level of signal output. At higher target concentrations riboswitches adopt the bound conformation. In this state the blue segment of the riboswitch interacts with the target. This frees up the green segment and allows it to interact with the RBS. The resulting structured RBS inhibits translation initiation of the reporter gene and leads to a low signal output.

Three parallel versions of the sensor were constructed that are identical other than the reporter gene used. Each of these probes employs a different mode of detection: β -galactosidase (colorimetric), firefly luciferase (luminescent) and DsRed-Express red fluorescent protein (fluorescent). Reporter genes were fused in frame to the riboswitch and cloned into vectors derived from the pBAD series of plasmids (Guzman et al., 1995). Complete details of these plasmids and their construction are provided in the appendix under chapter 3 supplemental materials (Figures S3-1 and supplemental materials and methods).

To test the ability of these constructs to detect AdoCbl they were transformed into the *E. coli* strain BW25113 (Datsenko and Wanner, 2000), henceforth referred to as wild-type (WT) cells. Reporter activity was measured for transformants grown both in media containing 1 μ M of AdoCbl and in media lacking AdoCbl or a suitable metabolic precursor (referred to as ‘no B₁₂’ conditions). All three reporter constructs showed a strong, ~100-fold repression in response of AdoCbl (Figure 3-2A). For each reporter, vectors featuring mutant riboswitches were also tested. These were identical to the WT sensors other than a double mutation (A₁₄₉A₁₅₀>TT) to a region of the riboswitch previously shown to disrupt AdoCbl binding (Nahvi et al., 2002). None of these mutant constructs showed a detectable response to AdoCbl, indicating that the riboswitch is indeed responsible for the observed regulation.

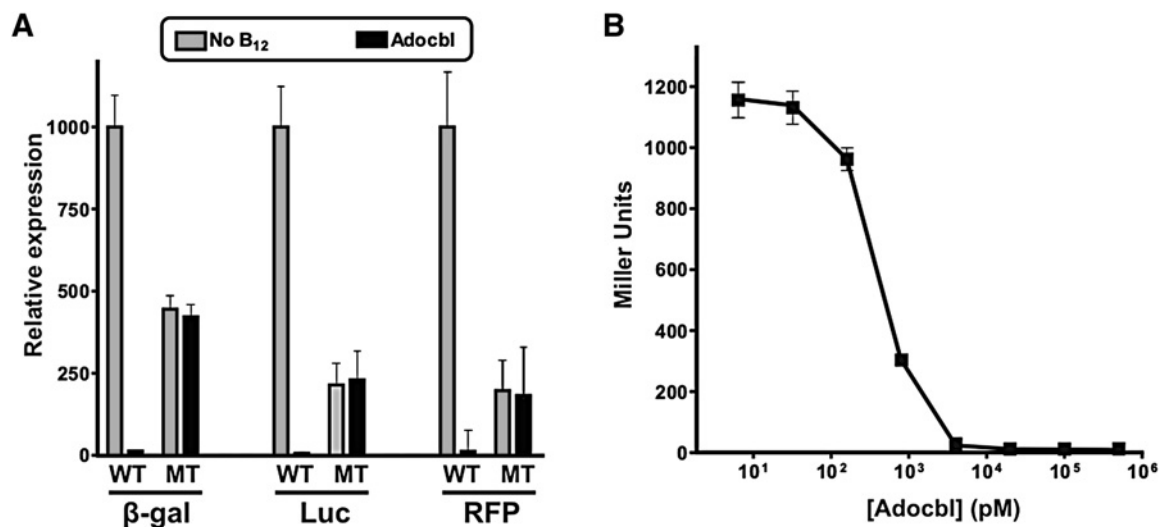


Figure 3-2 Intracellular Detection of AdoCbl by Riboswitch-Based Sensors. (A) AdoCbl detection using sensor constructs featuring three different reporter genes: β -galactosidase (β -gal), firefly luciferase (Luc), and the DsRed-Express RFP. For each reporter, a parallel construct featuring a double mutation that disrupts ligand binding (MT) was also tested. Assays were conducted using cells grown in medium either lacking vitamin B₁₂ (no B₁₂) or containing 1 mM AdoCbl. Reporter activity was normalized for all samples by setting the values for the WT sensors in the no B₁₂ conditions to 1000. (B) Dose response of the WT β -galactosidase sensor over a wide range of AdoCbl concentrations. Error bars represent standard deviation in all graphs.

We next sought to determine the dose response of our sensors over a broad range of target concentrations. We were able to detect AdoCbl at high picomolar media concentrations, observing a saturated response by the low-mid nanomolar range (Figure 3-2B). The observed sensitivity in this low concentration range is reflective of both the high affinity of the riboswitch and the ability of *E. coli* to concentrate vitamin B₁₂ compounds intracellularly using an effective transport system. *E. coli* cells grown in media containing 10 nM vitamin B₁₂ have been found to accumulate an intracellular

concentration of ~25 μM (Reynolds et al., 1980). This suggests that the intracellular AdoCbl concentration required to saturate our sensors is in the low micromolar range. This is roughly consistent with the high nanomolar apparent K_d determined for this riboswitch *in vitro* (Gallo et al., 2008; Nahvi et al., 2002).

We were impressed by the observed sensitivity and signal amplitude of the sensors. However, we were keen to assess their ability to be applied in the study of biological processes that involve AdoCbl, as well as to more rigorously test their capacity to detect physiologically relevant fluctuations in AdoCbl concentrations. In the following sections we apply the riboswitch-based AdoCbl sensors to monitor the changing levels of their target in response to genetic manipulations and varying environmental conditions. The aim of these experiments was to examine the synthesis and import of AdoCbl in *E. coli*, with an emphasis on the importance of the genetic factors involved in these processes.

3.5.2 Riboswitch-based sensors can be used to monitor AdoCbl metabolism.

AdoCbl is a large, chemically complex molecule containing a central cobalt atom that is coordinated by six ligands (see Figure 3-3). Four of these interactions lie in a plane with a porphyrin-like structure known as the corrin ring. The final two ligands for cobalt are an adenosyl moiety (upper axial ligand) and a nucleotide loop (lower axial ligand). *E. coli* cannot synthesize vitamin B₁₂ compounds *de novo* because it lacks the ~20 dedicated genes required to synthesize the central ring structure. *E. coli* is, however, able to salvage incomplete corrinoid molecules from the environment by carrying out the final steps in AdoCbl synthesis that involve the addition of the upper and lower axial ligands for cobalt

(Lawrence and Roth, 1996). To do so they require cobinamide (Cbi) as a minimal metabolic precursor. 5,6-dimethylbenzimidazole (DMB), a base incorporated into the nucleotide loop of cobalamins, must also be provided exogenously to produce elevated levels of AdoCbl (see below).

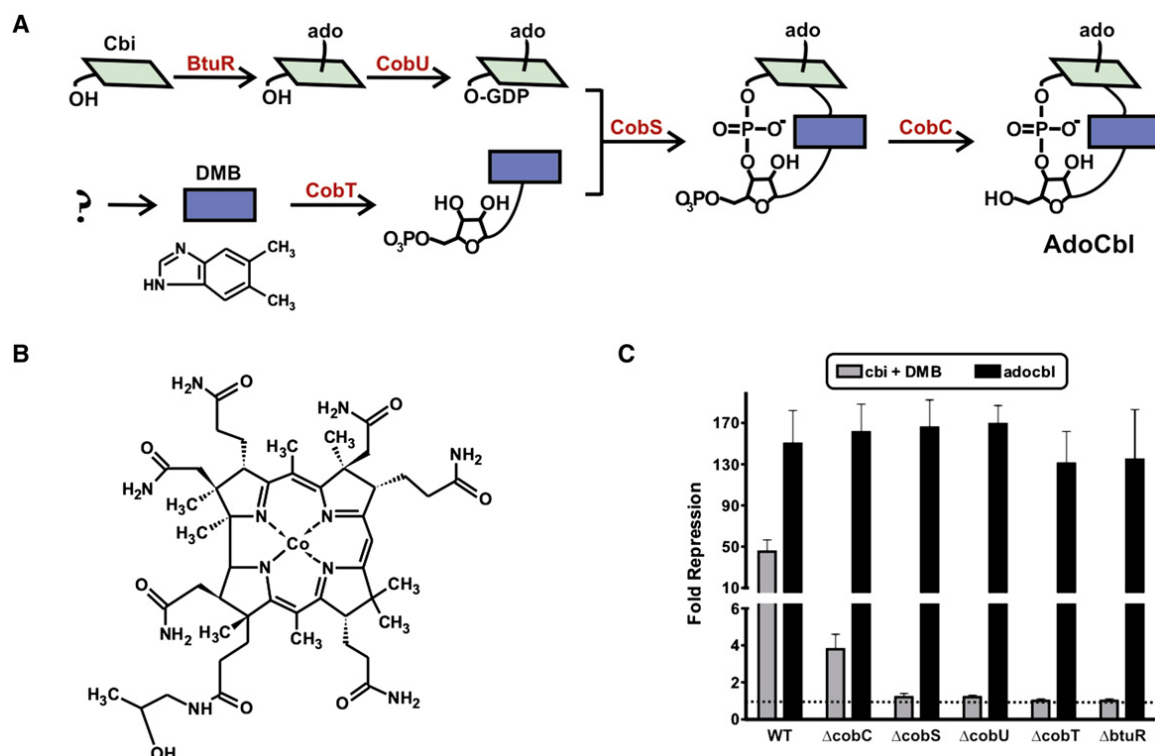


Figure 3-3 Using Riboswitch-Based Sensors to Probe AdoCbl Metabolism. (A) Schematic of the predicted *E. coli* AdoCbl biosynthetic pathway from minimal precursors. Ado, 5'-deoxyadenosyl. (B) Chemical structure of Cbi. (C) Assay probing AdoCbl concentrations in strains carrying deletions to each of the genes from the pathway shown in (A). The data is expressed as the fold repression observed for cells grown in the presence of Cbi and DMB or AdoCbl relative to the no B12 control. The dotted line indicates no repression and thus no detectable AdoCbl. Error bars represent the standard deviation.

Studies of vitamin B₁₂ metabolism have generally focused on organisms such as *S. enterica* and *P. denitrificans* that are able to carry out the full *de novo* synthesis. Figure 3-3A shows a simplified schematic of the predicted AdoCbl biosynthetic pathway in *E. coli*, based largely on studies of homologous genes in related *Salmonella* species. To determine the importance of each of the enzymes shown in this pathway, we tested the ability of deletion strains to synthesize AdoCbl. The β -galactosidase version of the sensor was transformed into each strain and was used to monitor AdoCbl levels in cells grown in medium containing the metabolic precursors Cbi and DMB, as well as in control media containing no B₁₂ (Figure 3-3C). For each strain, the reporter activity in the no B₁₂ control was compared to the activity in the media containing Cbi and DMB to provide a fold-repression. A fold repression greater than 1 implies that AdoCbl is being produced. A minor variation in the raw reporter values was observed (< 2-fold) for the unrepressed expression from strain to strain as shown in appendix, chapter 3 supplemental materials, Figure S3-2. However, this day-to-day variation was not observed for the fold-repression values, which were very consistent. Additionally, for each strain the mutant construct was tested (Figure S3-3). As expected, the mutant construct was not responsive to AdoCbl in any of the strains and the raw reporter activities only varied to a small extent from strain to strain, as observed for the wild-type sensor.

The reporter activities show that the $\Delta cobC$ strain retained a partial ability to synthesize AdoCbl, while the other deletions could not produce sufficient levels to be detected by our sensors. For *btuR*, *cobS* and *cobU*, this is consistent with previous

observations that indicate that these genes are essential for cobalamin synthesis (Lawrence and Roth, 1995; Lundrigan and Kadner, 1989; O'Toole et al., 1993). By contrast, strains lacking CobC and CobT activity maintain sufficient levels of cobalamin synthesis to grow under methylcobalamin-requiring conditions (Lawrence and Roth, 1995; O'Toole et al., 1993; O'Toole et al., 1994). While this observation has not yet been explained for *cobC*, genetic evidence from *S. enterica* suggests that the loss of *cobT* can be partially compensated for by the activity of CobB, a member of the SIR2 family of regulators (Tsang and Escalante-Semerena, 1998).

We were keen to investigate how the levels of AdoCbl detected by the sensors in these strains relate to physiological processes that involve AdoCbl in *E. coli*. Toward this end, we assessed the ability of the metabolic mutants to grow in media where ethanolamine was the sole nitrogen source. Under these conditions AdoCbl is required for growth as it is an essential cofactor for the first step in the ethanolamine utilization pathway (Bradbeer, 1965; Chang and Chang, 1975). Both metabolic mutants previously shown to be dispensable for methylcobalamin-dependent growth ($\Delta cobT$ and $\Delta cobC$) were assayed, as were WT and $\Delta cobS$ cells, which served as controls. None of the tested deletion strains were able produce sufficient levels of AdoCbl to support growth under these conditions (Figure 3-4). When AdoCbl was provided in the medium each of the strains grew like the WT strain, confirming that these strains were experiencing a lesion in AdoCbl metabolism. These results indicate that the cobalamin synthesis that occurs in the $\Delta cobT$ strain, while below the detection threshold of our sensors, is also less than that required for ethanolamine utilization. This is consistent with previous results for *S.*

enterica (Anderson et al., 2008). Our sensors can detect the cobalamins produced in $\Delta cobC$ strain, which is also unable to support growth on ethanolamine. However, it is possible that this could be due to the riboswitch detecting CobC's phosphorylated substrate (see Discussion).

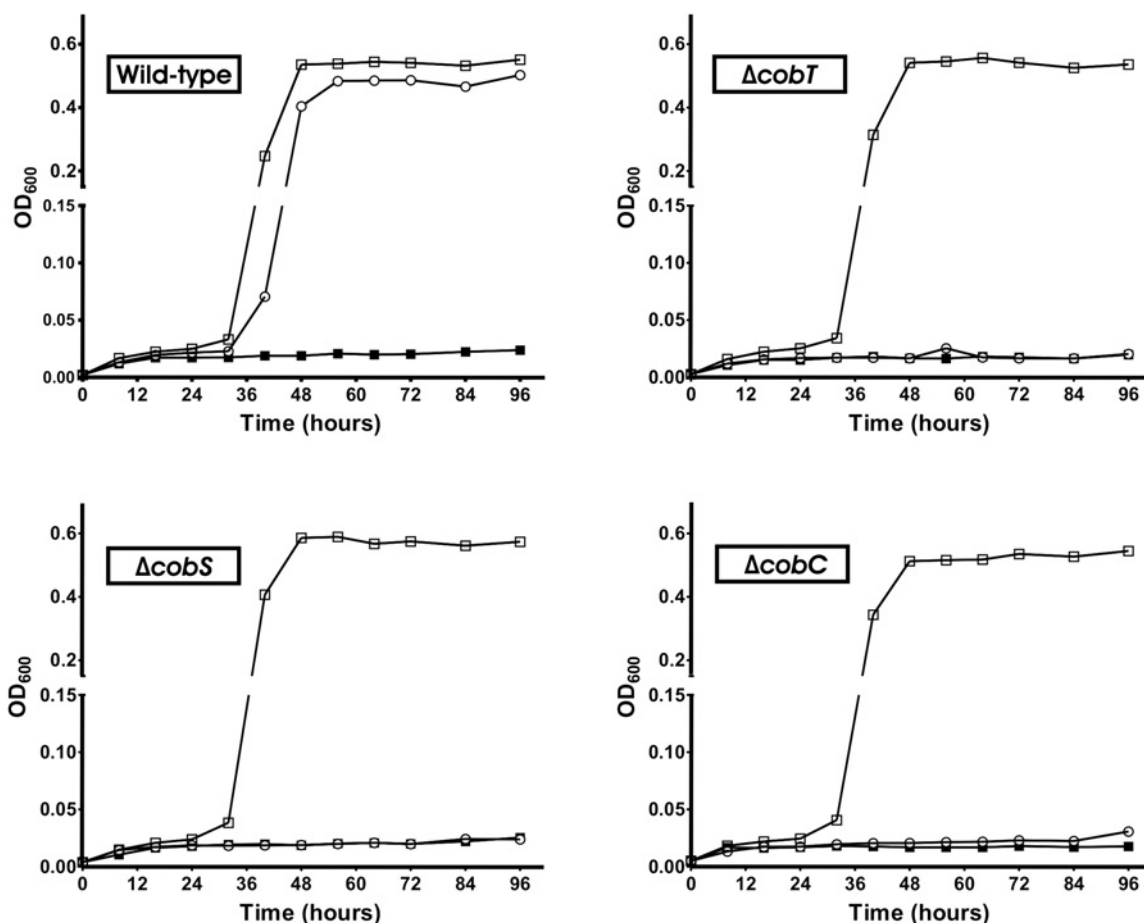


Figure 3-4 Growth of Metabolic Deletion Strains Using Ethanolamine as the Sole Nitrogen Source. The ability of selected strains to produce sufficient AdoCbl to sustain AdoCbl-dependent ethanolamine utilization was assessed in media containing Cbi and DMB (open circles), as well as in control media containing no B12 (filled boxes: negative control) and AdoCbl (open boxes: positive control). Panels are identical other than the strain used, which is indicated in the top left corner of each graph. Data points represent the mean of three biological replicates.

3.5.3 Riboswitch-based sensors can detect the availability of B₁₂ lower axial ligands

We next turned our attention to DMB, the base incorporated as the lower axial ligand in AdoCbl. Previous studies have shown that *S. enterica* produces DMB at very low levels from flavonoid precursors (Keck et al., 1998). These levels are sufficiently low to incite the suggestion that the DMB produced by *S. enterica* is through a non-enzymatic chemical reaction or as a side product of flavin-utilizing enzymes (Maggio-Hall et al., 2003; Taga and Walker, 2008). Consistent with this, when Cbi is provided without DMB our sensors show a very low level of repression of less than 2-fold (Figure 3-5A). Under these conditions it is likely that some or all of the repression observed is due to the synthesis of *pseudo*-AdoCbl, a functional variant of AdoCbl in which DMB is replaced by adenine. In *S. enterica*, while DMB is used preferentially, adenine-containing cobalamins are produced using the same metabolic pathway and can be used to support B₁₂-requiring processes when DMB is not available (Anderson et al., 2008). An AdoCbl analog in which DMB was replaced with 2-methyladenine was recently found to bind the *E. coli* *btuB* riboswitch with a very minor reduction in affinity, indicating that this riboswitch has not evolved a strong selectivity for DMB (Gallo et al., 2008). To determine whether our sensors also detect *Pseudo*-AdoCbl, we assayed cells grown in media containing Cbi and different concentrations of adenine. We observed a dose-dependent repression of reporter activity from our sensor. This demonstrates that *pseudo*-AdoCbl is produced at an elevated level under these conditions, which can be detected by our sensors (Figure 3-5A).

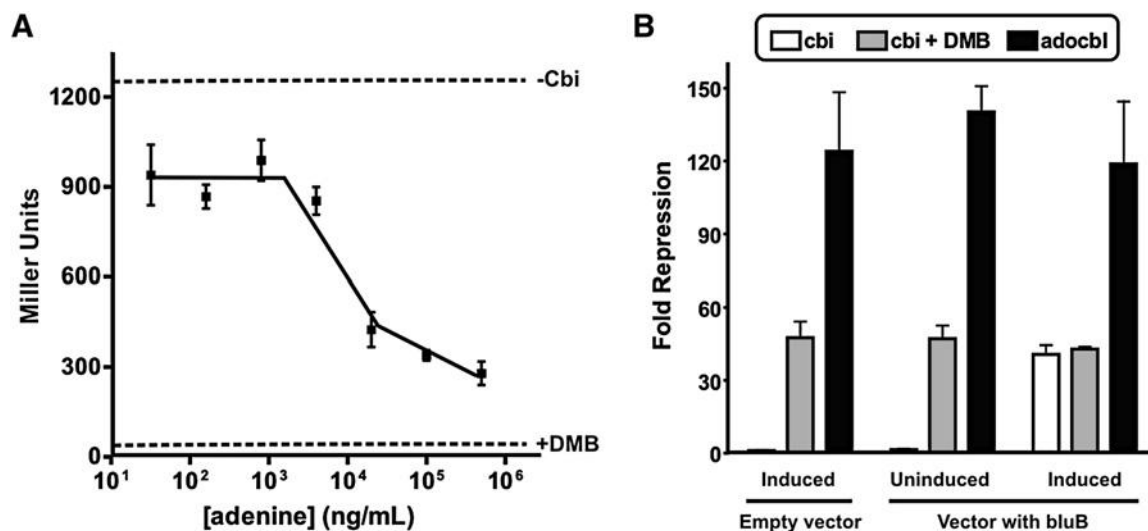


Figure 3-5 Monitoring the Availability of Lower Axial Ligands Using Riboswitch-Based Sensors. (A) The affect of adenine concentration on sensor response in media containing Cbi but not DMB. Dotted lines indicate the signal for controls lacking Cbi and containing both Cbi and DMB. (B) The effect of *M. smegmatis* BluB expression on AdoCbl concentrations in the absence of exogenous DMB. BluB was expressed from pBAD18 by adding the inducer L-arabinose and compared to the empty vector and uninduced controls. The ability of BluB to produce DMB for AdoCbl synthesis was assessed in media containing Cbi alone and compared to control samples grown in media containing Cbi and DMB or AdoCbl. Error bars represent the standard deviation for all graphs.

While the source of DMB remains unclear for *E. coli*, it was recently uncovered that some organisms boast an enzyme, BluB, which catalyzes the conversion of reduced flavin mononucleotide to DMB (Campbell et al., 2006; Gray and Escalante-Semerena, 2007; Taga et al., 2007). While *E. coli* lacks an obvious BluB homologue, we were intrigued by the possibility that BluB could be recombinantly expressed in *E. coli* to provide an enhanced ability to synthesize DMB. To test this hypothesis, the *bluB* gene from *Mycobacterium smegmatis* (*M. smegmatis*) was cloned and inducibly expressed in

WT *E. coli* cells also carrying the β -galactosidase sensor plasmid. Remarkably, recombinantly expressed BluB produced AdoCbl at the same level as samples where excess DMB was provided exogenously (Figure 3-5B). This demonstrates that a single gene from a distantly related organism can provide *E. coli* with an efficient means of synthesizing DMB, where one does not appear to exist. Furthermore, this displays the ability of the riboswitch sensors to detect genetic modifications that result in a gain of function, in addition to the loss of function detected for the metabolic deletions.

3.5.4 Riboswitch-based sensors can be used to monitor AdoCbl transport

Vitamin B₁₂ import is amongst *E. coli*'s most intricate and impressive transport systems. Because it is too large to permeate outer membrane porins, vitamin B₁₂ is actively transported across both the outer membrane (OM) and the cytoplasmic membrane (CM) in a process that requires at least seven genes (shown schematically in Figure 3-6A). To assess the importance of each of these genes in the transport process, we tested the effects of their deletion on intracellular AdoCbl levels for cells grown in media containing 10 nM and 1 μ M AdoCbl (Figure 3-6B). Raw reporter expression of the unrepressed samples and the response of the mutant sensor for each strain were consistent with those observed for the metabolic deletion strains (shown in Figures S3-2 and S3-3). As expected, deletion of each of the transport genes had a significant impact on AdoCbl import. However, when present in the media at 1 μ M, AdoCbl could be detected intracellularly in each of the strains, presumably due to a very low level of diffusion. This

is consistent with previous studies that found genetically isolated mutations to *btuB*, *tonB*, *btuC* and *btuD* could survive in vitamin B₁₂-requiring conditions when it was supplied at a sufficiently high concentration (Bassford and Kadner, 1977; Di Girolamo et al., 1971). Also consistent with these results was our observation that deletions of OM factors had a more drastic effect than those of IM factors. AdoCbl could be detected intracellularly at the 10 nM concentration for the $\Delta btuC$, $\Delta btuD$ and $\Delta btuF$ strains, likely because AdoCbl would be concentrated in the periplasm in these strains, leading to a higher level of incidental entry into the cytoplasm, as proposed previously (Roth et al., 1996). Importantly, deletions to the IM transport proteins have been shown to be unable to support growth on ethanolamine when supplied with 5 μ M vitamin B₁₂ (Cadieux et al., 2002). This indicates that the sensors presented here can detect AdoCbl levels lower than those required for growth on ethanolamine as the sole nitrogen source.

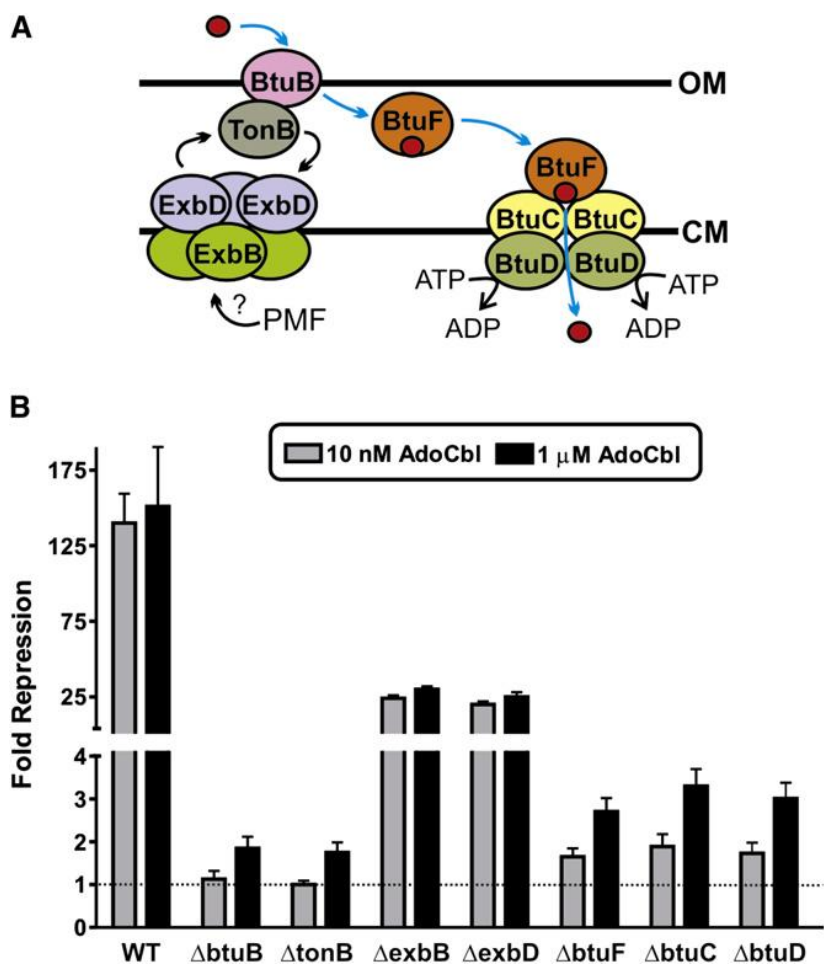


Figure 3-6 Using Riboswitch-Based Sensors to Probe AdoCbl Transport. (A) Schematic of vitamin B12 import in *E. coli*. The transport of vitamin B12 (small red circle) is instigated at the OM by the BtuB receptor, which binds vitamin B12 and a wide range of structural analogs with sub-nanomolar affinity. BtuB-mediated import into the periplasm requires an interaction with TonB, which harnesses the proton motive force (PMF) to energize import in conjunction with CM proteins ExbB/ExbD. Once in the periplasm vitamin B12 is bound by BtuF, which delivers its cargo to the CM transporter made up of BtuC/BtuD. BtuC/BtuD, an ABC transporter, uses ATP hydrolysis to pump vitamin B12 into the cytoplasm. Many aspects of the energetics of OM transport are unclear. The shuttling of TonB between membranes as shown is speculative, as is the stoichiometry of the active TonB/ExbB/ExbD complex. (B) Riboswitch sensor assay monitoring AdoCbl import in strains carrying deletions to the genes shown in (A). Assays conducted as described in the caption for Figure 3C. Error bars represent the standard deviation.

A far less severe phenotype was observed for deletions of the *exbB* and *exbD* genes than for any of the other five genes. This was not surprising, since it has been previously shown that *tonB* function can be partially maintained by the TolQ/TolR in the absence of ExbB/ExbD (Braun and Herrmann, 1993). Despite this compensation, the sensors can readily detect a defect in these strains, showing a 5-fold change in signal compared to the WT strain.

3.6 DISCUSSION

This study demonstrates the capacity of riboswitches to be used as intracellular probes of a target metabolite. The sensors constructed here demonstrate several desirable properties. They display remarkable sensitivity, detecting AdoCbl levels at the low end of the biologically relevant range, a particularly impressive feat given the low abundance of this metabolite. Furthermore, deletions to metabolic genes showed a total lack of AdoCbl synthesis, indicating that the sensors are extremely specific with respect to Cbi and other metabolic precursor molecules that presumably accumulate under these conditions. One potential exception to this is the phosphorylated AdoCbl variant that presumably accumulates in the $\Delta cobC$ strain. The sensors detected a 3-4 fold repression in this strain when supplied with Cbi and DMB, which could reflect a detection of CobC's substrate. Alternatively, another phosphatase could partially compensate for the loss of CobC, leading to a reduced level of AdoCbl synthesis. Further experiments will be required to distinguish these possibilities.

For each of the genes predicted to be involved in AdoCbl transport and metabolism, defects were readily detected using the riboswitch-based sensors. Importantly, the extent of the resulting deficiency could also be assessed. This is most clear for the transport deletions, which showed the expected hierarchy with the OM transport genes showing the most severe phenotype, followed by the IM genes and finally *exbB/exbD*, whose loss can be compensated for by the *tolQ/tolR* genes. An important advantage of using the sensors presented here rather than the growth phenotypes typically used to assess these mutants is that there is no selective pressure to accumulate mutations. Such suppressor mutations have been observed previously in growth experiments featuring AdoCbl mutants (Cadieux et al., 2002) and can lead to misleading data or incorrect conclusions.

We were unable to detect AdoCbl synthesis in the $\Delta cobT$ strain despite previous data indicating that a low level of cobalamin production is retained when *cobT* is disrupted (Lawrence and Roth, 1995; O'Toole et al., 1993). These results are based on the viability of methylcobalamin-requiring mutants, which can grow using as few as 20 cobalamin molecules per cell (Di Girolamo et al., 1971). Under conditions where cobalamins are available at extremely low quantities it is likely that a significant portion of this pool is not in the free, adenosylated form that is detected by the sensors presented here. Furthermore, for applications where the detection of an extremely low concentration range is desirable, chromosomal integration of the sensor could provide higher sensitivity. Plasmid-encoded sensors were desirable for this study to facilitate transfer between strains. However, the multi-copy nature of the plasmid sensors could result in a

scenario where, at sufficiently low concentrations, the target is not in excess. This is particularly applicable for cobalamins because they are required at such low levels.

One drawback of riboswitch-based sensors is that they do not directly detect the target molecule, a common limitation of intracellular detection methods. The signal measured for riboswitch-based sensors relies on the level of expression of a reporter gene, which can be influenced by varying environmental and genetic conditions. The use of these probes therefore requires proper controls to correct for factors that influence reporter expression in a non-specific manner. In the present study, data are expressed as a fold repression in signal when different strains are assayed. This allows for samples lacking the target molecule to act as controls for any non-specific effects that influence reporter expression. For riboswitch ligands where a “no target” control is not possible, mutant constructs that negate riboswitch binding (such as those presented in Figure 3-2) would serve as useful controls. The riboswitch sensors, as constructed, are meant to provide information on relative concentrations. Technologies that monitor macromolecule levels, such as western blots or oligonucleotide microarrays, can be very informative without providing quantitative information on cellular concentrations. Applying the appropriate controls, the same is true for small molecule detection.

The sensor presented here makes use of a naturally occurring riboswitch. Natural riboswitches make ideal sensors as they presumably evolve to detect physiologically relevant target concentrations. A growing number of riboswitch classes have been identified, consisting of a diverse set of biologically important molecules and key

metabolic nodes (Serganov, 2009). There is therefore a great deal of opportunity to monitor a wide range of biological processes using riboswitch sensors derived from natural sources. Importantly, we and others have shown the capacity to convert artificially isolated aptamers into riboswitches and to modify the characteristics of naturally occurring riboswitches (Desai and Gallivan, 2004; Dixon et al.; Fowler et al., 2008; Lynch et al., 2009; Nomura and Yokobayashi, 2007; Weigand et al., 2008). This provides a great deal of flexibility to riboswitch-based sensing, presenting the possibility of creating custom sensors to detect a specific metabolite of interest. The development of an array of riboswitch-based sensors would open the door to multiplexing experiments where bacterial strains could be constructed that contain a number of sensors, each detecting a unique metabolite using a distinct reporter. Furthermore, riboswitch-based sensing need not be limited to prokaryotic organisms. While the majority of riboswitches that have been isolated to date have been found in prokaryotes, the thiamine pyrophosphate riboswitch is also found in higher organisms, including yeast and plants (Bocobza et al., 2007; Cheah et al., 2007; Kubodera et al., 2003; Wachter et al., 2007). When this riboswitch was fused to fluorescent proteins and introduced into *Arabidopsis*, the levels of fluorescence from leaves could be controlled by the addition of exogenous thiamine (Bocobza et al., 2007; Wachter et al., 2007). This suggests there may be potential for riboswitch-based sensors as non-invasive intracellular probes in eukaryotic organisms as well, a prospect discussed in a recent review (Bocobza and Aharoni, 2008).

A major advantage of riboswitch-based sensing is its capacity for high throughput applications. Because the readout is a simple reporter assay, this means of following

small molecule concentrations could readily be scaled up to multi-well plate format to screen a large number of environmental or genetic conditions. One exciting prospect is screening single gene deletion or over-expression libraries looking for genetic factors that affect metabolite levels. Such a screen could presumably identify not only genes directly involved in the target's metabolism or import, but also genes involved in regulation, export or in related metabolic pathways. Another intriguing possibility is screening for small molecule inhibitors of a target metabolic pathway using riboswitch-based sensors to monitor the concentration of the end product. Such a screen would have the advantage of using a living organism to provide more biologically relevant data and would also provide insight into the target of hit molecules.

3.7 REFERENCES

Abeles, R.H., and Lee, H.A., Jr. (1961). An intramolecular oxidation-reduction requiring a cobamide coenzyme. *J. Biol. Chem.* 236, 2347–2350.

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

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2, 2006.0008.

Barrick, J.E., and Breaker, R.R. (2007). The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol.* 8, R239. Bassford, P.J., Jr., and Kadner,

- R.J. (1977). Genetic analysis of components involved in vitamin B12 uptake in *Escherichia coli*. *J. Bacteriol.* *132*, 796–805.
- Binkowski, B., Fan, F., and Wood, K. (2009). Engineered luciferases for molecular sensing in living cells. *Curr. Opin. Biotechnol.* *20*, 14–18.
- Bocobza, S.E., and Aharoni, A. (2008). Switching the light on plant riboswitches. *Trends Plant Sci.* *13*, 526–533.
- Bocobza, S., Adato, A., Mandel, T., Shapira, M., Nudler, E., and Aharoni, A. (2007). Riboswitch-dependent gene regulation and its evolution in the plant kingdom. *Genes Dev.* *21*, 2874–2879.
- Bourgogne, A., Hilsenbeck, S.G., Dunny, G.M., and Murray, B.E. (2006). Comparison of OG1RF and an isogenic *fsrB* deletion mutant by transcriptional analysis: the *Fsr* system of *Enterococcus faecalis* is more than the activator of gelatinase and serine protease. *J. Bacteriol.* *188*, 2875–2884.
- Bradbeer, C. (1965). The clostridial fermentations of choline and ethanolamine. II. Requirement for a cobamide coenzyme by an ethanolamine deaminase. *J. Biol. Chem.* *240*, 4675–4681.
- Braun, V., and Herrmann, C. (1993). Evolutionary relationship of uptake systems for biopolymers in *Escherichia coli*: cross-complementation between the TonB-ExbB-ExbD and the TolA-TolQ-TolR proteins. *Mol. Microbiol.* *8*, 261–268.
- Cadieux, N., Bradbeer, C., Reeger-Schneider, E., Koster, W., Mohanty, A.K., Wiener, M.C., and Kadner, R.J. (2002). Identification of the periplasmic cobalamin-binding protein BtuF of *Escherichia coli*. *J. Bacteriol.* *184*, 706–717.
- Campbell, G.R., Taga, M.E., Mistry, K., Lloret, J., Anderson, P.J., Roth, J.R., and Walker, G.C. (2006). *Sinorhizobium meliloti* bluB is necessary for production of 5,6-dimethylbenzimidazole, the lower ligand of B12. *Proc. Natl. Acad. Sci. USA* *103*, 4634–4639.
- Chang, G.W., and Chang, J.T. (1975). Evidence for the B12-dependent enzyme ethanolamine deaminase in *Salmonella*. *Nature* *254*, 150–151.
- Cheah, M.T., Wachter, A., Sudarsan, N., and Breaker, R.R. (2007). Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. *Nature* *447*, 497–500.
- Cho, E.J., Lee, J.-W., and Ellington, A.D. (2009). Applications of aptamers as sensors. *Annu. Rev. Anal. Chem.* *2*, 241–264.
- Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* *97*, 6640–6645.
- Desai, S.K., and Gallivan, J.P. (2004). Genetic screens and selections for small molecules based on a synthetic riboswitch that activates protein translation. *J. Am. Chem. Soc.* *126*, 13247–13254.

- Di Girolamo, P.M., Kadner, R.J., and Bradbeer, C. (1971). Isolation of vitamin B12 transport mutants of *Escherichia coli*. *J. Bacteriol.* *106*, 751–757.
- Dixon, N., Duncan, J.N., Geerlings, T., Dunstan, M.S., McCarthy, J.E., Leys, D., and Micklefield, J. (2010). Reengineering orthogonally selective riboswitches. *Proc. Natl. Acad. Sci. USA* *107*, 2830–2835.
- Ellington, A.D., and Szostak, J.W. (1990). In vitro selection of RNA molecules that bind specific ligands. *Nature* *346*, 818–822.
- Fan, F., Binkowski, B.F., Butler, B.L., Stecha, P.F., Lewis, M.K., and Wood, K.V. (2008). Novel genetically encoded biosensors using firefly luciferase. *ACS Chem. Biol.* *3*, 346–351.
- Fehr, M., Frommer, W.B., and Lalonde, S. (2002). Visualization of maltose uptake in living yeast cells by fluorescent nanosensors. *Proc. Natl. Acad. Sci. USA* *99*, 9846–9851.
- Fowler, C.C., Brown, E.D., and Li, Y. (2008). A FACS-based approach to engineering artificial riboswitches. *ChemBioChem* *9*, 1906–1911.
- Franklund, C.V., and Kadner, R.J. (1997). Multiple transcribed elements control expression of the *Escherichia coli* *btuB* gene. *J. Bacteriol.* *179*, 4039–4042.
- Gallo, S., Oberhuber, M., Sigel, R.K., and Krautler, B. (2008). The corrin moiety of coenzyme B12 is the determinant for switching the *btuB* riboswitch of *E. coli*. *ChemBioChem* *9*, 1408–1414.
- Gray, M.J., and Escalante-Semerena, J.C. (2007). Single-enzyme conversion of FMNH₂ to 5,6-dimethylbenzimidazole, the lower ligand of B12. *Proc. Natl. Acad. Sci. USA* *104*, 2921–2926.
- Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* *177*, 4121–4130.
- Huang, Y., Leming, C.L., Suyemoto, M., and Altier, C. (2007). Genome-wide screen of *Salmonella* genes expressed during infection in pigs, using in vivo expression technology. *Appl. Environ. Microbiol.* *73*, 7522–7530.
- Joseph, B., Przybilla, K., Stuhler, C., Schauer, K., Slaghuis, J., Fuchs, T.M., and Goebel, W. (2006). Identification of *Listeria monocytogenes* genes contributing to intracellular replication by expression profiling and mutant screening. *J. Bacteriol.* *188*, 556–568.
- Keck, B., Munder, M., and Renz, P. (1998). Biosynthesis of cobalamin in *Salmonella typhimurium*: transformation of riboflavin into the 5,6-dimethylbenzimidazole moiety. *Arch. Microbiol.* *171*, 66–68.
- Klumpp, J., and Fuchs, T.M. (2007). Identification of novel genes in genomic islands that contribute to *Salmonella typhimurium* replication in macrophages. *Microbiology* *153*, 1207–1220.

- Korbel, J.O., Doerks, T., Jensen, L.J., Perez-Iratxeta, C., Kaczanowski, S., Hooper, S.D., Andrade, M.A., and Bork, P. (2005). Systematic association of genes to phenotypes by genome and literature mining. *PLoS Biol.* 3, e134.
- Kubodera, T., Watanabe, M., Yoshiuchi, K., Yamashita, N., Nishimura, A., Nakai, S., Gomi, K., and Hanamoto, H. (2003). Thiamine-regulated gene expression of *Aspergillus oryzae* thiA requires splicing of the intron containing a riboswitch-like domain in the 5'-UTR. *FEBS Lett.* 555, 516–520.
- Lawhon, S.D., Frye, J.G., Suyemoto, M., Porwollik, S., McClelland, M., and Altier, C. (2003). Global regulation by CsrA in *Salmonella typhimurium*. *Mol. Microbiol.* 48, 1633–1645.
- Lawrence, J.G., and Roth, J.R. (1995). The cobalamin (coenzyme B12) biosynthetic genes of *Escherichia coli*. *J. Bacteriol.* 177, 6371–6380.
- Lawrence, J.G., and Roth, J.R. (1996). Evolution of coenzyme B12 synthesis among enteric bacteria: evidence for loss and reacquisition of a multigene complex. *Genetics* 142, 11–24.
- Lundrigan, M.D., and Kadner, R.J. (1989). Altered cobalamin metabolism in *Escherichia coli* btuR mutants affects btuB gene regulation. *J. Bacteriol.* 171, 154–161.
- Lundrigan, M.D., Koster, W., and Kadner, R.J. (1991). Transcribed sequences of the *Escherichia coli* btuB gene control its expression and regulation by vitamin B12. *Proc. Natl. Acad. Sci. USA* 88, 1479–1483.
- Lynch, S.A., Topp, S., and Gallivan, J.P. (2009). High-throughput screens to discover synthetic riboswitches. *Methods Mol. Biol.* 540, 321–333.
- Maadani, A., Fox, K.A., Mylonakis, E., and Garsin, D.A. (2007). *Enterococcus faecalis* mutations affecting virulence in the *Caenorhabditis elegans* model host. *Infect. Immun.* 75, 2634–2637.
- Maggio-Hall, L.A., Dorrestein, P.C., Escalante-Semerena, J.C., and Begley, T.P. (2003). Formation of the dimethylbenzimidazole ligand of coenzyme B(12) under physiological conditions by a facile oxidative cascade. *Org. Lett.* 5, 2211–2213.
- Medintz, I.L. (2006). Recent progress in developing FRET-based intracellular sensors for the detection of small molecule nutrients and ligands. *Trends Biotechnol.* 24, 539–542.
- Miller, J.H. (1972). *Experiments in Molecular Genetics* (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory).
- Mironov, A.S., Gusarov, I., Rafikov, R., Lopez, L.E., Shatalin, K., Kreneva, R.A., Perumov, D.A., and Nudler, E. (2002). Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* 111, 747–756.
- Miyawaki, A., Llopis, J., Heim, R., McCaffery, J.M., Adams, J.A., Ikura, M., and Tsien, R.Y. (1997). Fluorescent indicators for Ca²⁺ based on green fluorescent proteins and calmodulin. *Nature* 388, 882–887.

- Nahvi, A., Sudarsan, N., Ebert, M.S., Zou, X., Brown, K.L., and Breaker, R.R. (2002). Genetic control by a metabolite binding mRNA. *Chem. Biol.* *9*, 1043.
- Navani, N.K., and Li, Y. (2006). Nucleic acid aptamers and enzymes as sensors. *Curr. Opin. Chem. Biol.* *10*, 272–281.
- Nomura, Y., and Yokobayashi, Y. (2007). Reengineering a natural riboswitch by dual genetic selection. *J. Am. Chem. Soc.* *129*, 13814–13815.
- Nou, X., and Kadner, R.J. (1998). Coupled changes in translation and transcription during cobalamin-dependent regulation of *btuB* expression in *Escherichia coli*. *J. Bacteriol.* *180*, 6719–6728.
- Nou, X., and Kadner, R.J. (2000). Adenosylcobalamin inhibits ribosome binding to *btuB* RNA. *Proc. Natl. Acad. Sci. USA* *97*, 7190–7195.
- O’Toole, G.A., Rondon, M.R., and Escalante-Semerena, J.C. (1993). Analysis of mutants of *Salmonella typhimurium* defective in the synthesis of the nucleotide loop of cobalamin. *J. Bacteriol.* *175*, 3317–3326.
- O’Toole, G.A., Trzebiatowski, J.R., and Escalante-Semerena, J.C. (1994). The *cobC* gene of *Salmonella typhimurium* codes for a novel phosphatase involved in the assembly of the nucleotide loop of cobalamin. *J. Biol. Chem.* *269*, 26503–26511.
- Paulmurugan, R., and Gambhir, S.S. (2006). An intramolecular folding sensor for imaging estrogen receptor-ligand interactions. *Proc. Natl. Acad. Sci. USA* *103*, 15883–15888.
- Reynolds, P.R., Mottur, G.P., and Bradbeer, C. (1980). Transport of vitamin B12 in *Escherichia coli*. Some observations on the roles of the gene products of *BtuC* and *TonB*. *J. Biol. Chem.* *255*, 4313–4319.
- Romoser, V.A., Hinkle, P.M., and Persechini, A. (1997). Detection in living cells of Ca²⁺-dependent changes in the fluorescence emission of an indicator composed of two green fluorescent protein variants linked by a calmodulin binding sequence. A new class of fluorescent indicators. *J. Biol. Chem.* *272*, 13270–13274.
- Roth, J.R., Lawrence, J.G., and Bobik, T.A. (1996). Cobalamin (coenzyme B12): synthesis and biological significance. *Annu. Rev. Microbiol.* *50*, 137–181.
- Serganov, A. (2009). The long and the short of riboswitches. *Curr. Opin. Struct. Biol.* *19*, 251–259.
- Taga, M.E., and Walker, G.C. (2008). Pseudo-B12 joins the cofactor family. *J. Bacteriol.* *190*, 1157–1159.
- Taga, M.E., Larsen, N.A., Howard-Jones, A.R., Walsh, C.T., and Walker, G.C. (2007). *BluB* cannibalizes flavin to form the lower ligand of vitamin B12. *Nature* *446*, 449–453.
- Toraya, T., Honda, S., and Fukui, S. (1979). Fermentation of 1,2-propanediol with 1,2-ethanediol by some genera of *Enterobacteriaceae*, involving coenzyme B12-dependent diol dehydratase. *J. Bacteriol.* *139*, 39–47.

- Tsang, A.W., and Escalante-Semerena, J.C. (1998). CobB, a new member of the SIR2 family of eucaryotic regulatory proteins, is required to compensate for the lack of nicotinate mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase activity in cobT mutants during cobalamin biosynthesis in *Salmonella typhimurium* LT2. *J. Biol. Chem.* 273, 31788–31794.
- Tuerk, C., and Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249, 505–510.
- Wachter, A., Tunc-Ozdemir, M., Grove, B.C., Green, P.J., Shintani, D.K., and Breaker, R.R. (2007). Riboswitch control of gene expression in plants by splicing and alternative 3' end processing of mRNAs. *Plant Cell* 19, 3437–3450.
- Weigand, J.E., Sanchez, M., Gunnesch, E.B., Zeiher, S., Schroeder, R., and Suess, B. (2008). Screening for engineered neomycin riboswitches that control translation initiation. *RNA* 14, 89–97.
- Yamamoto, N., Nakahigashi, K., Nakamichi, T., Yoshino, M., Takai, Y., Touda, Y., Furubayashi, A., Kinjyo, S., Dose, H., Hasegawa, M., et al. (2009). Update on the Keio collection of *Escherichia coli* single-gene deletion mutants. *Mol. Syst. Biol.* 5, 335.
- Zhang, X., and Bremer, H. (1995). Control of the *Escherichia coli* rrnB P1 promoter strength by ppGpp. *J. Biol. Chem.* 270, 11181–11189.

Chapter 4

Characterization of Intermolecular Interactions Involving the Periplasmic Binding Protein BtuF Using Riboswitch-Based Sensors

4.1 AUTHOR'S PREFACE

The third and final project described in this thesis assesses the utility of riboswitches in addressing specific and relevant research questions. This study examines completely unexplored aspects of vitamin B12 transport using the riboswitch sensors described in the previous chapter as the primary experimental methodology. Importantly, the extent and the nature of the information provided by the riboswitch sensors in this study could not have been acquired practically using existing alternative technologies. The specific focus of this chapter is the ABC transporter that is responsible for the uptake of periplasmic vitamin B12 into the cytoplasm for *E. coli* cells. Several important molecular contacts made by the periplasmic component of this system, BtuF, are dissected. The results presented revealing novel insights into the transport mechanism that are also relevant for many related, physiologically important transport systems.

Unlike the two preceding chapters, the work featured in this chapter has not yet been published. The research is, however, presented in the same manner as the previous chapters. A manuscript describing this study is expected to be submitted for publication prior to the date of my thesis defence. I will again be the first author on this manuscript.

I was principally responsible for the conception of this project as well as the experimental design and interpretation. I carried out all experiments presented and have written this chapter without assistance. Dr. Yingfu Li and Dr. Eric Brown once again provided important guidance and discussions relating to this work. Dr. Murray Junop provided a great deal of assistance with experimental design and interpretation relating to the structural biology involved in this project. With my guidance, Seiji Sugiman-Marangos created the images shown Figures 4-2A and 4-3A. The title and authorship of this manuscript will be as follows:

Fowler CC, Sugiman-Marangos S, Junop MS, Brown ED and Li Y. Characterization of Intermolecular Interactions Involving the Periplasmic Binding Protein BtuF Using Riboswitch-Based Sensors.

4.2 ABSTRACT

E. coli employs an intricate systems to transport vitamin B12-related molecules across by both its outer and periplasmic membranes. The periplasmic system is an ABC transporter consisting the cytoplasmic ATPase BtuD, the transmembrane channel-forming protein BtuC and the periplasmic binding protein BtuF that form a BtuC₂D₂F complex. The availability of a remarkable wealth of high resolution structural information for this transporter makes it a useful model for the study of the many similar

bacterial uptake systems. In this report we aim to address the specific functional and mechanistic role of several interesting intermolecular interactions observed in previous structural reports. Specifically, we have investigated the interactions of the periplasmic binding protein BtuF with both its cognate ABC transporter and with its substrate. Our results indicate that two highly conserved BtuF glutamic acid residues play an important but unequal role in transport and that a disruption of these residues disproportionately disrupts uptake of larger substrates. A number of key ligand-binding residues were also investigated and a particularly vital role was noted for residues that form a rigid wall in the binding pocket. Transport activity in this study was assessed in growing cells using a genetically-encoded riboswitch sensor that detects intracellular coenzyme B12 concentrations. The successful and flexible application of these sensors in this report illustrates the potential of such sensors as tools for the study of biological transport.

4.3 INTRODUCTION

Cells face the formidable challenge of acquiring desirable small molecules from the environment, while extruding and maintaining a barrier against potentially harmful compounds. Intricate transport systems that harness energy to transfer specific substrates across biological membranes are crucial to addressing this challenge and are essential to the welfare of all organisms. Altered activity of human transporters has been linked to a range of disease states (Borst and Elferink, 2002). Bacteria, as single-celled organisms that typically interact directly with a harsh environment, also rely heavily on a large

collection of transporters for both uptake and efflux of a wide variety of molecules (Davidson et al., 2008). Many genes of this category are putative or established virulence factors for human pathogens (Eitinger et al., 2011). Additionally, the acquisition of efflux pumps that prevent antibiotics from accessing their cellular targets represents one of the most common routes to antibiotic resistance for clinically relevant pathogens (Lubelski et al., 2007). A detailed understanding of bacterial transport machinery is therefore valuable for developing strategies to combat microbial infections and antibiotic resistance.

E. coli is one of many organisms that have elaborate uptake systems to procure vitamin B12 and closely related molecules from the environment. In an active form, vitamin B12 acts as a cofactor for a range of chemical reactions. For enteric bacteria, vitamin B12 is perhaps most notably employed in reactions essential to the breakdown of secondary energy sources such as ethanolamine and propanediol (Roth et al., 1996). *E. coli*, unlike many of its close relatives, cannot produce vitamin B12 *de novo* and thus it is completely reliant on the activity of its transport system to support its vitamin B12-requiring functions (Lawrence and Roth, 1995). Unlike most metabolites taken up by Gram negative bacteria, vitamin B12 cannot diffuse through porins. Instead it is actively transported across the outer membrane in a process which serves to concentrate this scarce cofactor in the periplasm (Kadner, 1990). Outer membrane transport requires the BtuB receptor as well as the TonB/ExbB/ExbD complex, which exploits the proton motive force to provide the required energy (Postle and Kadner, 2003).

Once in the periplasm, vitamin B12 is actively transported across the cytoplasmic membrane using the BtuCD-F ATP-binding cassette (ABC) transporter. Members of the ABC family of transporters, which can be recognized by their highly conserved ATPase subunits, are found across all phylogeny (Holland and Blight, 1999). Bacterial ABC transporters are commonly employed for both uptake and efflux functions. Import systems in Gram negative bacteria, such as *E. coli* BtuCD-F, are used to acquire useful nutrients that have traversed into the periplasm at the expense of ATP, which is hydrolyzed to energize the directional transport of the substrate. These systems are composed of two cytoplasmic ATPase domains or proteins, two transmembrane channel-forming domains or proteins and a periplasmic binding protein (PBP) (Davidson et al., 2008). The role of the PBP is to specifically bind its designated substrate and deliver it to its cognate ABC transporter at the periplasmic entrance of the transport channel. ATP hydrolysis is then thought to power structural rearrangements that allow the substrate to pass through the channel and for the dissociation of the unloaded PBP. For *E. coli* BtuCD-F, ATPase activity is provided by homodimeric BtuD which forms an extremely stable heterotetramer with the homodimeric transmembrane component BtuC. BtuF, the PBP component, binds vitamin B12 with very high affinity (Cadieux et al., 2002) and has also been shown to form a strong complex with BtuCD, yielding the pentamer BtuC₂D₂F (Hvorup et al., 2007). The unusually stable nature of the interactions amongst the components of the BtuCD-F transporter have facilitated structural studies (Lewinson et al., 2010). As a result, an impressive collection of high resolution structures of this system have been solved including: apo-BtuF (Karpowich et al., 2003), BtuF in complex

with vitamin B12 (cyanocobalamin or CNCbl) (Borths et al., 2002; Karpowich et al., 2003), BtuC₂D₂ (Locher et al., 2002) and, most impressively, BtuC₂D₂F (Hvorup et al., 2007). The availability of this incredible wealth of structural information makes *E. coli* BtuCD-F a useful model for the study of the many related transporter systems.

The study of transporters can be hampered by a lack of suitable methods that can be used to track their activity. Establishing *in vitro* assay systems can be complicated and the conclusions drawn from these experiments might not hold true for living, growing cells. To monitor transport in live cells, one must be able to track the substrate as it enters or exits the cell. We have recently developed a series of genetically encoded intracellular sensors that detect the relative concentrations of an active form of vitamin B12, coenzyme B₁₂ (adenosylcobalamin or AdoCbl) (Fowler et al., 2010). These sensors allow for the fast and reliable detection of AdoCbl concentrations within living *E. coli* cells. Furthermore, they have previously been applied to successfully identify lesions in both the outer membrane and cytoplasmic membrane vitamin B12 transport systems. These sensors therefore represent valuable tools in the study of vitamin B12 transport, as they allow for the functional consequences of perturbations to these systems to be readily assessed within living cells.

The abundant structural data on the BtuCD-F provides an impression of the molecular interactions that take place between the components of this apparatus. However, the role that these interactions play in the transport process has not yet been investigated. In this study we dissect the functional importance of several intermolecular

interactions of special interest that have been observed for the PBP BtuF. Guided by the available structural data, we selected nine BtuF residues that appeared to have a pivotal role in its interaction with either its cognate ABC transporter or with its ligand. A large collection of mutations at these positions were designed and expressed within an *E. coli* BtuF deletion strain. A riboswitch-based coenzyme B12 sensor was then used to reveal the activity of these mutants within living cells. The consequences of select mutations on the growth of *E. coli* under vitamin B12-requiring conditions was also assessed to provide an understanding of the biological impact of the molecular interactions examined. The results presented provide insights into the role of highly conserved salt bridges formed between PBPs and channel-forming proteins and also present a detailed characterization of BtuF's vitamin B12 binding pocket. Furthermore, our findings illustrate the utility of riboswitch sensors in probing the molecular details of biological transport.

4.4 MATERIALS AND METHODS

4.4.1 Strains, plasmids and molecular cloning.

Assays were conducted using the $\Delta btuF$ strain from the Keio *E. coli* single-gene deletion library (Baba et al., 2006). The kanamycin resistance cassette of this strain was removed using FLP recombinase as described (Baba et al., 2006). This strain contains an

in-frame deletion of the *btuF* gene, but is otherwise identical to the BW25113 strain (*rrnB*_{T14} *ΔlacZ*_{WJ16} *hsdR514* *ΔaraBAD*_{AH33} *ΔrhaBAD*_{LD78}) (Datsenko and Wanner, 2000), which serves as the WT strain for Figure 4-1B. Molecular cloning was carried out using *E. coli* NovaBlue cells (Novagen). The previously described pRsBs-ACbl-βgal-MC vector (Fowler et al., 2010) (the medium copy version of the pRsBs-ACbl-βgal-HC plasmid) was used in all instances to monitor vitamin B12 levels. BtuF, WT as well as all described mutants, was cloned into the pBAD18-kan vector (Guzman et al., 1995) using the EcoRI/XbaI restriction sites to produce pBAD18-BtuF. WT BtuF, WT BtuF with a C-terminal His6 tag, as well as the E72 and E202 mutants were created in house using PCR with the appropriate restriction sites built into the primers and standard molecular cloning techniques. All constructs were confirmed by sequencing (Mobix lab, McMaster University). Saturation mutagenesis libraries for residues W85, G88, W196, D242 and R246 were synthesized and sequenced by ATG:biosynthetics (Merzhausen, Germany) and contained a C-terminal His6 tag.

4.4.2 Low throughput reporter assays.

The *ΔbtuF* strain was transformed with the pRsBs-ACbl-βgal-MC vector and single colonies of the sensor-containing strain were selected, made competent and transformed with the indicated pBAD18-BtuF vector. This yielded single colonies of a sensor-containing strain wherein the deleted chromosomal copy of *btuF* was complemented with a plasmid encoded copy of WT or mutant *btuF*. Colonies were

selected and grown overnight at 37°C with shaking in a rich, chemically defined medium (previously described in (Fowler et al., 2010)) supplemented with 40 µg/ml kanamycin and 75 µg/ml ampicillin. Saturated cultures were split out into fresh media (1/1000 dilution) containing either no vitamin B12 or the indicated concentrations of CNCbl, AdoCbl or Cbi (Cbi supplied as cobinamide dicyanide) and 5 µM DMB (all purchased from Sigma-Aldrich) and grown for 6 hours to mid-late log phase. Cells were then assayed for β-galactosidase activity as described below.

β-galactosidase assays were performed using the Galacto-Star assay system (Applied Biosciences), which employs a luminescent substrate. For each sample, 20 µl of cell culture was added to 80 µl of permeabilization buffer (100 mM Na₂HPO₄, 20 mM KCl, 2 mM MgSO₄, 0.8 mg/ml cetrimonium bromide, 0.4 mg/ml sodium deoxycholate, 5.4 µl/ml β-mercaptoethanol) and incubated at room temperature for 30 minutes. Following this, 10 µl of the permeabilization mixture was then added to 100 µl of Galacto-Star substrate, incubated for 1 hour at room temperature. Following this incubation, luminescence was measured using an Envision luminometer (Perkin-Elmer). Raw luminescence values were normalized for cell density using OD₆₀₀ measurements taken with a VERSAmax spectrophotometer (Molecular Devices). Fold repression values were obtained by dividing the calculated reporter activity for each sample grown in the absence of vitamin B₁₂ compounds by that calculated for the parallel sample containing CNCbl, AdoCbl or Cbi. All data shown represent the average of at least three biological replicates, confirmed by at least two independent experiments.

4.4.3 Reporter assays for saturation mutagenesis libraries.

pBAD18-BtuF vectors containing individual BtuF mutants as well as control vectors were transformed into competent *ΔbtuF* cells carrying the pRsBs-ACbl-βgal-MC sensor vector. Individual colonies were selected and grown overnight in 96-well plates that were subsequently used to make permanent glycerol stocks. Each plate contained eight scattered WT BtuF controls, eight scattered empty vector controls, four scattered E202A controls and four scattered samples for each of the 19 mutants being assessed. An example of the plate arrangement is shown in the appendix, chapter 4 supplemental materials, Figure S4-1. Frozen stocks were pinned and grown overnight in 96-well plates containing 250 μl of rich defined media. Overnight cultures were then diluted ~100-fold into 250 μl of fresh media that contained either no vitamin B12 or 50 nM CNCbl and were grown for 5 hours at 37°C shaking. Samples were then assayed for β-galactosidase activity as described in section 4.4.2. Two technical replicates of this process (from pinning through to assay completion) were performed, yielding eight total fold-repression values per mutant. The highest and lowest fold-repression values for each sample were automatically disregarded and the data presented represents the mean of the remaining six values.

4.4.4 Ethanolamine utilization assays.

For all strains tested, overnight cultures were grown in rich, chemically defined media. A 10 µl aliquot of these saturated overnight cultures was then added to 4 ml of a growth medium in which ethanolamine was the sole nitrogen source. This medium consisted of glycerol (0.2%), ethanolamine (0.1%), KH₂PO₄ (3 mg/ml), K₂HPO₄ (7 mg/ml), MgSO₄ (2 mM), CaCl₂ (100 µM) and the indicated vitamin B₁₂-related molecules at a concentration of 50 nM (5 µM DMB was also added to the Cbi samples). Prior to sterilization and the addition of vitamin B₁₂ compounds, the pH of the media was adjusted to ~7.2 using HCl. At the indicated time points, 200 µl of these cultures was removed and the OD₆₀₀ was measured using a VERSAmax spectrophotometer (Molecular Devices).

4.5 RESULTS

4.5.1 An assay system to detect the activity of BtuF mutants

We previously described the construction and characterization of plasmid-encoded coenzyme B₁₂ (AdoCbl) sensors (Fowler et al., 2010). These vectors exploit the remarkable ligand recognition properties of an AdoCbl-responsive riboswitch, which is used to regulate the expression of co-transcribed reporter proteins. The reporter activity of *E. coli* cells carrying these vectors is tightly tied to the cytoplasmic concentration of AdoCbl, such that higher AdoCbl concentrations result in stronger repression of reporter

expression and thus lower reporter values. In order to track vitamin B₁₂ transport in the present study, *E. coli* strains carrying the medium copy, β -galactosidase version of the sensor plasmid were employed. Cells were grown in a medium that contained vitamin B₁₂ molecules and were compared to parallel samples grown in a medium lacking vitamin B₁₂ or a suitable precursor. The extent of reporter gene repression observed as a result of adding vitamin B₁₂ is proportional to the amount of substrate imported into the cytoplasm for that sample. AdoCbl precursors, including both CNCbl and Cbi are quickly converted to AdoCbl by metabolic enzymes following transport, allowing for the uptake of these molecules to be easily measured as well. The data are expressed as fold repression values, such that a value of one would indicate that no detectable AdoCbl is present in the cytoplasm. It is worth noting that, when present at sufficiently high media concentrations, low levels of cytoplasmic AdoCbl can be detected in strains with genetic deletions to components of the BtuCD-F transporter (Bassford and Kadner, 1977; Fowler et al., 2010). This can likely be attributed to the accumulation of vitamin B₁₂ molecules in the periplasm, coupled with a very low level of incidental diffusion into the cytoplasm.

In order to track the effects of BtuF mutations on its cellular activity, we sought to compliment the $\Delta btuF$ strain with a plasmid encoded copy of *btuF*. A plasmid complementation system was advantageous in this instance as it would allow us to easily assess the activity of a large library of BtuF mutants. BtuF was cloned into the pBAD18 vector (Guzman et al., 1995) under the control of an arabinose-inducible promoter and transformed into the $\Delta btuF$ strain. Both the wild-type (WT) and the C-terminal his₆-tagged version of BtuF fully complimented the chromosomal deletion in the absence of

arabinose (Figure 4-1B). The three test strains showed the same level of transport for three different vitamin B₁₂ compounds at both sub-saturating and saturating concentrations (2 nM and 50 nM). As expected, the deletion strain carrying the empty vector control showed low, yet detectable, levels of cytoplasmic coenzyme B₁₂. Consistent with the hypothesis that this is due to passive diffusion across the cytoplasmic membrane, the greatest level of cytoplasmic coenzyme B₁₂ was observed when the smallest substrate, cobinamide (Cbi), was supplied. To our knowledge, this marks the first direct evidence that BtuF transports Cbi and AdoCbl. While this finding is not surprising, there is precedent for distinct PBPs delivering different substrates to the same ABC transporter (Higgins and Ames, 1981). The addition of arabinose to induce BtuF expression led to a mild defect in transport (data not shown). This is not surprising, as an excess of PBPs has been previously observed to obstruct transport for similar systems (Prossnitz et al., 1989).

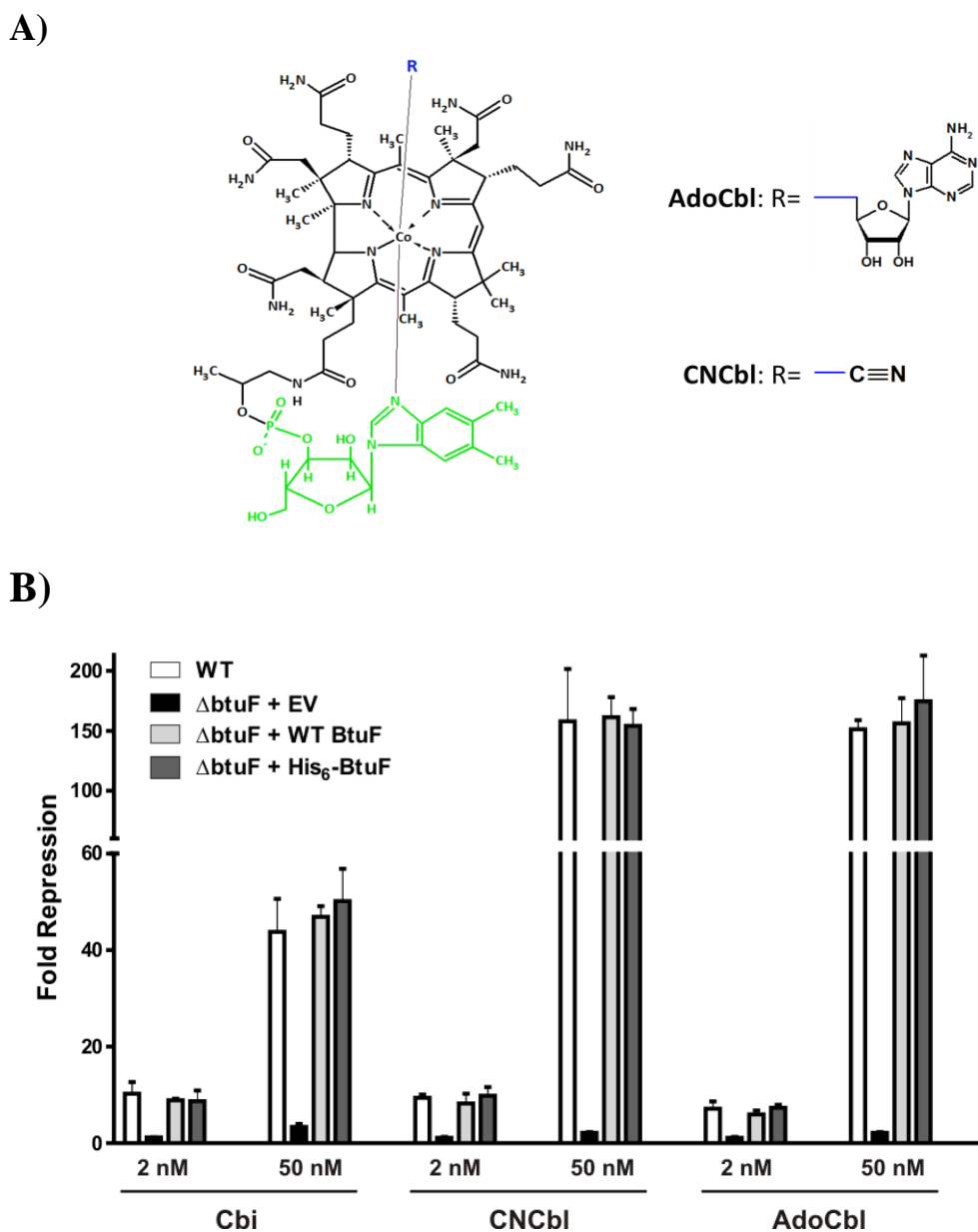


Figure 4-1 The transport activity of chromosomal and plasmid-encoded BtuF. (A) Chemical structures of three BtuF substrates, adenosylcobalamin (AdoCbl), cyanocobalamin (CNCbl) and cobinamide (Cbi). Cobinamide is supplied as dicyanocobinamide, which lacks the nucleotide loop (shown in green) and which contains a cyano group for both the upper and lower axial ligands for cobalt. (B) Assay comparing the vitamin B12 transport activity of a WT strain to that of a $\Delta btuF$ strain carrying either

an empty vector control (EV), plasmid-encoded WT BtuF (WT) or plasmid-encoded WT BtuF with a C-terminal His6 tag (His6-BtuF). Transport activity is compared using 2 nM and 50 nM concentrations of the compounds described in (A). Data are expressed as mean fold repression values (described in the text) and error bars represent standard deviation.

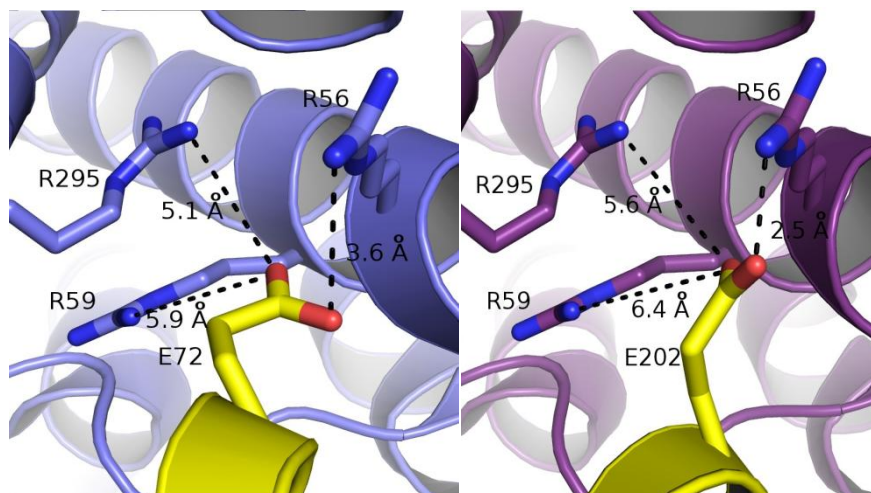
4.5.2 Assessing the activity of BtuC interaction mutants

We were interested in exploring the nature of BtuF's interaction with BtuC, the transmembrane channel-forming component of its cognate ABC transporter. The structure of BtuF features N-terminal and C-terminal lobes with an intervening ligand-binding pocket (Borths et al., 2002; Karpowich et al., 2003). Each lobe of BtuF has been observed to interact with one monomer of the BtuC homodimer, which positions the bound ligand over the translocation pathway (Hvorup et al., 2007). Two highly conserved glutamic acid residues, one in each lobe of BtuF, have been previously proposed (Borths et al., 2002) and observed (Hvorup et al., 2007) to interact with the same three arginine residues on either monomer of BtuC (Figure 4-2A). This arrangement is also exploited by a number of other ABC importers. The significance of these interactions has been explored previously for the *E. coli* FecCDE-B iron transporter, where disruptions were observed to decrease iron uptake (Braun and Herrmann, 2007).

We wanted to determine the functional importance of these glutamic acid-arginine interactions for the BtuCD-F transporter and to compare these observations to those previously made for parallel residues of iron transporters. A series of vectors were

constructed that contain BtuF with single mutations to the glutamic acid residues at both position 72 and 202. For each of these residues, aspartic acid, alanine and lysine substitutions were made and these six mutant constructs were assessed for their ability to transport vitamin B12 using the riboswitch sensors (Figure 4-2B). As was observed for the iron uptake system, each of the two BtuF glutamic acid residues clearly has a significant functional role (Braun and Herrmann, 2007). Interestingly, even very conservative mutations to aspartic acid resulted in a detectable loss of activity for both residues. Aspartic acid mutations, which shorten the side chain by a single carbon unit, result in the carboxyl group inserting more shallowly into the positively charged BtuC pocket. While it is conceivable that R56 of BtuC could rotate to compensate for this, any interaction with the more distant arginines (R59 and R295) is likely lost. Each of the alanine mutations, which completely negate these electrostatic interactions, resulted in a further decrease in vitamin B12 uptake. Mutations to lysine, which are expected to be highly repulsive, completely abolished transporter activity.

A)



B)

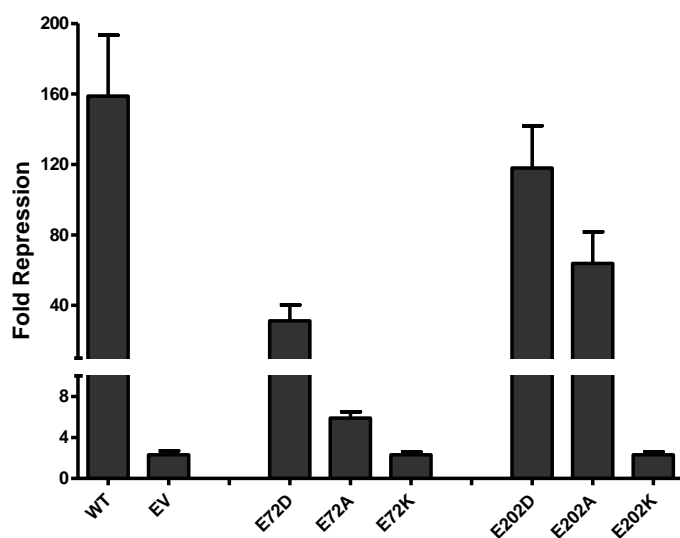


Figure 4-2 The transport activity of BtuF E72 and E202 mutants. (A) Structural context of the two pertinent glutamic acid residues: E72 (left) and E202 (right). BtuF is shown in yellow, the BtuC monomer than interacts with E72 is shown in blue and the BtuC monomer than interacts with E202 is shown in purple. (B) Assay examining the transport activity of BtuF carrying single mutations to BtuF E72 and E202 to aspartic acid, alanine and lysine. Assay was conducted using 50 nM CNCbl. Data are expressed

as mean fold repression values (described in section 4.5.1) and error bars represent standard deviation.

Perhaps the most notable finding was that mutations to E72 were clearly more disruptive than those to E202. This is particularly interesting in light of the fact that each of these residues is interacting with an identical pocket on either monomer of the BtuC homodimer. The static image provided by the BtuCD-F structure shows that the atomic distances between E202 and R56 point to a slightly stronger salt bridge than is observed between R56 and E72 (Hvorup et al., 2007). While R59 and R295 are slightly closer to E72, the observed distances indicate that the resulting energetic contribution would be negligible. One possible explanation for the increased importance of the interactions involving E72 could be that the surrounding interactions are weaker at the N-terminal lobe and thus less able to compensate for the loss of the salt bridge. Conversely, the N-terminal lobe could hypothetically be the more strongly interacting lobe and could bind first to anchor BtuF to the surface of BtuC. These possibilities are discussed further in section 4.6. Interestingly, mutations to the parallel glutamic acid at the N-terminal lobe of FecB were also observed to be more disruptive than those to the corresponding C-terminal residue for the FecCDE-B transporter (Braun and Herrmann, 2007).

4.5.3 Assessing the activity of mutations to the BtuF ligand binding pocket

Two separate groups have published high resolution crystal structures of BtuF in complex with vitamin B12 that are essentially identical, showing a root mean standard deviation (RMSD) of ~ 0.6 Å, well within the error of resolution (Borths et al., 2002;

Karpowich et al., 2003). These structures show vitamin B12 bound between the two lobes of BtuF by means of a many direct and water-mediated interactions between BtuF and the perimeter of the ligand. While a great deal of vitamin B12's surface area is buried within the binding pocket, a significant portion of the molecule is also accessible. The binding pocket generally appears to be quite flexible, apparently relying on a loose grasp on a large area of the ligand for its high affinity ($K_d \sim 15$ nM) (Cadieux et al., 2002). This arrangement is consistent with the premise that BtuF undergoes relatively subtle structural changes when binding and releasing its cargo. This hypothesis is also supported by the observation that BtuF makes considerably fewer direct contacts with its ligand has been observed for vitamin B12-utilizing enzymes (Karpowich et al., 2003).

General shape complementarity appears to be the primary source of BtuF's high affinity for vitamin B12. This, coupled with the large amount of buried surface area, suggests that individual interactions with the ligand could potentially be interrupted without much effect on BtuF activity. However, the considerable degree of evolutionary conservation observed for BtuF residues that contact vitamin B12 contradict this sentiment (Karpowich et al., 2003). We set out to explore the nature of the BtuF binding pocket and to assess how these molecular interactions contribute to overall transport activity. Guided by structural data, we selected five highly conserved amino acids within the BtuF binding pocket with diverse roles in ligand recognition: W85, G88, W196, D242 and R246 (Figure 4-3A). For each of these residues, saturation mutagenesis libraries were constructed that consisted of single mutations to each of the 19 possible alternatives. The resulting collection of 95 mutants was designed to provide an intricate perspective on

the chemical requirements and molecular interactions required for vitamin B12 recognition and transport.

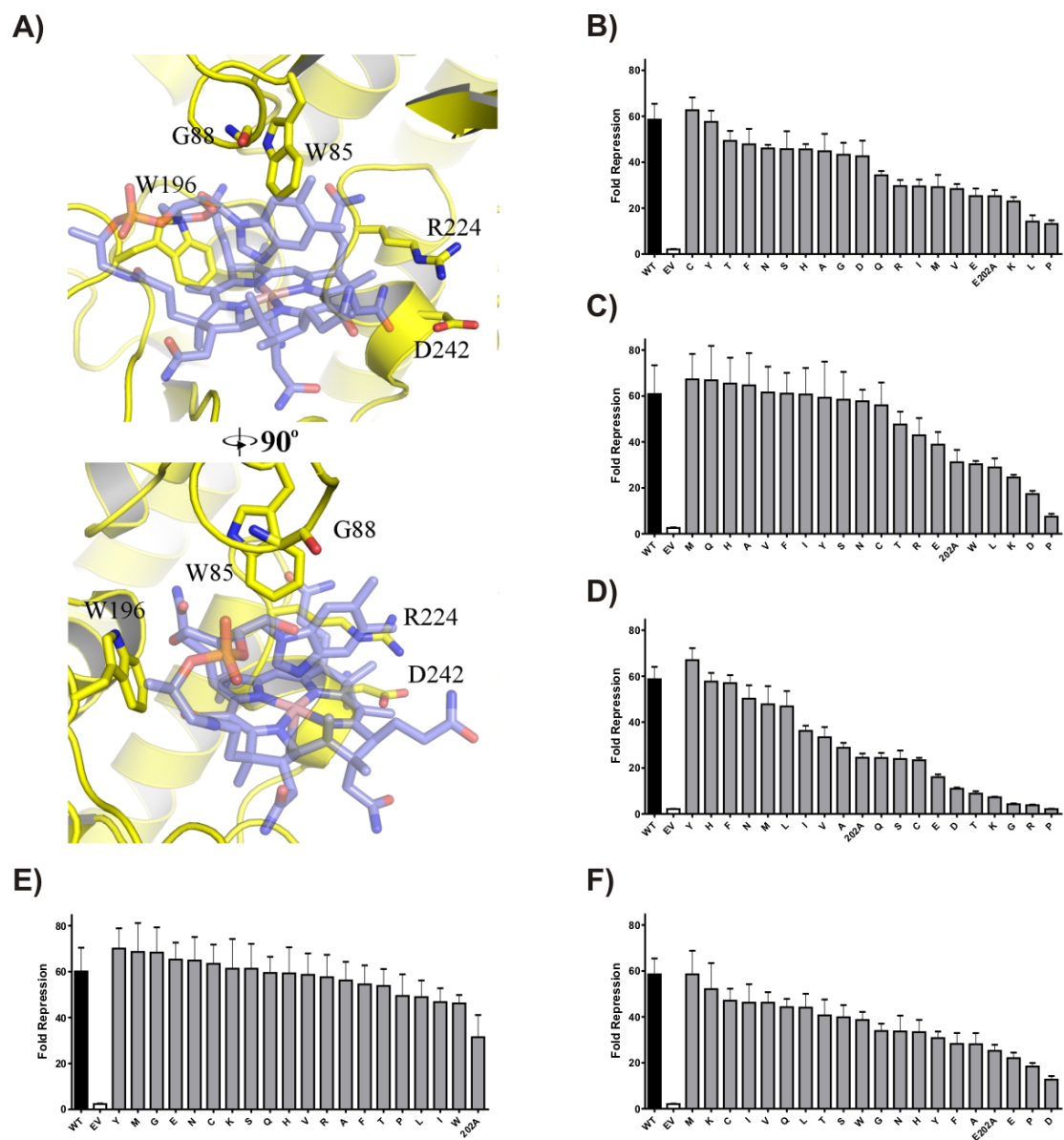


Figure 4-3 The transport activity of BtuF ligand binding mutants. (A) Structural context of the five BtuF residues involved in vitamin B12 binding investigated in this study. Vitamin B12 is shown in purple and BtuF in yellow. Panels B-F show the results of assays measuring the transport activity of saturation mutagenesis libraries for each of the residues shown in panel A. Mutants are arranged left to right according to activity.

(B) W85. (C) G88. (D) W196. (E) D242. (F) R246. The assays were conducted using 50 nM CNCbl. The data are expressed as mean fold repression values and error bars represent standard deviation. BtuF mutant E202A was included as an additional control for each of the libraries.

Mutant libraries for each residue were transformed into the $\Delta btuF$ strain containing the sensor and several clones of each mutant were grown in 96-well plates. Also included on each plate were several replicates of the WT and empty vector controls as well as the E202A mutant, which provided a control for a partial impairment of transport activity. Libraries were assayed for transport activity using the genetically encoded AdoCbl sensors in a relatively high-throughput, 96-well plate format (Figure 4-3B-F).

Other than D242, each of the positions tested yielded several mutations that demonstrated an obvious drop in activity, affirming the importance of these residues. With few exceptions, however, these mutants maintained a significant level of activity, indicating that none of these residues has an essential function in transport. Amongst the positions tested were two tryptophan residues that provide extensive van der Waals contact with the ligand: W85 (Figure 4-3B) and 196 (Figure 4-3D). W85 contacts the DMB moiety of vitamin B12's lower axial ligand. A simple loss of these interactions appears to have a very minor impact of activity, however introduction of small hydrophobic or charged residues at this position are more disruptive. Mutations to W196, which contacts hydrophobic extremities emerging from the central corrin ring, showed the most drastic effects of any position tested. While the beneficial contacts made by this

residue provide valuable binding energy, the more important function of W196 might be its role as an inflexible wall of the binding pocket that holds the ligand in a beneficial location and orientation. While substitutions to similar amino acids such as tyrosine and phenylalanine maintain activity, smaller amino acids, which would not maintain this rigid hold on the ligand, are less active. Due to its position at an inflexible area of the binding pocket it is not surprising that residues expected to clash with the surroundings, such as the charged residues, are particularly poorly tolerated.

G88 (Figure 4-3C) was the only residue selected that does directly interact with vitamin B12 in the published structures. G88 appears to be conserved because its small size allows it to fit into a crowded area near the binding interface. Initially, we were surprised to observe that many substitutions at this position are well tolerated. Upon closer inspection of the binding pocket, we noted a nearby compartment into which many residues could rotate to avoid steric clashes. Mutations that do not permit a simple rotation into this pocket are therefore less active. This includes residues that are not accommodated well in this pocket, such as certain charged residues, and those that are less flexible to rotate into this pocket, such as leucine.

The final two residues selected, D242 and R246 (Figures 4-3E & 4-3F), each form hydrogen bonds with a propionamide side chain from the corrin ring of vitamin B12. Interestingly, while D242 can be mutated freely without a significant effect on BtuF function, most R246 mutants were less active than WT. The lack of an effect for the D242 mutants indicates that the simple loss of this hydrogen bond is not detrimental

under the conditions tested. This is also apparent from the R246 data, where a methionine mutation that negates this hydrogen bond is completely tolerated. While the D242-vitamin B12 interaction consists of a flexible hydrogen bond, R246, much like W196, defines part of a wall within the binding pocket. It therefore appears that the role of R246 in stabilizing and positioning the ligand is more important than its contributions to the overall binding energy. Methionine's ability to replace arginine at this position strongly supports this notion, since its side chain can adopt a similar conformation to that observed for arginine, thereby maintaining the integrity binding pocket's edge. Overall, the data observed support the hypothesis that the primary role of the conserved molecular interactions between vitamin B12 and BtuF in CNCbl transport is to maintain the ligand in a secure orientation that is appropriate for substrate release and transport. Mutations that result in clashes with this orientation or that negate important contacts that fix the ligand in such a position therefore disrupt BtuF activity.

4.5.4 Dose-response of BtuF mutants

The above characterization of BtuF's interactions with BtuC and vitamin B₁₂ was conducted using CNCbl at concentrations that are saturating for WT BtuF. To gain further perspective on the effects of the described mutations, the activity of representative mutants for each of the tested residues was assayed over a broad range of CNCbl concentrations (Figure 4-4). Remarkably, only the highly impaired E72A mutant shows a detectable loss of activity at the low end of the responsive concentration range. This is in

strong agreement with the hypothesis stated above that the principal disruption caused by the vitamin B12 binding site mutants is to ligand orientation and transport rather than to overall binding affinity. At higher CNCbl concentrations, transport defects become obvious for each of the mutants other than D242; the activity of each mutant at this high concentration range is consistent with the data presented in Figures 4-2 and 4-3. A plausible explanation for this trend is that at higher CNCbl concentrations, BtuF-ligand binding ceases to be the rate limiting step of the transport process, allowing for defects at other stages of transport to be observed. CNCbl concentrations above ~25 nM failed to yield further increases in cytoplasmic AdoCbl concentrations for any of the strains tested. This is explained by previous experiments that have shown that BtuB, the receptor responsible for transport of vitamin B12 across the outer membrane, becomes saturated around this concentration range (Bradbeer and Gudmundsdottir, 1990).

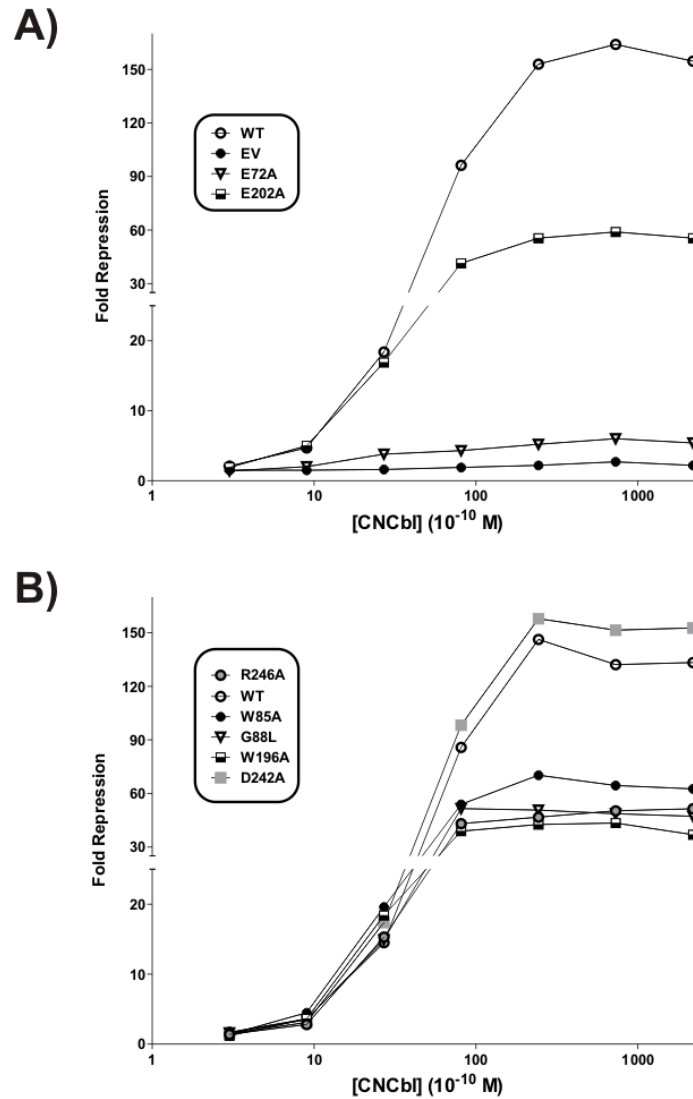


Figure 4-4 The transport activity of BtuF variants as a function of CNCbl concentration. Assay measuring the transport activity of strains carrying various BtuF constructs over a range of CNCbl concentrations. (A) BtuC binding mutants. (B) Vitamin B12 binding mutants. Transport activity is expressed as a mean fold repression value.

4.5.5 Substrate-specific transport defects of BtuF mutants

As observed in Figure 4-1B, BtuF recognizes and facilitates the transport a number related ligands. While its substrates are all expected to contain a common corrin ring, cobalt's lower and upper axial ligands, which can account for more than 40% of the mass of the molecule, are variable. The importance of the molecular interactions under investigation in this report might therefore vary greatly between different substrates. To investigate this possibility, we assessed the transport activity of representative mutants for each of the nine pertinent residues using three different substrates: Cbi, CNCbl and AdoCbl (Figure 4-5). Relative to CNCbl, Cbi lacks the DMB moiety that acts as cobalt's lower axial ligand as well as the nucleotide loop appendage that connects DMB back to the corrin ring. In AdoCbl, the small cyano group that acts as cobalt's upper axial ligand in CNCbl is replaced by a 5'-deoxyadenosyl group (see figure 4-1A).

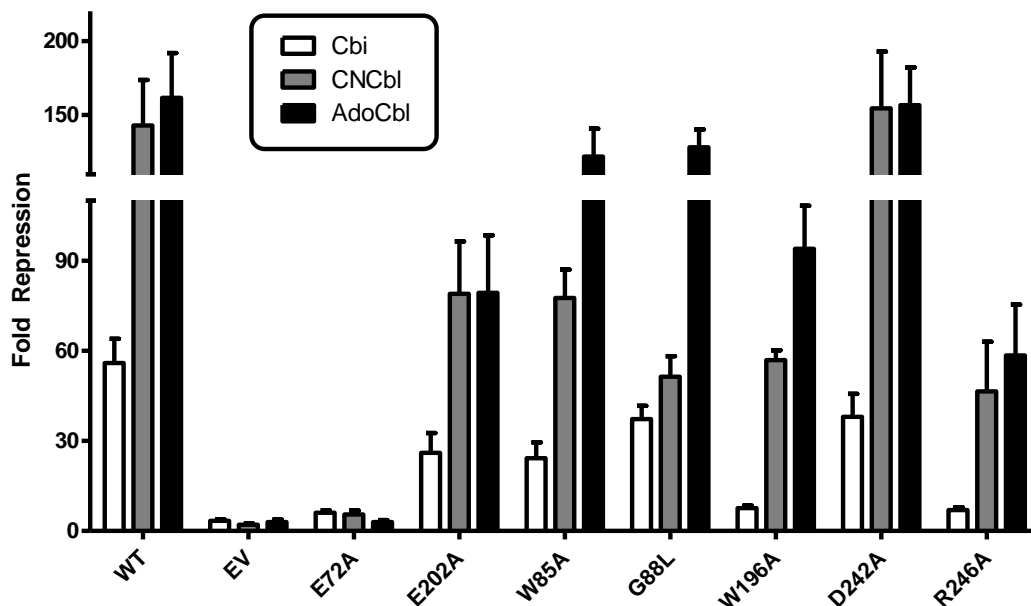


Figure 4-4 The transport activity of BtuF variants using different substrates. Assay measuring the transport activity of strains carrying various BtuF constructs using either 50 nM Cbi and 5 μ M DMB, 50 nM CNCbl or 50 nM AdoCbl. Transport activity is expressed as a fold repression value and error bars indicate the standard deviation.

In general, defects in CNCbl transport for ligand binding mutants were at least partially negated by using AdoCbl as a ligand. By contrast, a significant difference in the transport of these two substrates was not observed for D242A and the WT control. This indicates that the adenosyl moiety might be recognized by BtuF, or at least that this bulky group restricts ligand mobility, thereby helping to overcome the loss of stabilizing interactions. Transport of Cbi was more drastically impaired than that of CNCbl for many ligand binding mutants. This is particularly apparent for the two residues predicted to play a key role in defining binding pocket wall, W196 and R246. Interestingly, the

D242 mutant, which retains full activity for AdoCbl and CNCbl transport, shows decreased Cbi uptake compared to the WT. This might indicate that the loss of binding energy associated with disrupting individual BtuF-vitamin B12 molecular interactions becomes functionally relevant for Cbi. This is not surprising since Cbi is expected to make significantly less contact with the binding pocket than the other substrates.

Intriguingly, higher levels of cytoplasmic AdoCbl were detected for the E72A mutant when grown in media containing Cbi compared to when AdoCbl is provided. This is in stark contrast to each of the other BtuF constructs. Given the position and apparent function of E72, it is highly unlikely that this effect is due to differences in binding these two substrates. We hypothesize that the loss of this salt bridge affects the BtuC-induced spreading of the two BtuF lobes that is expected to trigger ligand release and transport. Alternatively, this phenomenon may be due to a subtle disruption of the alignment of the BtuC transport channel and the ligand binding pocket of BtuF. In either scenario, smaller substrates would presumably be transported more efficiently.

4.5.6 The effects of BtuF mutations on coenzyme B₁₂-dependent growth

E. coli cells require AdoCbl in order to carry out the first step in the breakdown of ethanolamine into useful sources of carbon, nitrogen and energy (Scarlett and Turner, 1976). In certain nutrient-poor environments, ethanolamine utilization might be extremely valuable and it has been proposed that this function could be the evolutionary driving force that maintains *E. coli*'s elaborate vitamin B12 uptake systems (Roth et al.,

1996). We wanted to investigate the consequences of the transport defects observed in previous sections of this study on ethanolamine-dependent growth. This was intended to provide information on the biological significance of the pertinent intermolecular interactions. Cells expressing the assortment of BtuF variants examined in Figures 4-4 and 4-5 were therefore assessed for their ability to grow in media in which ethanolamine was the only nitrogen source. Parallel samples were tested that contained either 50 nM Cbi and 5 μ M DMB or 50 nM AdoCbl (Figure 4-6). No growth was observed for samples grown in the absence of vitamin B12 molecules (data not shown). CNCbl was not assessed here because it is a potent competitive inhibitor of ethanolamine ammonia lyase, the enzyme for which AdoCbl serves as a cofactor (Blackwell and Turner, 1978).

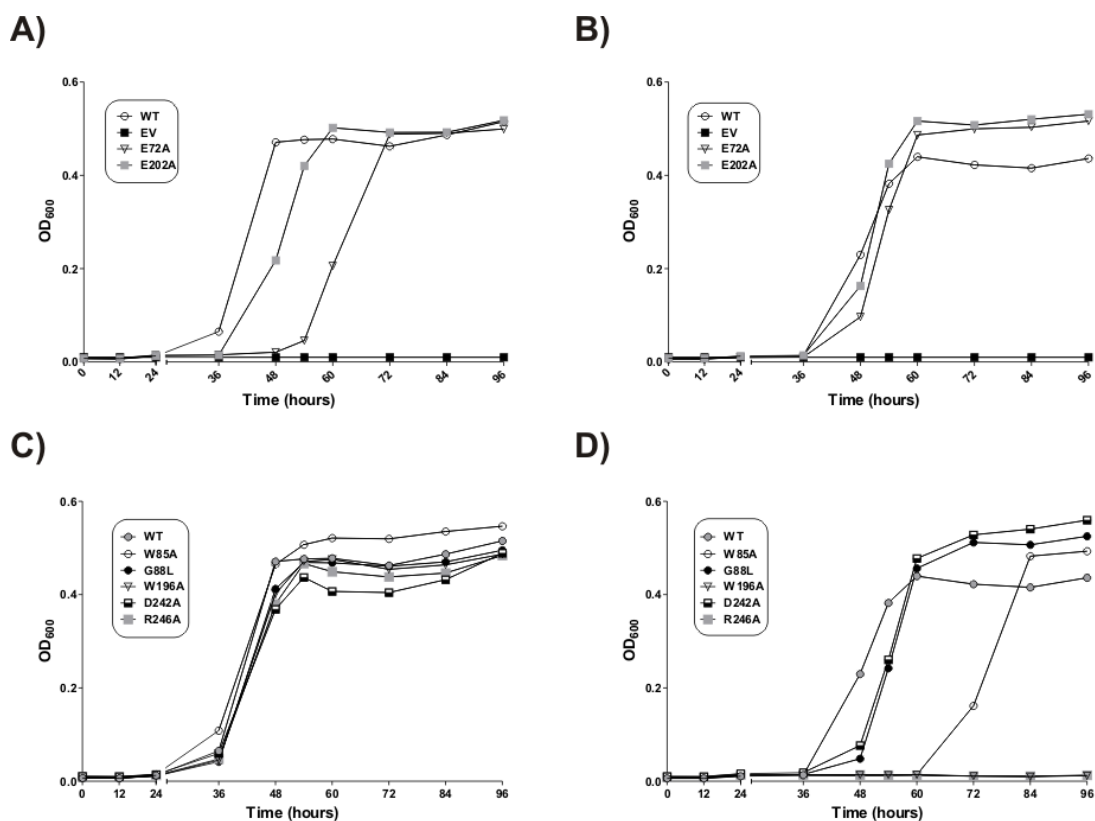


Figure 4-6 AdoCbl-dependent growth of BtuF mutants. Growth assays in a medium where ethanolamine was the sole nitrogen source. Under these conditions AdoCbl is required for growth. Strains expressing the indicated BtuF constructs were grown in the presence of either AdoCbl or the AdoCbl precursors Cbi and DMB. (A) BtuC binding mutants and controls grown with 50 nM AdoCbl. (B) BtuC binding mutants and controls grown with 50 nM Cbi. (C) Vitamin B12 binding mutants grown with 50 nM AdoCbl. (D) Vitamin B12 binding mutants grown with 50 nM Cbi. The WT control data is included as a reference for both the BtuC and vitamin B12 binding mutant graphs. Parallel samples for each strain were grown in media lacking vitamin B12 molecules. No growth was observed under these conditions (data not shown).

Mutant constructs displayed either WT levels of growth, an extended lag phase, or did not grow depending on the mutant and the conditions tested. Drastic changes in growth rate were not observed and all viable samples ultimately grew to approximately the same density. Relative to WT, the BtuC binding mutants displayed a significant

delay in growth in the presence of AdoCbl (Figure 4-6A) and a very slight delay in the presence of Cbi (Figure 4-6B). This difference is particularly evident for E72A, which reaches maximal density at least 12 hours earlier when grown in Cbi than when grown in AdoCbl. This confirms the earlier observation that these mutants transport Cbi more efficiently than AdoCbl, and further supports a hypothetical role for E72 in aligning the binding pocket with the transport channel or in inducing BtuF conformational changes. The more severe growth defects observed for E202A when supplied with AdoCbl compared to Cbi indicates that nature of the disruptions caused by this mutation could be similar to those caused by E72A.

None of the vitamin B12 binding mutants showed a change in growth when supplied with AdoCbl (Figure 4-6C). By contrast, all five mutants tested displayed growth defects in the Cbi-containing medium and strains carrying two of these mutants, W196A and R246A, were not viable under these conditions (Figure 4-6D). This is consistent with the results observed in section 4.5.5, where vitamin B12 binding mutants were observed to transport Cbi relatively poorly and AdoCbl relatively well. The previous transport activity assay also showed that W196A and R246A are far more impaired for Cbi transport than the uptake of other vitamin B12 binding mutants; this observation is confirmed by their lack of growth in Figure 4-6D.

A comparison of the performance of BtuC binding and vitamin B12 binding mutants presents some findings not anticipated based on the transport data in Figure 4-5. For example, E72A experienced only a minor delay in growth in ethanolamine when

provided with Cbi, whereas W196A and R246A did not grow under these conditions despite showing similar levels of Cbi transport in Figure 4-5. Similarly, E202A grew more quickly in ethanolamine medium supplemented with Cbi than a number of vitamin B12 binding mutants that showed similar or superior Cbi transport in Figure 4-5. This effect could be explained by the different nature of the defects associated with these two classes of mutants in Cbi transport. While BtuC binding mutants suffer from a decreased rate of transport, over time in slowly growing cells they could potentially accumulate higher cytoplasmic vitamin B12 concentrations than vitamin B12 binding mutants, which have a higher initial rate of transport but a decreased affinity for Cbi (see discussion).

4.6 DISCUSSION

In this report we have thoroughly examined the functional significance of a number of highly conserved residues of the vitamin B12 PBP BtuF. The resulting data provide mechanistic insights into the interaction of BtuF with both its ligand and its cognate ABC transporter. With regard to its interaction with the channel-forming protein BtuC, we investigated two BtuF glutamic acid residues, each of which contacts three arginine residues on either monomer of the BtuC dimer. Mutational analysis demonstrated that the interactions involving both of these residues are important but not essential for BtuF activity, which is in agreement with a previous study that focused on analogous residues for an iron siderophore PBP (Braun and Herrmann, 2007). As

expected, our data support a role for these interactions in the efficiently associating with BtuC to carry out productive substrate transport, rather than an effect on ligand binding. This is supported by the undisrupted activity of E202A at low CNCbl concentrations and the relatively high levels of transport detected for these mutants for ligands predicted to form weaker interactions with BtuF.

Both transport assays and ethanolamine growth assays indicate that mutations to these residues result in a less severe defect for Cbi uptake than for the uptake of the larger substrates. We hypothesize that this is due to one of the following scenarios: (i) a subtle change in the alignment of BtuF and BtuC upon disruption of these salt bridges that disrupts ligand access to transport channel or (ii) these glutamic acid residues play an important role in the BtuC-induced spreading of the two lobes of BtuF that is required for ligand release. The increased activity of these mutants with smaller substrates is not explained by a simple role in maintaining suitable binding rates or affinity between BtuC and BtuF. A more intricate functional role for these glutamic acid residues would explain their conservation amongst several classes of PBP.

We observed a consistent and significant difference in the defects associated with mutations to E72 and E202. This speaks to the previously documented asymmetry of this transport complex (Hvorup et al., 2007). The increased importance of the glutamic acid at position 72 implies that other interactions at the N-terminal lobe are less able to compensate for the loss of this salt bridge. In the context of the putative roles that we predict for these residues, it appears that the C-terminal lobe may have additional factors

that help maintain the structural rearrangements or alignment responsibilities. A serine residue at position 200, which also interacts with the arginine-rich BtuC pocket, is a likely candidate for this; a similar residue is not present at the N-terminal lobe. Interestingly, a similar disparity between the conserved N-terminal and C-terminal glutamic acids was previously observed in an iron siderophore PBP. The possibility that the uneven contributions of these interactions is mechanistically significant is an intriguing one.

Our examination of several key residues within the BtuF's ligand binding pocket also revealed several interesting, mechanistically relevant observations. Overall, these data support a scenario in which the large contact area between BtuF and the ligand is able to compensate for the loss of individual molecular interactions to maintain the requisite binding affinity for the transport of CNCbl and AdoCbl. However, the disruption of molecular interactions that provide rigid barriers in the binding pocket are detrimental to transport activity due to the "loose" nature of BtuF's grasp on these molecules. Cbi makes considerably less contact with the binding pocket and therefore the loss of key molecular interactions appears to yield a significant drop in binding affinity. Mutations to key residues that maintain a beneficial ligand orientation affect Cbi transport even more drastically than they affect the transport of larger ligands. The transport of Cbi and similar smaller ligands is therefore likely to be a strong driving force in maintaining the high degree of conservation within the ligand binding site. While *E. coli* lacks the majority of the vitamin B12 biosynthetic pathway, it has maintained the ability to synthesize the nucleotide loop (Lawrence and Roth, 1995). Precursors such as

Cbi must therefore be both environmentally available and physiologically important. Periplasmic concentrations of BtuF ligands are destined to fall within the limited range dictated by its outer membrane transport system. While evolution would drive BtuF to efficiently capture its ligands in this concentration range, it must also avoid binding them so tightly as to inhibit their subsequent release. BtuF has presumably evolved to exhibit the minimum binding energy required to efficiently capture its lowest affinity ligands. It is therefore not surprising the mutations that decrease this affinity further have a detrimental effect on transport activity of these substrates.

When provided with Cbi, the two classes of mutations investigated in this study behaved differently in transport activity assays than in ethanolamine growth assays. Specifically, BtuC binding mutants performed relatively well in the ethanolamine growth assays and relatively poorly in the transport activity assays, whereas the reverse was true for the ligand-binding mutants. We propose the following hypothesis to explain this apparent discrepancy: The effect of the BtuC binding mutants is to decrease the rate of transport, however over time slowly growing cells can still accumulate high levels of Cbi and thus AdoCbl. While this may lead to growth delays, ultimately sufficient AdoCbl will be accumulated to support ethanolamine breakdown and growth. Ligand binding mutants significantly affect BtuF's affinity for Cbi however the rate of transport at high Cbi concentrations may be similar or superior to that of the BtuC binding mutants. However, as cytoplasmic vitamin B12 concentrations increase, the expression of the outer membrane transporter is shut down. The resulting decrease in periplasmic Cbi

concentrations has a greater effect on ligand binding mutants and thus lower levels of AdoCbl accumulate and ethanalamine utilization is more strongly affected.

At present we cannot conclusively rule out the prospect of altered expression, folding or localization of the mutants we have created on their observed activity. However, several lines of evidence suggest that these artefacts are not prevalent in our mutant library. First, the library was carefully designed to avoid selecting residues that appeared to play an important structural role. Second, the different mutations created for each residue tested generally produce results that can be explained from a structural standpoint. The BtuC-binding mutants follow the expected activity trend (D>A>K) and more conservative mutations to vitamin B12-binding residues were generally observed to be more active. Finally, the mutants examined generally showed ligand-specific effects. This is much more consistent with a specific transport defect than with a defect related to expression, localization or folding.

Further investigation of the BtuF mutants examined here will be instrumental in evaluating the hypotheses prompted by this study. Of particular interest will be to study the interaction of these mutants with ligands and with BtuCD *in vitro* using purified components. An impressive study was recently published in which the dynamics of the interaction between BtuF and BtuCD were uncovered (Lewinson et al., 2010). The effects of adding substrates or ATP on complex formation were also assessed in this study. Applying similar methodology using the mutants created here should provide a

means to test the hypotheses proposed relating to the mechanistic contributions of these conserved molecular interactions.

In this study, vitamin B12 transport activity was monitored in living cells using genetically encoded riboswitch sensors. The sensors provided precise and consistent measurements across a range of biologically-relevant AdoCbl concentrations. These assays are very user-friendly, inexpensive and do not require the use of radioactive substrates. Another significant advantage of employing riboswitch-based sensors to track transport is their compatibility with higher throughput experiments. This allowed us assess the activity of a very large number of mutants and to examine a range of different substrates and substrate concentrations with several of these variants. Naturally evolved riboswitches that recognize a range of important metabolites (Barrick and Breaker, 2007) have now been characterized and methods to engineer custom riboswitches have also been established (Topp and Gallivan, 2010). This presents the enticing prospect of applying similar riboswitch sensors toward the study of other transporter systems of interest.

4.7 REFERENCES

- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2, 2006 0008.
- Barrick, J.E., and Breaker, R.R. (2007). The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol* 8, R239.

- Bassford, P.J., Jr., and Kadner, R.J. (1977). Genetic analysis of components involved in vitamin B12 uptake in *Escherichia coli*. *J Bacteriol* 132, 796-805.
- Blackwell, C.M., and Turner, J.M. (1978). Microbial metabolism of amino alcohols. Purification and properties of coenzyme B12-dependent ethanolamine ammonia-lyase of *Escherichia coli*. *Biochem J* 175, 555-563.
- Borst, P., and Elferink, R.O. (2002). Mammalian ABC transporters in health and disease. *Annu Rev Biochem* 71, 537-592.
- Borths, E.L., Locher, K.P., Lee, A.T., and Rees, D.C. (2002). The structure of *Escherichia coli* BtuF and binding to its cognate ATP binding cassette transporter. *Proc Natl Acad Sci U S A* 99, 16642-16647.
- Bradbeer, C., and Gudmundsdottir, A. (1990). Interdependence of calcium and cobalamin binding by wild-type and mutant BtuB protein in the outer membrane of *Escherichia coli*. *J Bacteriol* 172, 4919-4926.
- Braun, V., and Herrmann, C. (2007). Docking of the periplasmic FecB binding protein to the FecCD transmembrane proteins in the ferric citrate transport system of *Escherichia coli*. *J Bacteriol* 189, 6913-6918.
- Cadieux, N., Bradbeer, C., Reeger-Schneider, E., Koster, W., Mohanty, A.K., Wiener, M.C., and Kadner, R.J. (2002). Identification of the periplasmic cobalamin-binding protein BtuF of *Escherichia coli*. *J Bacteriol* 184, 706-717.
- Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97, 6640-6645.
- Davidson, A.L., Dassa, E., Orelle, C., and Chen, J. (2008). Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol Biol Rev* 72, 317-364, table of contents.
- Eitinger, T., Rodionov, D.A., Grote, M., and Schneider, E. (2011). Canonical and ECF-type ATP-binding cassette importers in prokaryotes: diversity in modular organization and cellular functions. *FEMS Microbiol Rev* 35, 3-67.
- Fowler, C.C., Brown, E.D., and Li, Y. (2010). Using a riboswitch sensor to examine coenzyme B(12) metabolism and transport in *E. coli*. *Chem Biol* 17, 756-765.
- Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* 177, 4121-4130.
- Higgins, C.F., and Ames, G.F. (1981). Two periplasmic transport proteins which interact with a common membrane receptor show extensive homology: complete nucleotide sequences. *Proc Natl Acad Sci U S A* 78, 6038-6042.
- Holland, I.B., and Blight, M.A. (1999). ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. *J Mol Biol* 293, 381-399.

- Hvorup, R.N., Goetz, B.A., Niederer, M., Hollenstein, K., Perozo, E., and Locher, K.P. (2007). Asymmetry in the structure of the ABC transporter-binding protein complex BtuCD-BtuF. *Science* 317, 1387-1390.
- Kadner, R.J. (1990). Vitamin B12 transport in *Escherichia coli*: energy coupling between membranes. *Mol Microbiol* 4, 2027-2033.
- Karpowich, N.K., Huang, H.H., Smith, P.C., and Hunt, J.F. (2003). Crystal structures of the BtuF periplasmic-binding protein for vitamin B12 suggest a functionally important reduction in protein mobility upon ligand binding. *J Biol Chem* 278, 8429-8434.
- Lawrence, J.G., and Roth, J.R. (1995). The cobalamin (coenzyme B12) biosynthetic genes of *Escherichia coli*. *J Bacteriol* 177, 6371-6380.
- Lewinson, O., Lee, A.T., Locher, K.P., and Rees, D.C. (2010). A distinct mechanism for the ABC transporter BtuCD-BtuF revealed by the dynamics of complex formation. *Nat Struct Mol Biol* 17, 332-338.
- Locher, K.P., Lee, A.T., and Rees, D.C. (2002). The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism. *Science* 296, 1091-1098.
- Lubelski, J., Konings, W.N., and Driessen, A.J. (2007). Distribution and physiology of ABC-type transporters contributing to multidrug resistance in bacteria. *Microbiol Mol Biol Rev* 71, 463-476.
- Postle, K., and Kadner, R.J. (2003). Touch and go: tying TonB to transport. *Mol Microbiol* 49, 869-882.
- Prossnitz, E., Gee, A., and Ames, G.F. (1989). Reconstitution of the histidine periplasmic transport system in membrane vesicles. Energy coupling and interaction between the binding protein and the membrane complex. *J Biol Chem* 264, 5006-5014.
- Roth, J.R., Lawrence, J.G., and Bobik, T.A. (1996). Cobalamin (coenzyme B12): synthesis and biological significance. *Annu Rev Microbiol* 50, 137-181.
- Scarlett, F.A., and Turner, J.M. (1976). Microbial metabolism of amino alcohols. Ethanolamine catabolism mediated by coenzyme B12-dependent ethanolamine ammonia-lyase in *Escherichia coli* and *Klebsiella aerogenes*. *J Gen Microbiol* 95, 173-176.
- Topp, S., and Gallivan, J.P. (2010). Emerging applications of riboswitches in chemical biology. *ACS Chem Biol* 5, 139-148.

Chapter 5

Discussion and Future Directions

This thesis describes an extensive and successful investigation into a novel method to detect small molecules within bacterial cells. The projects herein span the development of a technique that facilitates the engineering of custom sensors, the construction and characterization of a sensor that detects an important metabolite and the application of this sensor to provide novel insights into bacterial transport mechanisms. Overall, the data presented in this thesis validate the central hypothesis that riboswitches can be used as the basis of very effective intracellular metabolite sensors that can be used to address relevant and specific research questions. In order to expand the utility of this detection method and to overcome some of its current limitations, novel sensors that detect new ligands and that incorporate modifications into their design will be required. As a conclusion to this thesis, this section provides a discussion of some pertinent issues that emerged from each of the described research projects. Furthermore, the future outlook for the field of riboswitch-based sensors is considered.

The first project in this report describes the development of a novel method to create custom riboswitches from rationally designed libraries. Ultimately, this work resulted in the isolation of a theophylline-responsive riboswitch that activates gene expression ~7-fold in the presence of excess ligand. This induction ratio is similar to those reported in several other synthetic riboswitch publications (Desai and Gallivan,

2004; Suess et al., 2004), however it is modest compared to the most efficient ratios that have been reported which approach 100-fold (Lynch and Gallivan, 2009). The primary goal of this work was the development of the FACS-based selection methodology. A major conclusion from this work was that the library design was not optimal, but that the selection strategy was demonstrated to be proficient. This notion has since been validated by a report that described the isolation of an extremely efficient riboswitch using essentially the same FACS-based methodology (Lynch and Gallivan, 2009). The primary difference between our study and this one was the library design; while we targeted a transcriptional mechanism, the other report sought a translational mechanism.

A major shortcoming of the first project was our inability to define a regulatory mechanism for the riboswitches we identified. However, several experiments that were conducted using the most efficient riboswitch isolated did provide some insight into a number of possibilities. We were able to rule out the intended transcriptional mechanism, as the transcriptional terminator inserted upstream of the regulated gene was not important for riboswitch activity. The prospect of a mechanism that mimics the canonical RBS-sequestration regulatory scheme was also assessed and seemed unlikely based on mutational analysis and in line probing results. It is clear that both aptamer binding and the specific sequence isolated for the randomized portion of the library are essential for the activity of this riboswitch. Although no future experiments are currently planned to look into this matter further, my speculation remains that ligand binding affects ribosome access to the transcript in a manner that does not involve Watson-Crick

base-pairing of the RBS. One means of exploring this possibility would be to conduct *in vitro* binding assays using purified ribosomes (Nou and Kadner, 2000).

Despite its prevalence in naturally evolved riboswitches, the transcriptional termination mechanism has not yet been successfully targeted by a synthetic riboswitch project. While this could be at least partially due to a dearth of studies that have explored this possibility, this mechanism might also be inherently more challenging to re-create. One possibility is that the expression machinery for *E. coli*, the organism generally employed for riboswitch selection experiments, is better suited to accommodate a translational mechanism. This is supported by the fact that transcriptional riboswitches are relatively rare amongst those that have been identified for *E. coli* and related species (Barrick and Breaker, 2007). The underlying mechanism of this predisposition is not clear and thus this concern is entirely speculative. Another potential difficulty associated with isolating synthetic transcriptional riboswitches is the kinetic component of this mechanism (Wickiser et al., 2005). In order to regulate the formation of the terminator stem in the brief period in which RNA polymerase is paused on the template, riboswitch aptamers must quickly bind their ligands. Aptamers isolated *in vitro* have generally not been selected on the basis of fast rates of ligand association and therefore might not be compatible with such a mechanism. Future endeavours to address these issues would be worthwhile, as transcriptional riboswitches have the considerable advantage that they can be used in tandem. This opens up the possibility of employing multiple riboswitches that respond to either the same or to different small molecules to regulate a single gene or operon.

At the time of this project's beginnings, no general strategies for selection or screening riboswitch libraries had been established. In the time since, several such approaches have been described (see section 1.4). A disadvantage of FACS is the challenge associated with growing and maintaining a large library of clones in a single liquid culture. The loss of library diversity is a concern at each FACS step, particularly the stage that seeks cells expressing high levels of the fluorescent protein. Cells must be grown for a significant amount of time to allow for differences in fluorescent protein expression to become apparent. During this growth period, cells that express lower levels of the fluorescent protein might be subject to a minor growth advantage. The capacity of FACS to screen very large numbers of cells in a short period of time is likely to limit the impact of this effect, however, as it should be possible to identify cells with relatively rare phenotypes. There are several beneficial features of the FACS-based methodology. The stringency of each selection step can easily be modified by adjusting the gating settings. FACS is therefore particularly well suited to select for riboswitches with specific expression characteristics and for especially efficient riboswitches. The FACS approach is also flexible; it can be used with any strain or organism and can select for sequences either that activate or repress expression in response to ligand binding.

In the second project described in this thesis the concept of riboswitch-based sensors is assessed. The AdoCbl-responsive riboswitch that regulates the *btuB* gene was selected for the pioneering sensors. Prior to the construction of the final version of the sensor that is presented in chapter 3, several alternative designs were investigated. The relatively simple design of the prevailing plasmids was practical for the initial steps into

riboswitch-based detection. More elaborate systems, discussed later in this section, might be beneficial for future ventures. It should be noted that several plasmids were constructed using the same riboswitch that did not respond significantly to AdoCbl (data not shown). It is not possible to pinpoint the reason for this inactivity since there were several differences between these vectors and the final, functional constructs. The inactive plasmids had a different backbone, used a different promoter, had some additional sequence at the 5' end of the transcript and fused slightly less of the *btuB* coding sequence. Regardless of the underlying cause of this inactivity, this illustrates the point that engineering efficient sensors using established riboswitches might not always be as simple as it originally appears. Future endeavours should therefore carefully evaluate the performance of a number of parallel designs. In some instances, screening strategies such as the FACS-based method described in the first project could be useful to further optimize the performance of the probes.

Three different reporter genes were tested in conjunction with the AdoCbl-responsive riboswitch in the second project. Some benefits and drawbacks to each of these reporter methods are outlined in section 1.6. The fluorescent construct featuring DsRed-Express RFP produced a very weak signal. Upon the addition of saturating concentrations of AdoCbl, this signal was indistinguishable for that of a control strain lacking a RFP, as indicated by the relatively large error bars for that sample in Figure 3-2A. While the sensor is housed on a high copy plasmid, the P_{btuB} promoter that drives the expression of the reporter genes is thought to be relatively weak (White et al., 1973). Furthermore, it is possible that the addition of the N-terminal fusion influences the

folding of this protein. To produce fluorescent AdoCbl sensors with an improved signal intensity, another student is currently assessing a panel of other engineered fluorescent proteins. Preliminary results indicate that some of these options do produce a somewhat stronger signal, but that these reporters still lag well behind the β -galactosidase and firefly luciferase versions in this respect (Zohaib Ghazi, personal communication).

The mysterious source of the AdoCbl precursor DMB was amongst the aspects of vitamin B12 biology investigated in chapter 3. We demonstrated that the recently-identified *bluB* gene from the organism *M. smegmatis*, which breaks down FMN to form DMB, could be recombinantly expressed by *E. coli* to eliminate the need for exogenously supplied DMB. While genes fulfilling this metabolic need have long been sought, it is curious that only a small number of microbial species have an obvious *bluB* homologue (Campbell et al., 2006). It is entirely possible that many species lack this functionality, as genetic screens aiming to isolate such a protein in enteric bacteria have repeatedly failed to uncover candidate genes (Anderson et al., 2008). The genetic approach is limited in this instance, however, due to the ability of adenine to substitute for DMB and to the low concentrations of vitamin B12 required to be viable. The riboswitch sensors, which are able to distinguish between cells producing an intermediate and a high level of AdoCbl, could therefore be very useful in identifying such a factor if it exists. Indeed, the capacity to identify proteins with a more subtle role in metabolism or uptake makes riboswitch sensors a potentially valuable tool for the screening of genetic libraries. While such factors are often missed by genetic approaches that only perceive outcomes that affect

cell growth, methods that detect metabolite levels are more sensitive to less drastic effects.

While the second project included a number of experiments that examine vitamin B12-related biology, the emphasis of this project was the ability of riboswitch sensors to detect certain effects. The major objective of the third and final project was the investigation of vitamin B12 transport and the AdoCbl probes are applied as the primary methodology to address the research goals. Shifting the emphasis away from the detectors and focussing on a highly relevant and unexplored area of biology was an important step in demonstrating the value and utility of riboswitch sensors. The nature of the experiments in this study demanded that cytoplasmic AdoCbl concentrations be determined accurately and sensitively over a broad range. Furthermore, over one hundred unique constructs were tested under different conditions. Several constructs were tested using many different transport substrates and at a range of concentrations. In order for such a large number of samples to be accurately processed, AdoCbl concentrations needed to be measured in a high-throughput format. Carrying out these experiments using alternative technologies would not be practical. Studies of vitamin B12 transport in live cells have previously been conducted using radioactively-labelled substrates (Bradbeer, 1993; Cadieux et al., 2002; Fuller-Schaefer and Kadner, 2005). These substrates are not currently commercially available at usable concentrations. The process of produce and purify these compounds is very involved and requires the appropriate licenses for the use of radioactive cobalt. Finally, determining cytoplasmic vitamin B12 levels using radioactive substrates requires separation and wash steps that

severely hinders the amenability of these protocols to higher throughput experiments. The project described in chapter 4 therefore demonstrates that riboswitch sensors can provide a valuable and unique tool for biological research.

The third project provided a number of mechanistic insights into the transport of vitamin B12 across the cytoplasmic membrane in *E. coli* cells. Importantly, the high degree of phylogenetic conservation of this class of transporters makes this information relevant to many other uptake systems as well. The close parallels between the uptake machinery for many metal chelating compounds and vitamin B12 are particularly evident for both outer and cytoplasmic membrane transport. The most celebrated members of this family transport iron siderophores, which play an important role in bacterial virulence by allowing pathogens to effectively compete for the extremely limited pools of bioavailable iron. The information uncovered concerning the interactions between the PBP and the channel-forming component of the vitamin B12 transporter is presumably directly applicable to these systems, since the molecular interactions investigated are completely conserved in iron siderophore ABC transporters. The specific arrangement of the ligand binding pockets is obviously different for iron siderophore PBPs. However, the limited structural rearrangement involved in ligand release and the overall structure of these PBPs is thought to closely parallel BtuF (Borths et al., 2002; Braun and Herrmann, 2007). The nature of ligand binding required to satisfy the transport mechanism is therefore likely to be similar. The information regarding the vitamin B12 BtuF interactions that is described should therefore provide a useful basis for studying ligand interactions in iron siderophore PBPs as well. The follow up experiments proposed in

section 4.6 to further evaluate the hypotheses presented in Chapter 4 would therefore be very worthwhile, as this mechanistic information could be useful in developing strategies to inhibit this class of transporter.

While the advantages of riboswitch sensors have been emphasized throughout this report, it is important to acknowledge their limitations as well. One obvious limitation of these sensors is the fact that they do not provide information on the absolute concentration of their cognate metabolite. As mentioned in section 3.6, the relative information that they provide can be tremendously useful for a very wide range of experiments if proper controls are employed. However, for certain applications an absolute quantification is required. One example of this would be comparing metabolite levels to affinity constants of relevant enzymes or binding proteins. To provide numerical concentrations using riboswitch sensors, *in vitro* quantification values and reporter values could be compared over a wide range of ligand concentrations. These data could be used to provide a standard curve such that reporter values could be related back to ligand concentrations. This tactic would have strict limitations, however, since factors that affect gene expression or growth rate would invalidate the standard curve and lead to inaccurate results.

Another limitation of these sensors is the poor time resolution of the detection. This drawback will be difficult to completely overcome without drastically changing the sensing paradigm. The recognition event and the signalling event are, by definition, separated by the processes of translation and protein folding for a riboswitch sensor.

Using reporter proteins that fold quickly and that have short lifetimes would help improve time resolution. Furthermore, mathematic modelling of translation times, folding rates and reporter lifetimes could be applied in an attempt to compensate for the requisite separation of recognition and signalling events. Ultimately, however, the genetically-encoded metabolite sensing proteins described in section 1-1 that produce signals based on transient binding interactions are better suited to applications that require superior time resolution.

The success of the pioneering AdoCbl detectors has been encouraging and should inspire the construction of additional riboswitch-based sensors. The design of future probes should be guided by the specific downstream applications for which they are intended. There are countless small molecules that have an interesting role in prokaryotic biology that could potentially be the focus of future endeavours. Amongst the metabolites recognized by characterized naturally evolved riboswitches, an exciting candidate is cyclic-di-GMP. This molecule is currently a considerable focus of microbial pathogenesis research due to its diverse and prominent role in dictating the physiological mode of a bacterium. The contributions of the perplexingly high number enzymes that produce and break down this molecule in many species (Hengge, 2009) and further defining the environmental stimuli that regulate its levels are amongst the many potential applications of such a sensor. Many other riboswitch metabolites would also make good candidates, such as those that are essential to the viability of bacteria. Disrupting the synthesis or uptake of these molecules is a promising strategy for the development of novel antimicrobial compounds. Given the ability of riboswitch sensors to monitor these

biosynthetic and transport processes, they could be instrumental as tools for studying or discovering inhibitory compounds.

Future riboswitch-based sensors could potentially benefit from more elaborate designs. One possibility that we have considered is using a second reporter gene for normalization purposes. This reporter could be expressed from an identical promoter and carry an identical riboswitch with point mutations to the metabolite binding site. The differences observed between the expression of the two reporter constructs could therefore be reliably attributed to ligand binding. A caveat of this strategy is that the two reporters require similar folding and characteristics; certain pairs of fluorescent proteins might satisfy this stipulation. Another exciting possibility is combining riboswitches in tandem, a strategy employed by some natural riboswitches. Using tandem riboswitches that respond to the same metabolite, sensors could be constructed that are more sensitive to subtle changes in ligand concentrations. Tandem riboswitches could also be used to expand upon the dynamic range and signal amplitude of the probes. A related concept would be to combine two riboswitch sensors with different affinities for a common metabolite. One sensor could then be used to detect relatively low concentrations of the target and the other applied to probe higher levels. Finally, combining multiple sensors with distinct targets could be used to allow for multiplexing experiments that investigate how a particular stimulus effects key metabolic nodes. This prospect is particularly enticing using two targets that are physiologically related in some way, or for molecules for which a putative connection is being investigated.

Several downstream projects can be envisioned that focus on the unique abilities of riboswitch sensors. For the AdoCbl sensor described here, many other questions pertaining to AdoCbl transport could be addressed. A putative physical interaction between the PBP BtuF and TonB, which provides energy for outer membrane transport, has been uncovered (James et al., 2009). The physiological importance of this interaction could be investigated using the AdoCbl sensors as an integral aspect of this study. Additionally, genetic screens to identify new factors involved in vitamin B12 biology could be conducted. Screening an over-expression library seeking novel genes that use AdoCbl as a cofactor is a challenging but exciting possibility. This screen could be conducted using low levels of AdoCbl in a $\Delta btuR$ strain, which would be impaired for the adenosylation of damaged cobalamins. Highly expressed AdoCbl-utilizing enzymes could therefore affect AdoCbl levels through binding and also through the damage that can occur as a result of the reaction mechanism. Additional projects employing the AdoCbl sensor could benefit from using *S. enterica* as the host strain. *S. enterica* is a closely related species that has a complete biosynthetic pathway for AdoCbl as well as additional AdoCbl-utilizing processes, which would open up several other research opportunities.

The high-throughput capacity of riboswitch-based sensors makes them especially well suited for functions such as screening genetic libraries, small molecule collections or panels of growth conditions. A few projects along these lines have been discussed above. High-throughput approaches are regularly used in research relating to the discovery of antimicrobial molecules. The medical importance of these compounds and their

prominence microbiological research makes them a logical choice as targets of future riboswitch sensors. One specific application of such sensors that relates to the work presented here is the study of efflux systems that provide resistance to clinically used antibiotics. Large scale screens of the genetic or environmental conditions that contribute to the activity of these transporters would be facilitated by the fast and reliable detection afforded by a riboswitch sensor. Using riboswitch sensors to screening collections of small molecules for inhibitors of these efflux systems is also an enticing concept that could ultimately produce compounds that have potential for combination therapies.

One of the major advantages of riboswitch sensors over *in vitro* detection methods is their amenability to be used in biologically complex environments. Microorganisms generally subsist in niches that contain an immense diversity of living species. Specifically isolating a particular bacterium from these environments and quantifying a particular metabolite therein is a daunting prospect. Riboswitch-based sensors, by contrast, do not require the isolation of the target bacterial species and are therefore more suitable for experiments of this nature. The reporter system used for such projects should be highly sensitive and should not require cell lysis nor the exogenous addition of a substrate. The lux reporter system is therefore an appropriate choice, since it produces a luminescent signal and includes a full biosynthetic pathway for the substrate (Tu and Mager, 1995). The lux system is also compatible with luminescent imaging technologies, which have been applied to the visualization of bacterial infections within animal models (Andreu et al., 2011; Hutchens and Luker, 2007). One of the many interesting potential applications of riboswitch sensors within biologically complex environments is the

imaging of metabolite levels within pathogenic bacteria in an infection model. This highly relevant area of biology is almost completely unexplored. This experimental approach could provide unprecedented insights into the chemical environments encountered by pathogens in different tissues and the metabolic processes they rely upon to flourish under these conditions.

In conclusion, the research described in this thesis comprises the construction, characterization and application of a novel class of sensors that detect small molecule levels within living bacterial cells. These riboswitch-based sensors were used to examine several unexplored aspects of the biology of vitamin B12 in *E. coli*. These investigations were especially fruitful in providing mechanistic insights into the transport of vitamin B12 across the cytoplasmic membrane. Equally importantly, these pioneering efforts pave the way for the development of similar probes that can be used to detect other interesting molecules. The simplicity, flexibility and reliability of riboswitch-based sensors and the endless list of potential applications for which they could be exploited suggests that they have a bright future as valuable tools for biological studies.

APPENDIX

CHAPTER 2 SUPPLEMENTAL MATERIALS:



Figure S2-1 Cloning schematic. Each of the elements cloned into the pRB374 plasmid are shown along with the restriction sites (yellow boxes) into which they were cloned. The *metI* transcriptional terminator (stem-loop followed by a polyU tract) was cloned in PstI/PacI with the KpnI site added in to allow the aptamer (Y-shape) plus linker element to be cloned in PstI/KpnI. The sequence of each of the individual elements is given below.

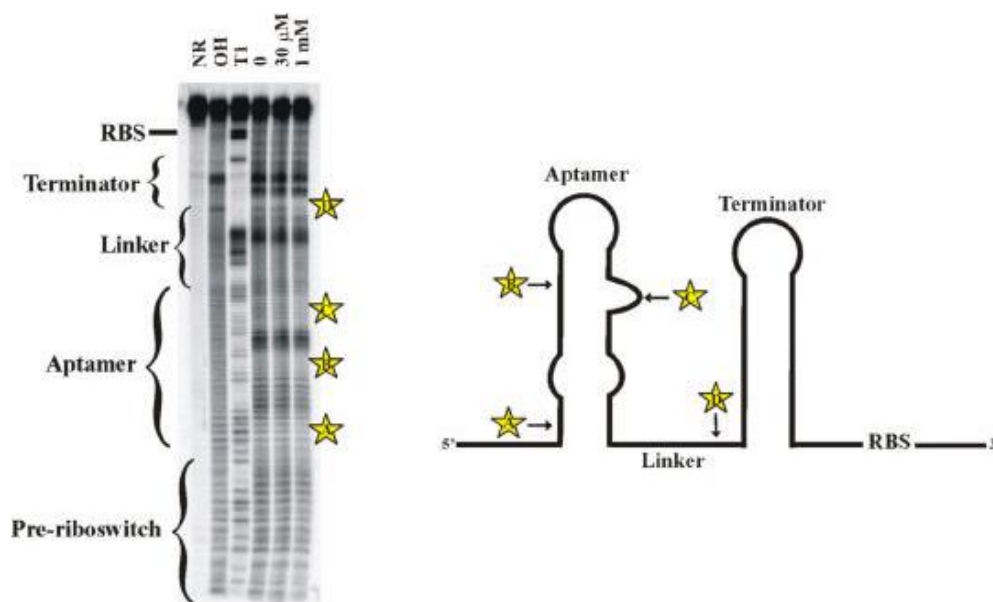


Figure S2-2 In Line Probing. Examination of the structural changes in the RS11A riboswitch in response to theophylline. The RNA was analyzed using 10% PAGE as shown on the left. Lanes 1-3 represent controls: NR: no reaction, OH: alkaline hydrolysis,

T1: T1 RNase digestion. Lanes 4-6 represent reactions with increasing concentrations of theophylline as shown. There are minor, yet reproducible, changes in the cleavage pattern that are shown by letters in yellow stars. The locations of the corresponding structural changes are shown on the diagram on the right. This diagram shows the location of the structural changes as they relate to significant elements in the 5'UTR. The structural changes noted lie in the aptamer domain as well as the selected linker domain.

Additional information on in line probing (Figure S2-2): In line probing is a method to examine the structure of RNA through its spontaneous breakdown over time. RNA linkages that assume a conformation that allow for the 2' OH to attack the neighbouring phosphate backbone more readily will be cleaved to a greater extent. By contrast, those linkages that are structurally constrained in a position that is not favourable for this transesterification reaction will experience less cleavage. This technique has been used extensively to study RNA-ligand interactions. Here, we observe a ligand-dependent decrease in cleavage in several regions of the aptamer sequence, indicating (presumably) that these regions are becoming more structured. We also note a decrease in cleavage for part of the selected linker region. The region showing a structural change using this technique is toward the end of the linker. This region is not sufficient for the switching activity, as mutants we have isolated to the start of the linker region have shown absolutely no activity (data not shown). Partial alkali digestion and partial T1 RNase digestion (specifically cleaves following G residues) were used as ladders.

Relevant Sequence Information for chapter 2:

1 - GFP element. Relevant restriction sites are shown in bold, the ribosome binding site is underlined, start and stop codons are shown in italics.

TTAATTAAGGAGGTTTCTTTA ATGGCTAGCAAAGGAGAAGAACTCTTCACTGGA
GTTGTCCCAATTCTTGTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTG
TCAGTGGAGAGGGTGAAGGTGATGCAACATACGGAAAACCTACCCTTAAATTTA
TTTGCCTACTGGAAAACCTACCTGTTCCATGGCCAACACTTGTCCTACTTTTAC
TTATGGTGTTCAGTGCTTTTCAAGATACCCGGATCATATGAAACGGCATGACTTT
TTCAAGAGTGCCATGCCC GAAGTTATGTACAGGAAAGAACTATATTTTCAAAG
ATGACGGGAACTACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTG
TTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGG
GCACAAATTGGAATACAAC TATAACTCACACAATGTATACATCATGGCAGACAAA
CAAAGAATGGAATCAAAGCTAACTTCAAATTAGACACAACATTGAAGATGGAA
GCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTG
TCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCC
CAACGAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGAT
TACACATGGCATGGATGAACTATACAAAGGATCCGGCTGCTAA CAAAGCCC
GAAAGGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCTTG
GGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCG
GATATCCACAGGACGGGTGTGCATGCAAGCTTGATATC**GAATTC**

2 – MetI transcriptional terminator. Relevant restriction sites are shown in bold.

CTGCAGCTTACAGGTC**GGTACC**TTAAACCCCTTCTTCTTATGAAGAAGGGGTT
TTTATTTTGA**TTAATTAA**

3 – Theophylline aptamer and linker region. Relevant restriction sites are shown in bold, the aptamer sequence is shown in italics. For linker sequences, N denotes a completely random residue and an asterisk denotes a residue that is present in 76% of sequences (with an 8% chance of each of the other three possible residues).

CTGCAGCAGCGATACGCAGGTGATACCAGCATCGTCTTGATGCCCTTGGCAGC
ACC- linker - **GGTACC**

Linker sequence for library 1: NNNA*T*A*A*G*A*A*G*A*A*G*

Linker sequence for library 2: A*C*T*A*T*A*T*G*A*A*G*A*G*

CHAPTER 3 SUPPLEMENTAL MATERIALS:

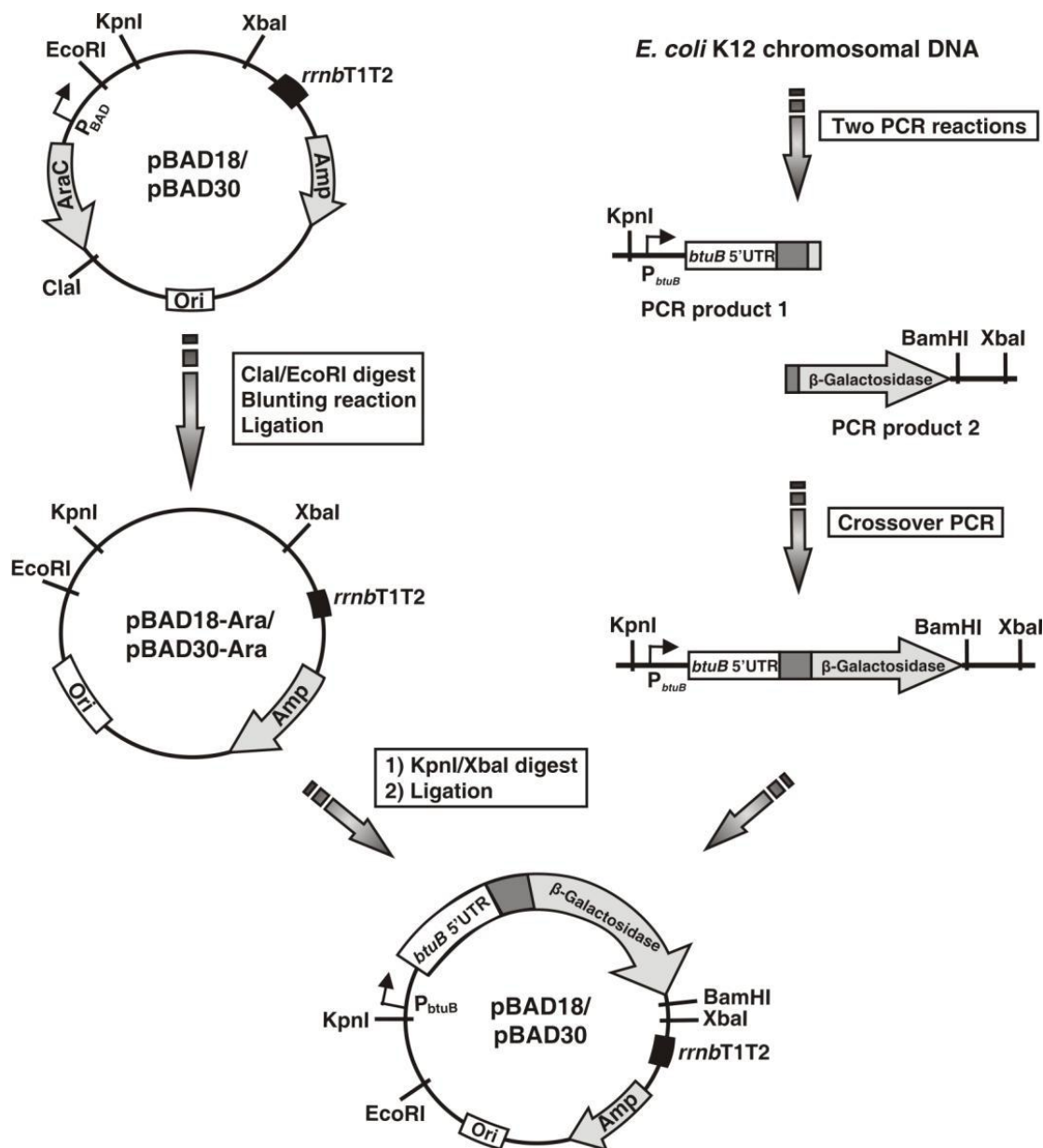


Figure S3-1 Cloning scheme for riboswitch sensor constructs showing the relevant plasmid features. Origins of replication are pBR322 (pBAD18, high copy) and p15A (pBAD30, medium copy). The dark gray shaded area included in the reporter constructs indicates the N-terminal truncation of BtuB. Constructs using alternate reporter genes were made in a synonymous manner. Details of molecular cloning are provided in Supplemental Experimental Procedures below.

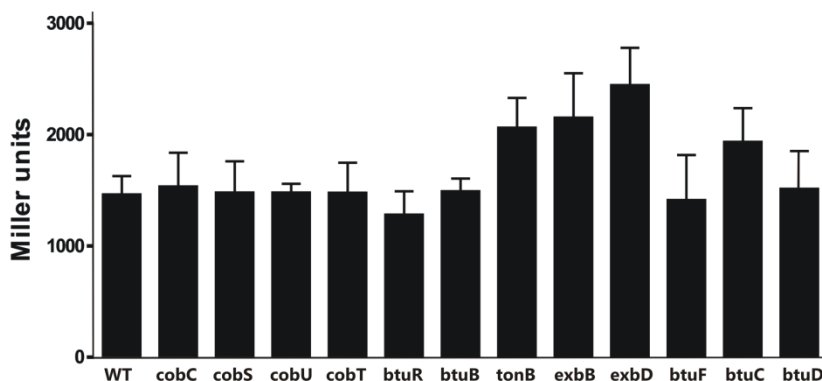


Figure S3-2 Raw reporter activity levels in each of the tested strains under unrepressed conditions. Reporter activity was generally consistent from strains to strain. The minor differences in the absolute intensity of the reporter activity observed are largely due to subtle day-to-day variation. A small effect stemming from the preparation of the competent cells is also possible, as strains prepared and stored in parallel tended to yield similar values.

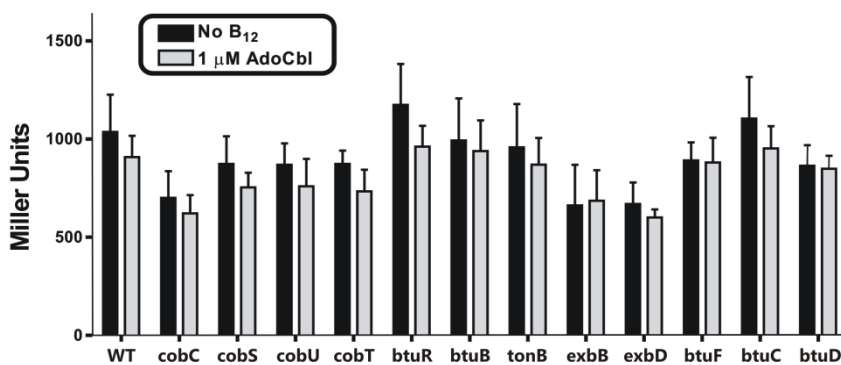


Figure S3-3 Activity of the mutant riboswitch construct in each of the strains tested. Raw reporter activity levels are shown under no B₁₂ and 1 μM AdoCbl conditions. As expected, there was no significant response to AdoCbl for mutant construct in any of the tested strains. As in Figure S2, some subtle variation in the raw reporter activity from strain to strain was observed.

Table S3-1. Description and sequence of oligonucleotide primers

Description	Restriction sites*	Sequence
Forward, for riboswitch amplification	KpnI	ctagtagggtaccag tcttgatgaattcctattgtg
Reverse, for riboswitch amplification - crossover with β -galactosidase	-	acgttgtaaacgacggccggaagacggcgagcacatcattgacc
Reverse, for riboswitch amplification - crossover with firefly luciferase	-	gtttttggcgtcttccatcggaagacggcgagcacatcattgacc
Reverse, for riboswitch amplification - crossover with RFP	-	gatgtcctcactgctcggccatcggaagacggcgagcacatcattgacc
Forward, for β -galactosidase amplification - crossover with riboswitch	-	gtgctcgccgctctccggccgctgtttacaacgtcgtgactggg
Forward, for firefly luciferase amplification - crossover with riboswitch	-	gtgctcgccgctctccgatggaagacgcaaaaacataaagaaag
Forward, for RFP amplification - crossover with riboswitch	-	gtgctcgccgctctccgatggcgagcagtgaggacatcatcaaggag
Reverse, for β -galactosidase amplification	BamHI, XbaI	atgatgctctagaggatccttattttgacaccagaccaactgg
Reverse, for firefly luciferase amplification	BamHI	atgatcgccgatccttacaatttgactttccgcccttc
Reverse, for RFP amplification	BamHI	atgatcgccgatccttaccaggaacaggtggtggcgccc
Forward, used to create riboswitch mutant construct	-	tagataccctccaagcccgttgacctgccgccaacgtcgc
Reverse, used to create riboswitch mutant construct	-	gcgacgttgcccgccaggtcaacggccttgagggggtatcta
Forward, for <i>M. smegmatis</i> BluB amplification	EcoRI	ctagtaggaattctaaggaggtttcttaatgcctcgacatgcgttcagccccagc
Reverse, for <i>M. smegmatis</i> BluB amplification	XbaI	cctagaatctagattacacggacgcgtgctcggccag
Forward, for <i>btuB</i> deletion construct (amplification of cm^I cassette)	-	ggaagtcatcatctttagtatcttagataccctccaagcccgaagaccttaa gaggtccaactttaccataa
Reverse, for <i>btuB</i> deletion construct (amplification of cm^I cassette)	-	ttagcacggcgtaacaacggcgtgctcggtaatcgattgcgcgcacgacttacg ccccccctgccactcatcg
Forward, for <i>cobT</i> deletion construct (amplification of cm^I cassette)	-	cttggtgaactggtcttctgctggcactgctataaagagacgtgagatgtaagag gtccaactttaccataa
Reverse, for <i>cobT</i> deletion construct (amplification of cm^I cassette)	-	aagtatagctttatcagactttctgttttaactgttcaaatcagaagtftacgccc gccctgccactcatcg

*Restriction sites incorporated into primers for subsequent cloning steps

Chapter 3 supplemental materials and methods:

Plasmid construction. The construction of the sensing plasmids is shown schematically in Figure S1. The first step in this process was the modification of the core plasmids, the high copy pBAD18 and medium copy pBAD30, to remove the arabinose-responsive elements. Both plasmids were modified in an identical manner to create β -galactosidase sensor plasmids (only pBAD18 steps are described here). pBAD30 versions of the other sensors have not been created. pBAD18 was digested with EcoRI and ClaI (New England Biolabs) followed by agarose gel purification of the large plasmid fragment. This fragment was then blunted using T4 DNA polymerase (Fermentas) and ligated to re-circularize. This modified vector, pBAD18-Ara, was used as the backbone for each of the sensor constructs.

The two sensing elements (riboswitch and reporter gene) were amplified separately by PCR and fused together using crossover PCR, as shown schematically in Figure S1, where a short segment from the 5'-end of β -galactosidase is added to the 3'-end of the riboswitch fragment, and a short segment of the 3'-end of the riboswitch fragment is added to the 5'-end of the β -galactosidase fragment. This is accomplished by building the appropriate overhanging sequences into the PCR primers used to amplify each of the fragments. Upon amplification, the two fragments contain an overlapping sequence to act as an annealing site that allows them to be fused in a second PCR that uses both fragments as a template, the forward primer from the riboswitch PCR and the reverse primer from the β -galactosidase PCR. Synonymous reactions were used to create

fusions with the other reporter genes, as well as to create the riboswitch binding mutations.

The riboswitch elements consisted of the promoter region (beginning at -70 relative to the transcriptional start site), 5'-UTR (+1 to +240) and an N-terminal coding sequence (+241 to +450) of the *E. coli btuB* gene. This fragment was PCR-amplified from *E. coli* K12 genomic DNA using a forward primer with a 5'-overhang containing a BamHI site immediately prior to the *btuB* sequence. The β -galactosidase gene was PCR amplified from *E. coli* K12 genomic DNA and the reaction used a reverse primer with a 5'-overhang featuring consecutive BamHI and XbaI sites. The reporter fragments were fused to the riboswitch using crossover PCR as described above, digested using KpnI and XbaI, and ligated into a similarly digested pBAD18-Ara vector (SI Fig. 1). This construct was named pRsBs-ACbl- β gal-HC (for RiboSwitch BioSensor - AdoCbl - β -galactosidase - high copy).

DsRed-Express RFP and firefly luciferase constructs were made in a synonymous manner. Because luciferase has an internal XbaI, BamHI (incorporated into the pRsBs-ACbl- β gal-HC vector) was used as the second restriction site for the other two reporter genes. Firefly luciferase was amplified using the pGEM-luc plasmid (Promega) as a template and RFP (DsRed-Express) was amplified from the pBAD-RFP_{EC} plasmid (Pfleger et al., 2005). This version of DsRed-Express contains a single valine to isoleucine mutation in the N-terminal region of the protein that was introduced to reduce translation-inhibiting structures, a feature that is not relevant for this study. Binding

mutants were constructed by incorporating the appropriate mutations into the overlapping segment of PCR primers used for crossover PCR (SI Table 1), which used the wild-type constructs as templates. PCR products were then digested and ligated in the same manner used to make the WT plasmids.

The BluB expression plasmid was also made using pBAD18 as a backbone, however the arabinose-responsive elements were not removed. The *bluB* gene was PCR-amplified from *M. smegmatis* genomic DNA with a forward primer that contained a consensus RBS and an EcoRI restriction site and a reverse primer that contained an XbaI restriction site. The fragment and pBAD18 were then digested with EcoRI/XbaI, purified by agarose gel and ligated. The resulting plasmid was named pBAD18-BluB.

All restriction digests described above used enzymes from New England Biolabs. All ligations were performed using T4 DNA ligase from Fermentas. PCR reactions were performed using High Fidelity PCR Enzyme mix (Fermentas). Reactions were carried out essentially according to supplier instructions. All plasmid constructs were confirmed by sequencing (Mobix labs).

Rich, chemically defined medium. For all reporter assays, cells were grown in a chemically defined medium intended to provide cells with a wealth of nutrients. Sterile M9 salts were diluted to a 1× final concentration and supplemented with the following nutrients that were sterilized by filtration prior to use: glucose (0.4%), MgSO₄ (2 mM), CaCl₂ (100 μM), DL-alanine (100 μg/mL), L-arginine (22 μg/mL), glycine (100 μg/mL),

L-histidine (22 µg/mL), L-leucine (20 µg/mL), L-isoleucine (20 µg/mL), L-methionine (20 µg/mL), L-proline (30 µg/mL), L-threonine (80 µg/mL), L-valine (40 µg/mL), L-lysine (88 µg/mL), L-cysteine (22 µg/mL), L-phenylalanine (20 µg/mL), L-tyrosine (20 µg/mL), L-tryptophan (20 µg/mL), L-serine (100 µg/mL), L-asparagine (100 µg/mL), L-glutamine (100 µg/mL), L-aspartic acid (100 µg/mL), L-glutamic acid (100 µg/mL), Niacin (1 µg/mL), calcium pantothenate (1 µg/mL), pyridoxine-HCl (1 µg/mL), thiamine-HCl (1 µg/mL), biotin (500 ng/mL), thymidine (40 µg/mL), thymine (40 µg/mL), uracil (40 µg/mL). The medium was also supplemented with 50 µg/mL ampicillin to maintain the plasmid-based sensor.

Plasmid sequences.

pRsBs-ACbI-βgal-HC:

```
cgatgCGTCCGGCGTAGAGGATCTGCTCATGTTTGACAGCTTATCATCGAATTCGAGCTCGGTACCAGATCT
tgatgaattcctatTTGTGAGCTACGTCTGGACAGTAACCTGTTACAACCTGTAGCATCCACTTGCCGGTCC
tGTGAGTTAATAGGGAATCCAGTGCGAATCTGGAGCTGACGCGCAGCGGTAAGGAAAGGTGCGATGATTGCG
ttatgCGGACACTGCCATTCGGTGGGAAGTCATCATCTCTTAGTATCTTAGATACCCCTCCAAGCCCGAAGA
cctgCCGGCCAACGTGCGATCTGGTTCTCATCTGCGTAATATTGATGAAACCTGCGGCATCCTTCTTCTA
ttgtggatgctttacaatgattaaaaaagcttcgctgctgacggcgtgttccgtcacggcattttccgcttg
ggcacaggataccagcccgatactctcgctcgttactgctaaccgttttgaacagccgCGCAGCACTGTGCT
TGCACCAACCACCGTTGTGACCCGTCAGGATATCGACCGCTGGCAGTCGACCTCGGTCAATGATGTGCTGCG
CCGTCTTCCGGCCGTGTTTTACAACGTGCTGACTGGGAAAACCCTGGCGTTACCCAACCTAATCGCCTTGC
AGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCAACAGTTGCG
CAGCCTGAATGGCGAATGGCGCTTTGCTGGTTTTCCGGCACCAGAAGCGGTGCCGGAAGCTGGCTGGAGTG
CGATCTTCTGAGGCCGATACTGTCTGCTGCCCTCAAACCTGGCAGATGCACGGTTACGATGCGCCCATCTA
CACCAACGTGACCTATCCATTACGGTCAATCCGCGTTTTGTTCCACGGAGAATCCGACGGGTTGTTACTC
GCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTATTTTTGATGGCGTTAACTC
```


gggattggtggcgcgactcctggagcccgctcagtatcggcgggaattccagctgagcgcgggtcgctacat
taccagttggtctggtgtcaaaaataaggatcctctagagtcgacctgcaggcatgcaagcttggctgtttt
ggcggatgagagaagattttcagcctgatacagattaaatcagaacgcagaagcggctctgataaaacagaat
ttgcctggcggcagtagcgcgggtgggtccacctgaccccatgccgaactcagaagtgaacgccgtagcgc
gatggtagtggtgggtctccccatgcgagagtagggaactgccaggcatcaaataaaacgaaaggctcagtc
gaaagactgggcctttcgttttatctgtttgttgcggtgaacgctctcctgagtaggacaaatccgccggg
agcggatttgaacgttgcaagcaacggcccgagggtggcggggcaggacgcccgcataaaactgccaggca
tcaaattaagcagaaggccatcctgacggatggcctttttgctttctacaaactcttttgtttatttttct
aaatacattcaaataatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaagga
agagtatgagtattcaacatttccgtgtcgcccttattccctttttgcggcattttgccttccctgtttttg
ctcaccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtggttacatcgaac
tggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcactttta
aagttctgctatgtggcgggtattatcccgtgttgacgccccgaagagcaactcggctgcgcgcatacact
attctcagaatgacttgggtgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagag
aattatgcagtgctgccataaccatgagtataacactgcccgaacttacttctgacaacgatcggaggac
cgaaggagctaaccgcttttttgacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagc
tgaatgaagccatacaciaacgacgagcgtgacaccacgatgcctgcagcaatggcaacaacgttgcgcaaac
tattaactggcgaactacttactctagcttccccggcaacaattaatagactggatggaggcggataaagttg
caggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccgggtgagcgtg
ggtctcgggtatcattgcagcactggggccagatggtaagccctccgctatcgtagttatctacacgacgg
ggagttaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggt
aactgtcagaccaagtttactcatatatacttttagattgatttacgcgccctgtagcggcgcattaagcgcg
gccccgtgtggtggttacgcgcagcgtgaccgctacacttgccagcgccttagcgcgccctcctttcgctttc
ttcccttcccttctgcaccgcttgcgggctttccccgtcaagctctaaatcgggggctccctttagggttc
cgatttagtgctttacggcacctcgacccccaaaaaacttgatttgggtgatggttcacgtagtgggccatcg
ccctgatagacgggtttttcgccctttgacgctggagtccacgctctttaatagtgactcttgttccaaact
tgaacaacactcaaccctatctcgggctattcttttgatttataagggattttgccgatttcggcctattgg
ttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaacaaaaatattaacgtttacaatttaaaag
gatctaggtgaagatccttttgataatctcatgacaaaaatcccttaacgtgagttttcgttccactgagc
gtcagacccccgtagaaaagatcaaaggatcttcttgagatccttttttctgcgcgtaatctgctgcttgca
aacaaaaaaaccaccgctaccagcgggtggtttgtttgcccggatcaagagctaccaactctttttccgaaggt
aactggcttcagcagagcgcagatacaciaactgtccttctagtgtagccgtagttaggccaccacttcaa
gaactctgtagcaccgcctacatacctcgctctgctaactcctgttaccagtggctgctgccagtgccgataa
gtcgtgtcttaccgggttgactcaagacgatagttaccggataaggcgcagcggctcgggctgaacgggggg
ttcgtgcacacagcccagcttgagcgaacgacctacaccgaactgagatacctacagcgtgagctatgaga
aagcggccacgcttcccgaaggagaaaggcggacaggtatccggtaagcggcaggggtcggaacaggagagcg

cacgagggagcttccagggggaaacgcctgggtatctttatagtcctgtcgggtttcgccacctctgacttga
gCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACCGGGCCTTTTTACG
gttCCTGGCCTTTTGTGGCCTTTTGTCCACATGTTCTTTCTGCGTTATCCCCTGATTCTGTGGATAACCG
tattaccgcctttgagtgagctgataccgctcgcgcgagccgaacgaccgagcgcagcagtcagtgagcga
ggaagcggaagagcgcctgatgCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTTACACCGCATATGGTG
cactctcagtacaatctgctctgatgCGCATAGTTAAGCCAGTATACTCCTCGCTATCGCTACGTGACTGG
gtcatggctgcgccccgacacccgccaacacccgctgacgcgcctgacgggcttGTCTGTCTCCCGGCATCC
gcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgc
gCGAGGCAGCAAGGAGATGGCGCCCAACAGTCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACA
agcGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAAC
CGCACCTGTGGCGCGGTGATGCCGGCCA

Green text, KpnI site

Red text, riboswitch element, including the promoter, riboswitch and the N-terminal segment of BtuB

Underlined text, start codon of reporter

Blue text, β -galactosidase sequence

Purple text, BamHI site

Orange text, XbaI site

pRsBs-ACbl-FFluc-HC:

Identical to the sequence for pRsBs-ACbl- β gal-HC (above), except the β -galactosidase sequence (blue) is replaced with the firefly luciferase sequence shown below:

```
atggaagacgccccaaaacataaagaaaggccccggcgccattctatcctctagaggatggaaccgctggagag
caactgcataaggctatgaagagatacgcacctgggttcctggaacaattgcttttacagatgcacatatcgag
gtgaacatcacgtacgcggaataacttcgaaatgtccggttcgggttgccagaagctatgaaacgatatgggctg
aatacaaatcacagaatcgtcgtatgcagtgaaaactctcttcaattctttatgcccgtgttgggcgcggtta
tttatcggagttgcagttgcgccccggaacgcacatataatgaacgtgaattgctcaacagtatgaacatt
tcgcagcctaccgtagtgtttgtttccaaaaaggggttgcaaaaaatgttgacgtgcaaaaaaattacca
ataatccagaaaattattatcatggattctaaaacggattaccagggatttcagtcgatgtacacgttcgtc
acatctcatctacctcccggttttaatgaatacgaatgttaccagagtcctttgatcgtgacaaaaaatt
gcactgataatgaattcctctggatctactgggttacctaaggggtgtggcccttccgcatagaactgcctgc
gtcagattctcgcgatgccagagatcctatgttttgcaatcaaatcattccggatactgagattttaagtgtt
gttccattccatcacggttttggaatgtttactacactcggatatttgatatgtggatttcgagtcgtctta
atgtatagatttgaagaagagctgtttttacgatcccttcaggattacaaaattcaaagtgcggttgctagta
ccaaccctatgttcttcgccccaaagcactctgattgacaaaatcgaatcttaatttacacgaaatt
gcttctggggggcgccacctcttcgaaagaagtccgggaagcgggttgcaaacgcttccatcttcagggata
cgacaaggatattgggctcactgagactacatcagctattctgattacaccgagggggatgataaacggggc
gcggtcggtaaaagtgttccatgtttttgaagcgaaggttgtggatctggataccgggaaaacgctgggctt
aatcagagaggcgaattatgtgtcagaggacctatgattatgtccggttatgtaacaatccggaagcgacc
aacgccttgattgacaaggatggatggctacattctggagacatagcttactgggacgaagacgaacacttc
ttcatagttgaccgcttgaagtctttaattaaatacaaaggatatcaggtggcccccgctgaattggaatcg
atattgttacaacacccccaacatcttcgacgcgggctggcaggtcttcccgcgatgacgcgggtgaactt
cccgcgccgcttgttgtttttggagcacggaaagacgatgacggaaaaagagatcgtggattacgtcgccagt
caagtaacaaccgcaaaaagtgtcgcgaggagttgtgtttgtggacgaagtaccgaaaggtcttaccgga
aaactcgacgcaagaaaaatcagagagatcctcataaaggccaagaagggcggaagtccaaattgtaa
```

pRsBs-ACbl-RFP-HC:

Identical to the sequence for pRsBs-ACbl- β gal-HC (above), except the β -galactosidase sequence (blue) is replaced with the DsRed-Express sequence shown below:

```
atggcgagcagtgaggacatcatcaaggagtcatgcgcttcaagggtgcatggagggctccgtgaacggc
cacgagttcgagatcgagggcgagggcgagggccgcccctacgagggcaccagaccgccaagctgaaggtg
accaagggcgccccctgcccttgcctgggacatcctgtccccccagttccagtacggctccaaggtgtac
gtgaagcaccgacatccccgactacaagaagctgtccttccccgagggcttcaagtgggagcgctg
atgaacttcgaggacggcggtggtgacctgacctcctcctgcaggacggctccttcatctac
aaggtgaagttcatcggtgaacttccccctccgacggccccgtaatgcagaagaagactatgggctgggag
gcctccaccgagcgctgtacccccgacggcggtgctgaagggcgagatccacaaggccctgaagctgaag
gacggcgccactacctgggtggagttcaagtccatctacatggccaagaagcccgtgcagctgcccggtac
tactacgtggactccaagctggacatcacctcccacaacgaggactacaccatcgtggagcagtagcgcgc
gccgagggccgcccaccacctgttctgtaa
```

References from chapter 3 supplemental material:

1. Pflieger BF, Fawzi NJ, & Keasling JD (2005) Optimization of DsRed production in *Escherichia coli*: effect of ribosome binding site sequestration on translation efficiency. *Biotechnol Bioeng* 92(5):553-558.

CHAPTER 4 SUPPLEMENTAL MATERIALS:

	1	2	3	4	5	6	7	8	9	10	11	12
A	WT	A	C	E	F	EV	WT	G	H	I	K	EV
B	L	T	S	R	Q	P	N	M	L	K	I	E
C	M	H	G	WT	EV	Y	W	V	T	S	R	F
D	N	Q	P	N	E202A	M	L	K	I	H	G	E202A
E	P	F	E	C	A	WT	EV	Y	W	V	T	Y
F	A	S	WT	EV	R	Q	P	N	A	C	E	W
G	C	F	G	H	I	K	E202A	EV	WT	L	M	V
H	EV	Q	R	S	T	WT	EV	V	W	Y	E202A	WT

Figure S4-1 Sample plate layout for the saturation mutagenesis library. The above plate layout for the BtuF D242 saturation mutagenesis library features four replicates of each mutant scattered throughout the 96-well plate. Also included are 8 wild-type (WT) control wells and 8 empty vector (EV) control wells that lack BtuF. Four controls of the BtuF E202A mutant are also randomly scattered.

References

- Anderson, P.J., Lango, J., Carkeet, C., Britten, A., Krautler, B., Hammock, B.D., and Roth, J.R. (2008). One pathway can incorporate either adenine or dimethylbenzimidazole as an alpha-axial ligand of B12 cofactors in *Salmonella enterica*. *J Bacteriol* *190*, 1160-1171.
- Andreu, N., Zelmer, A., and Wiles, S. (2011). Noninvasive biophotonic imaging for studies of infectious disease. *FEMS Microbiol Rev* *35*, 360-394.
- Ataide, S.F., Wilson, S.N., Dang, S., Rogers, T.E., Roy, B., Banerjee, R., Henkin, T.M., and Ibba, M. (2007). Mechanisms of resistance to an amino acid antibiotic that targets translation. *ACS Chem Biol* *2*, 819-827.
- Bajad, S.U., Lu, W., Kimball, E.H., Yuan, J., Peterson, C., and Rabinowitz, J.D. (2006). Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry. *J Chromatogr A* *1125*, 76-88.
- Banerjee, R., and Ragsdale, S.W. (2003). The many faces of vitamin B12: catalysis by cobalamin-dependent enzymes. *Annu Rev Biochem* *72*, 209-247.
- Bang, G.S., Cho, S., and Kim, B.G. (2005). A novel electrochemical detection method for aptamer biosensors. *Biosens Bioelectron* *21*, 863-870.
- Barrick, J.E., and Breaker, R.R. (2007). The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol* *8*, R239.
- Beisel, C.L., and Storz, G. (2010). Base pairing small RNAs and their roles in global regulatory networks. *FEMS Microbiol Rev* *34*, 866-882.
- Bertin, Y., Girardeau, J.P., Chaucheyras-Durand, F., Lyan, B., Pujos-Guillot, E., Harel, J., and Martin, C. (2011). Enterohaemorrhagic *Escherichia coli* gains a competitive advantage by using ethanolamine as a nitrogen source in the bovine intestinal content. *Environ Microbiol* *13*, 365-377.
- Bevis, B.J., and Glick, B.S. (2002). Rapidly maturing variants of the *Discosoma* red fluorescent protein (DsRed). *Nat Biotechnol* *20*, 83-87.
- Bittker, J.A., Phillips, K.J., and Liu, D.R. (2002). Recent advances in the in vitro evolution of nucleic acids. *Curr Opin Chem Biol* *6*, 367-374.
- Blair, D.F. (1995). How bacteria sense and swim. *Annu Rev Microbiol* *49*, 489-522.
- Blount, K.F., and Breaker, R.R. (2006). Riboswitches as antibacterial drug targets. *Nat Biotechnol* *24*, 1558-1564.
- Blount, K.F., Wang, J.X., Lim, J., Sudarsan, N., and Breaker, R.R. (2007). Antibacterial lysine analogs that target lysine riboswitches. *Nat Chem Biol* *3*, 44-49.
- Borths, E.L., Locher, K.P., Lee, A.T., and Rees, D.C. (2002). The structure of *Escherichia coli* BtuF and binding to its cognate ATP binding cassette transporter. *Proc Natl Acad Sci U S A* *99*, 16642-16647.
- Boyle, P.M., and Silver, P.A. (2009). Harnessing nature's toolbox: regulatory elements for synthetic biology. *J R Soc Interface* *6 Suppl 4*, S535-546.
- Bradbeer, C. (1993). The proton motive force drives the outer membrane transport of cobalamin in *Escherichia coli*. *J Bacteriol* *175*, 3146-3150.

- Braun, V. (1995). Energy-coupled transport and signal transduction through the gram-negative outer membrane via TonB-ExbB-ExbD-dependent receptor proteins. *FEMS Microbiol Rev* *16*, 295-307.
- Braun, V., and Herrmann, C. (2007). Docking of the periplasmic FecB binding protein to the FecCD transmembrane proteins in the ferric citrate transport system of *Escherichia coli*. *J Bacteriol* *189*, 6913-6918.
- Breaker, R.R. (1997). DNA aptamers and DNA enzymes. *Curr Opin Chem Biol* *1*, 26-31.
- Breaker, R.R. (2010). Riboswitches and the RNA World. *Cold Spring Harb Perspect Biol*.
- Bunka, D.H., Platonova, O., and Stockley, P.G. (2010). Development of aptamer therapeutics. *Curr Opin Pharmacol* *10*, 557-562.
- Cadieux, N., Bradbeer, C., Reeger-Schneider, E., Koster, W., Mohanty, A.K., Wiener, M.C., and Kadner, R.J. (2002). Identification of the periplasmic cobalamin-binding protein BtuF of *Escherichia coli*. *J Bacteriol* *184*, 706-717.
- Campbell, G.R., Taga, M.E., Mistry, K., Lloret, J., Anderson, P.J., Roth, J.R., and Walker, G.C. (2006). *Sinorhizobium meliloti* bluB is necessary for production of 5,6-dimethylbenzimidazole, the lower ligand of B12. *Proc Natl Acad Sci U S A* *103*, 4634-4639.
- Cao, X., Li, S., Chen, L., Ding, H., Xu, H., Huang, Y., Li, J., Liu, N., Cao, W., Zhu, Y., *et al.* (2009). Combining use of a panel of ssDNA aptamers in the detection of *Staphylococcus aureus*. *Nucleic Acids Res* *37*, 4621-4628.
- Cheah, M.T., Wachter, A., Sudarsan, N., and Breaker, R.R. (2007). Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. *Nature* *447*, 497-500.
- Chou, L.Y., Ming, K., and Chan, W.C. (2010). Strategies for the intracellular delivery of nanoparticles. *Chem Soc Rev* *40*, 233-245.
- Collins, J.A., Irnov, I., Baker, S., and Winkler, W.C. (2007). Mechanism of mRNA destabilization by the glmS ribozyme. *Genes Dev* *21*, 3356-3368.
- Cooper, J.a.C., T, ed. (2004). *Biosensors (Practical Approach)*, 2nd edn (Oxford University Press).
- Coppins, R.L., Hall, K.B., and Groisman, E.A. (2007). The intricate world of riboswitches. *Curr Opin Microbiol* *10*, 176-181.
- Dambach, M.D., and Winkler, W.C. (2009). Expanding roles for metabolite-sensing regulatory RNAs. *Curr Opin Microbiol* *12*, 161-169.
- Davidson, A.L., Dassa, E., Orelle, C., and Chen, J. (2008). Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol Biol Rev* *72*, 317-364, table of contents.
- Davis, J.H., and Szostak, J.W. (2002). Isolation of high-affinity GTP aptamers from partially structured RNA libraries. *Proc Natl Acad Sci U S A* *99*, 11616-11621.
- de Las Heras, A., Carreno, C.A., Martinez-Garcia, E., and de Lorenzo, V. (2010). Engineering input/output nodes in prokaryotic regulatory circuits. *FEMS Microbiol Rev* *34*, 842-865.
- Delehanty, J.B., Mattoussi, H., and Medintz, I.L. (2009). Delivering quantum dots into cells: strategies, progress and remaining issues. *Anal Bioanal Chem* *393*, 1091-1105.
- Desai, S.K., and Gallivan, J.P. (2004). Genetic screens and selections for small molecules based on a synthetic riboswitch that activates protein translation. *J Am Chem Soc* *126*, 13247-13254.
- Dettmer, K., Aronov, P.A., and Hammock, B.D. (2007). Mass spectrometry-based metabolomics. *Mass Spectrom Rev* *26*, 51-78.
- Deuschle, K., Fehr, M., Hilpert, M., Lager, I., Lalonde, S., Looger, L.L., Okumoto, S., Persson, J., Schmidt, A., and Frommer, W.B. (2005). Genetically encoded sensors for metabolites. *Cytometry A* *64*, 3-9.

- Drolet, D.W., Moon-McDermott, L., and Romig, T.S. (1996). An enzyme-linked oligonucleotide assay. *Nat Biotechnol* *14*, 1021-1025.
- Ellington, A.D. (1994). RNA selection. Aptamers achieve the desired recognition. *Curr Biol* *4*, 427-429.
- Ellington, A.D., and Szostak, J.W. (1990). In vitro selection of RNA molecules that bind specific ligands. *Nature* *346*, 818-822.
- Esposito C.L., C.S., de Franciscis V., Cerchia L. (2011). New Insight into Clinical Development of Nucleic Acid Aptamers. *Discovery Medicine*.
- Fang, X., Cao, Z., Beck, T., and Tan, W. (2001). Molecular aptamer for real-time oncoprotein platelet-derived growth factor monitoring by fluorescence anisotropy. *Anal Chem* *73*, 5752-5757.
- Fehr, M., Lalonde, S., Lager, I., Wolff, M.W., and Frommer, W.B. (2003). In vivo imaging of the dynamics of glucose uptake in the cytosol of COS-7 cells by fluorescent nanosensors. *J Biol Chem* *278*, 19127-19133.
- Ferre-D'Amare, A.R. (2010). The glmS ribozyme: use of a small molecule coenzyme by a gene-regulatory RNA. *Q Rev Biophys* *43*, 423-447.
- Fowler, C.C., Brown, E.D., and Li, Y. (2008). A FACS-based approach to engineering artificial riboswitches. *Chembiochem* *9*, 1906-1911.
- Fowler, C.C., Brown, E.D., and Li, Y. (2010). Using a riboswitch sensor to examine coenzyme B(12) metabolism and transport in *E. coli*. *Chem Biol* *17*, 756-765.
- Fox, K.A., Ramesh, A., Stearns, J.E., Bourgogne, A., Reyes-Jara, A., Winkler, W.C., and Garsin, D.A. (2009). Multiple posttranscriptional regulatory mechanisms partner to control ethanolamine utilization in *Enterococcus faecalis*. *Proc Natl Acad Sci U S A* *106*, 4435-4440.
- Frey, B., McCloskey, J., Kersten, W., and Kersten, H. (1988). New function of vitamin B12: cobamide-dependent reduction of epoxyqueuosine to queuosine in tRNAs of *Escherichia coli* and *Salmonella typhimurium*. *J Bacteriol* *170*, 2078-2082.
- Froese, D.S., Dobson, C.M., White, A.P., Wu, X., Padovani, D., Banerjee, R., Haller, T., Gerlt, J.A., Surette, M.G., and Gravel, R.A. (2009). Sleeping beauty mutase (sbm) is expressed and interacts with ygfD in *Escherichia coli*. *Microbiol Res* *164*, 1-8.
- Fuller-Schaefer, C.A., and Kadner, R.J. (2005). Multiple extracellular loops contribute to substrate binding and transport by the *Escherichia coli* cobalamin transporter BtuB. *J Bacteriol* *187*, 1732-1739.
- Gallo, S., Oberhuber, M., Sigel, R.K., and Krautler, B. (2008). The corrin moiety of coenzyme B12 is the determinant for switching the btuB riboswitch of *E. coli*. *Chembiochem* *9*, 1408-1414.
- Garsin, D.A. (2010). Ethanolamine utilization in bacterial pathogens: roles and regulation. *Nat Rev Microbiol* *8*, 290-295.
- Geiger, A., Burgstaller, P., von der Eltz, H., Roeder, A., and Famulok, M. (1996). RNA aptamers that bind L-arginine with sub-micromolar dissociation constants and high enantioselectivity. *Nucleic Acids Res* *24*, 1029-1036.
- Geissmann, T., Marzi, S., and Romby, P. (2009). The role of mRNA structure in translational control in bacteria. *RNA Biol* *6*, 153-160.
- Giepmans, B.N., Adams, S.R., Ellisman, M.H., and Tsien, R.Y. (2006). The fluorescent toolbox for assessing protein location and function. *Science* *312*, 217-224.

- Goodacre, R., Vaidyanathan, S., Dunn, W.B., Harrigan, G.G., and Kell, D.B. (2004). Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends Biotechnol* 22, 245-252.
- Gopinath, S.C. (2007). Methods developed for SELEX. *Anal Bioanal Chem* 387, 171-182.
- Gopinath, S.C., Balasundaresan, D., Akitomi, J., and Mizuno, H. (2006). An RNA aptamer that discriminates bovine factor IX from human factor IX. *J Biochem* 140, 667-676.
- Green, N.J., Grundy, F.J., and Henkin, T.M. (2010). The T box mechanism: tRNA as a regulatory molecule. *FEBS Lett* 584, 318-324.
- Gronewold, T.M., Glass, S., Quandt, E., and Famulok, M. (2005). Monitoring complex formation in the blood-coagulation cascade using aptamer-coated SAW sensors. *Biosens Bioelectron* 20, 2044-2052.
- Grundy, F.J., Rollins, S.M., and Henkin, T.M. (1994). Interaction between the acceptor end of tRNA and the T box stimulates antitermination in the *Bacillus subtilis* *tyrS* gene: a new role for the discriminator base. *J Bacteriol* 176, 4518-4526.
- Guo, K.T., SchAfer, R., Paul, A., Gerber, A., Ziemer, G., and Wendel, H.P. (2006). A new technique for the isolation and surface immobilization of mesenchymal stem cells from whole bone marrow using high-specific DNA aptamers. *Stem Cells* 24, 2220-2231.
- Haller, T., Buckel, T., Retey, J., and Gerlt, J.A. (2000). Discovering new enzymes and metabolic pathways: conversion of succinate to propionate by *Escherichia coli*. *Biochemistry* 39, 4622-4629.
- Halpern, J. (1985). Mechanisms of coenzyme B12-dependent rearrangements. *Science* 227, 869-875.
- Hampel, K.J., and Tinsley, M.M. (2006). Evidence for preorganization of the *glmS* ribozyme ligand binding pocket. *Biochemistry* 45, 7861-7871.
- Han, J., Datla, R., Chan, S., and Borchers, C.H. (2009). Mass spectrometry-based technologies for high-throughput metabolomics. *Bioanalysis* 1, 1665-1684.
- Hengge, R. (2009). Principles of c-di-GMP signalling in bacteria. *Nat Rev Microbiol* 7, 263-273.
- Hermann, T., and Patel, D.J. (2000). Adaptive recognition by nucleic acid aptamers. *Science* 287, 820-825.
- Hoffmann, B., Oberhuber, M., Stupperich, E., Bothe, H., Buckel, W., Konrat, R., and Krautler, B. (2000). Native corrinoids from *Clostridium cochlearium* are adenylcobamides: spectroscopic analysis and identification of pseudovitamin B(12) and factor A. *J Bacteriol* 182, 4773-4782.
- Holland, I.B., and Blight, M.A. (1999). ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. *J Mol Biol* 293, 381-399.
- Honda, A., Adams, S.R., Sawyer, C.L., Lev-Ram, V., Tsien, R.Y., and Dostmann, W.R. (2001). Spatiotemporal dynamics of guanosine 3',5'-cyclic monophosphate revealed by a genetically encoded, fluorescent indicator. *Proc Natl Acad Sci U S A* 98, 2437-2442.
- Hutchens, M., and Luker, G.D. (2007). Applications of bioluminescence imaging to the study of infectious diseases. *Cell Microbiol* 9, 2315-2322.
- Jain, V.K., and Magrath, I.T. (1991). A chemiluminescent assay for quantitation of beta-galactosidase in the femtogram range: application to quantitation of beta-galactosidase in *lacZ*-transfected cells. *Anal Biochem* 199, 119-124.
- James, K.J., Hancock, M.A., Gagnon, J.N., and Coulton, J.W. (2009). TonB interacts with BtuF, the *Escherichia coli* periplasmic binding protein for cyanocobalamin. *Biochemistry* 48, 9212-9220.

- Jenison, R.D., Gill, S.C., Pardi, A., and Polisky, B. (1994). High-resolution molecular discrimination by RNA. *Science* *263*, 1425-1429.
- Jin, Y., Watt, R.M., Danchin, A., and Huang, J.D. (2009). Use of a riboswitch-controlled conditional hypomorphic mutation to uncover a role for the essential *csrA* gene in bacterial autoaggregation. *J Biol Chem* *284*, 28738-28745.
- Joseph, B., Przybilla, K., Stuhler, C., Schauer, K., Slaghuis, J., Fuchs, T.M., and Goebel, W. (2006). Identification of *Listeria monocytogenes* genes contributing to intracellular replication by expression profiling and mutant screening. *J Bacteriol* *188*, 556-568.
- Joyce, G.F. (1994). In vitro evolution of nucleic acids. *Curr Opin Struct Biol* *4*, 331-336.
- Juskowiak, B. (2011). Nucleic acid-based fluorescent probes and their analytical potential. *Anal Bioanal Chem* *399*, 3157-3176.
- Keck, B., Munder, M., and Renz, P. (1998). Biosynthesis of cobalamin in *Salmonella typhimurium*: transformation of riboflavin into the 5,6-dimethylbenzimidazole moiety. *Arch Microbiol* *171*, 66-68.
- Keefe, A.D., and Cload, S.T. (2008). SELEX with modified nucleotides. *Curr Opin Chem Biol* *12*, 448-456.
- Klumpp, J., and Fuchs, T.M. (2007). Identification of novel genes in genomic islands that contribute to *Salmonella typhimurium* replication in macrophages. *Microbiology* *153*, 1207-1220.
- Lau, P.S., Coombes, B.K., and Li, Y. (2010). A general approach to the construction of structure-switching reporters from RNA aptamers. *Angew Chem Int Ed Engl* *49*, 7938-7942.
- Lawhon, S.D., Frye, J.G., Suyemoto, M., Porwollik, S., McClelland, M., and Altier, C. (2003). Global regulation by *CsrA* in *Salmonella typhimurium*. *Mol Microbiol* *48*, 1633-1645.
- Lawrence, J.G., and Roth, J.R. (1995). The cobalamin (coenzyme B12) biosynthetic genes of *Escherichia coli*. *J Bacteriol* *177*, 6371-6380.
- Lawrence, J.G., and Roth, J.R. (1996). Evolution of coenzyme B12 synthesis among enteric bacteria: evidence for loss and reacquisition of a multigene complex. *Genetics* *142*, 11-24.
- Lee, E.R., Baker, J.L., Weinberg, Z., Sudarsan, N., and Breaker, R.R. (2010). An allosteric self-splicing ribozyme triggered by a bacterial second messenger. *Science* *329*, 845-848.
- Lee, E.R., Blount, K.F., and Breaker, R.R. (2009). Roseoflavin is a natural antibacterial compound that binds to FMN riboswitches and regulates gene expression. *RNA Biol* *6*, 187-194.
- Lee, J.F., Stovall, G.M., and Ellington, A.D. (2006). Aptamer therapeutics advance. *Curr Opin Chem Biol* *10*, 282-289.
- Lemay, J.F., Desnoyers, G., Blouin, S., Heppell, B., Bastet, L., St-Pierre, P., Masse, E., and Lafontaine, D.A. (2011). Comparative study between transcriptionally- and translationally-acting adenine riboswitches reveals key differences in riboswitch regulatory mechanisms. *PLoS Genet* *7*, e1001278.
- Levy, M., Cater, S.F., and Ellington, A.D. (2005). Quantum-dot aptamer beacons for the detection of proteins. *ChemBiochem* *6*, 2163-2166.
- Locher, K.P., and Borths, E. (2004). ABC transporter architecture and mechanism: implications from the crystal structures of BtuCD and BtuF. *FEBS Lett* *564*, 264-268.
- Lundrigan, M.D., and Kadner, R.J. (1989). Altered cobalamin metabolism in *Escherichia coli* *btuR* mutants affects *btuB* gene regulation. *J Bacteriol* *171*, 154-161.

- Lundrigan, M.D., Koster, W., and Kadner, R.J. (1991). Transcribed sequences of the *Escherichia coli* *btuB* gene control its expression and regulation by vitamin B12. *Proc Natl Acad Sci U S A* *88*, 1479-1483.
- Lynch, S.A., Desai, S.K., Sajja, H.K., and Gallivan, J.P. (2007). A high-throughput screen for synthetic riboswitches reveals mechanistic insights into their function. *Chem Biol* *14*, 173-184.
- Lynch, S.A., and Gallivan, J.P. (2009). A flow cytometry-based screen for synthetic riboswitches. *Nucleic Acids Res* *37*, 184-192.
- Maadani, A., Fox, K.A., Mylonakis, E., and Garsin, D.A. (2007). *Enterococcus faecalis* mutations affecting virulence in the *Caenorhabditis elegans* model host. *Infect Immun* *75*, 2634-2637.
- Maharjan, R.P., and Ferenci, T. (2003). Global metabolite analysis: the influence of extraction methodology on metabolome profiles of *Escherichia coli*. *Anal Biochem* *313*, 145-154.
- Mandal, M., Lee, M., Barrick, J.E., Weinberg, Z., Emilsson, G.M., Ruzzo, W.L., and Breaker, R.R. (2004). A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science* *306*, 275-279.
- Mapelli, V., Olsson, L., and Nielsen, J. (2008). Metabolic footprinting in microbiology: methods and applications in functional genomics and biotechnology. *Trends Biotechnol* *26*, 490-497.
- Meyer, M.M., Hammond, M.C., Salinas, Y., Roth, A., Sudarsan, N., and Breaker, R.R. (2011). Challenges of ligand identification for riboswitch candidates. *RNA Biol* *8*, 5-10.
- Michaud, M., Jourdan, E., Villet, A., Ravel, A., Grosset, C., and Peyrin, E. (2003). A DNA aptamer as a new target-specific chiral selector for HPLC. *J Am Chem Soc* *125*, 8672-8679.
- Miethke, M., and Marahiel, M.A. (2007). Siderophore-based iron acquisition and pathogen control. *Microbiol Mol Biol Rev* *71*, 413-451.
- Miles, Z.D., McCarty, R.M., Molnar, G., and Bandarian, V. (2011). Discovery of epoxyqueuosine (oQ) reductase reveals parallels between halo-respiration and tRNA modification. *Proc Natl Acad Sci U S A* *108*, 7368-7372.
- Mironov, A.S., Gusarov, I., Rafikov, R., Lopez, L.E., Shatalin, K., Kreneva, R.A., Perumov, D.A., and Nudler, E. (2002). Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* *111*, 747-756.
- Montange, R.K., and Batey, R.T. (2008). Riboswitches: emerging themes in RNA structure and function. *Annu Rev Biophys* *37*, 117-133.
- Morris, M.C. (2009). Fluorescent biosensors of intracellular targets from genetically encoded reporters to modular polypeptide probes. *Cell Biochem Biophys* *56*, 19-37.
- Mulhbacher, J., St-Pierre, P., and Lafontaine, D.A. (2010). Therapeutic applications of ribozymes and riboswitches. *Curr Opin Pharmacol* *10*, 551-556.
- Nahvi, A., Barrick, J.E., and Breaker, R.R. (2004). Coenzyme B12 riboswitches are widespread genetic control elements in prokaryotes. *Nucleic Acids Res* *32*, 143-150.
- Nahvi, A., Sudarsan, N., Ebert, M.S., Zou, X., Brown, K.L., and Breaker, R.R. (2002). Genetic control by a metabolite binding mRNA. *Chem Biol* *9*, 1043.
- Navani, N.K., and Li, Y. (2006). Nucleic acid aptamers and enzymes as sensors. *Curr Opin Chem Biol* *10*, 272-281.
- Nimjee, S.M., Rusconi, C.P., and Sullenger, B.A. (2005). Aptamers: an emerging class of therapeutics. *Annu Rev Med* *56*, 555-583.
- Noinaj, N., Guillier, M., Barnard, T.J., and Buchanan, S.K. (2010). TonB-dependent transporters: regulation, structure, and function. *Annu Rev Microbiol* *64*, 43-60.

- Nomura, Y., and Yokobayashi, Y. (2007a). Dual selection of a genetic switch by a single selection marker. *Biosystems* 90, 115-120.
- Nomura, Y., and Yokobayashi, Y. (2007b). Reengineering a natural riboswitch by dual genetic selection. *J Am Chem Soc* 129, 13814-13815.
- Nou, X., and Kadner, R.J. (2000). Adenosylcobalamin inhibits ribosome binding to *btuB* RNA. *Proc Natl Acad Sci U S A* 97, 7190-7195.
- Nudler, E., and Mironov, A.S. (2004). The riboswitch control of bacterial metabolism. *Trends Biochem Sci* 29, 11-17.
- Ozawa, T. (2006). Designing split reporter proteins for analytical tools. *Anal Chim Acta* 556, 58-68.
- Ozsolak, F., and Milos, P.M. (2011). RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet* 12, 87-98.
- Patel, D.J., and Suri, A.K. (2000). Structure, recognition and discrimination in RNA aptamer complexes with cofactors, amino acids, drugs and aminoglycoside antibiotics. *J Biotechnol* 74, 39-60.
- Patel, D.J., Suri, A.K., Jiang, F., Jiang, L., Fan, P., Kumar, R.A., and Nonin, S. (1997). Structure, recognition and adaptive binding in RNA aptamer complexes. *J Mol Biol* 272, 645-664.
- Pfleger, B.F., Fawzi, N.J., and Keasling, J.D. (2005). Optimization of DsRed production in *Escherichia coli*: effect of ribosome binding site sequestration on translation efficiency. *Biotechnol Bioeng* 92, 553-558.
- Phillips, J.A., Lopez-Colon, D., Zhu, Z., Xu, Y., and Tan, W. (2008). Applications of aptamers in cancer cell biology. *Anal Chim Acta* 621, 101-108.
- Platt, T. (1986). Transcription termination and the regulation of gene expression. *Annu Rev Biochem* 55, 339-372.
- Reaves, M.L., and Rabinowitz, J.D. (2010). Metabolomics in systems microbiology. *Curr Opin Biotechnol* 22, 17-25.
- Regulski, E.E., and Breaker, R.R. (2008). In-line probing analysis of riboswitches. *Methods Mol Biol* 419, 53-67.
- Regulski, E.E., Moy, R.H., Weinberg, Z., Barrick, J.E., Yao, Z., Ruzzo, W.L., and Breaker, R.R. (2008). A widespread riboswitch candidate that controls bacterial genes involved in molybdenum cofactor and tungsten cofactor metabolism. *Mol Microbiol* 68, 918-932.
- Rfam-database (2011). Rfam database, S. Institute, ed.
- Robards, K., Jackson, P.E., Haddad, P.R., ed. (1995). *Principles and Practice of Modern Chromatographic Methods* (Academic Press (Elsevier Science)).
- Romig, T.S., Bell, C., and Drolet, D.W. (1999). Aptamer affinity chromatography: combinatorial chemistry applied to protein purification. *J Chromatogr B Biomed Sci Appl* 731, 275-284.
- Roth, J.R., Lawrence, J.G., and Bobik, T.A. (1996). Cobalamin (coenzyme B12): synthesis and biological significance. *Annu Rev Microbiol* 50, 137-181.
- Scarlett, F.A., and Turner, J.M. (1976). Microbial metabolism of amino alcohols. Ethanolamine catabolism mediated by coenzyme B12-dependent ethanolamine ammonia-lyase in *Escherichia coli* and *Klebsiella aerogenes*. *J Gen Microbiol* 95, 173-176.
- Sefah, K., Phillips, J.A., Xiong, X., Meng, L., Van Simaey, D., Chen, H., Martin, J., and Tan, W. (2009). Nucleic acid aptamers for biosensors and bio-analytical applications. *Analyst* 134, 1765-1775.

- Seiwert, S.D., Stines Nahreini, T., Aigner, S., Ahn, N.G., and Uhlenbeck, O.C. (2000). RNA aptamers as pathway-specific MAP kinase inhibitors. *Chem Biol* 7, 833-843.
- Serganov, A. (2009). Determination of riboswitch structures: light at the end of the tunnel? *RNA Biol* 7, 98-103.
- Shaner, N.C., Steinbach, P.A., and Tsien, R.Y. (2005). A guide to choosing fluorescent proteins. *Nat Methods* 2, 905-909.
- Sharma, C.M., and Vogel, J. (2009). Experimental approaches for the discovery and characterization of regulatory small RNA. *Curr Opin Microbiol* 12, 536-546.
- Sharma, V., Nomura, Y., and Yokobayashi, Y. (2008). Engineering complex riboswitch regulation by dual genetic selection. *J Am Chem Soc* 130, 16310-16315.
- Shendure, J., and Ji, H. (2008). Next-generation DNA sequencing. *Nat Biotechnol* 26, 1135-1145.
- Sinha, J., Reyes, S.J., and Gallivan, J.P. (2010). Reprogramming bacteria to seek and destroy an herbicide. *Nat Chem Biol* 6, 464-470.
- Smith, A.M., Fuchs, R.T., Grundy, F.J., and Henkin, T.M. (2010). The SAM-responsive S(MK) box is a reversible riboswitch. *Mol Microbiol* 78, 1393-1402.
- Soukup, G.A., and Breaker, R.R. (1999). Relationship between internucleotide linkage geometry and the stability of RNA. *RNA* 5, 1308-1325.
- Stulke, J. (2002). Control of transcription termination in bacteria by RNA-binding proteins that modulate RNA structures. *Arch Microbiol* 177, 433-440.
- Stupperich, E., Eisinger, H.J., and Krautler, B. (1988). Diversity of corrinoids in acetogenic bacteria. P-cresolylcobamide from *Sporomusa ovata*, 5-methoxy-6-methylbenzimidazolylcobamide from *Clostridium formicoaceticum* and vitamin B12 from *Acetobacterium woodii*. *Eur J Biochem* 172, 459-464.
- Sudarsan, N., Cohen-Chalamish, S., Nakamura, S., Emilsson, G.M., and Breaker, R.R. (2005). Thiamine pyrophosphate riboswitches are targets for the antimicrobial compound pyrithiamine. *Chem Biol* 12, 1325-1335.
- Sudarsan, N., Hammond, M.C., Block, K.F., Welz, R., Barrick, J.E., Roth, A., and Breaker, R.R. (2006). Tandem riboswitch architectures exhibit complex gene control functions. *Science* 314, 300-304.
- Sudarsan, N., Lee, E.R., Weinberg, Z., Moy, R.H., Kim, J.N., Link, K.H., and Breaker, R.R. (2008). Riboswitches in eubacteria sense the second messenger cyclic di-GMP. *Science* 321, 411-413.
- Sudarsan, N., Wickiser, J.K., Nakamura, S., Ebert, M.S., and Breaker, R.R. (2003). An mRNA structure in bacteria that controls gene expression by binding lysine. *Genes Dev* 17, 2688-2697.
- Suess, B., Fink, B., Berens, C., Stentz, R., and Hillen, W. (2004). A theophylline responsive riboswitch based on helix slipping controls gene expression in vivo. *Nucleic Acids Res* 32, 1610-1614.
- Takeuchi, M., Nagaoka, Y., Yamada, T., Takakura, H., and Ozawa, T. (2010). Ratiometric bioluminescence indicators for monitoring cyclic adenosine 3',5'-monophosphate in live cells based on luciferase-fragment complementation. *Anal Chem* 82, 9306-9313.
- Taneoka, A., Sakaguchi-Mikami, A., Yamazaki, T., Tsugawa, W., and Sode, K. (2009). The construction of a glucose-sensing luciferase. *Biosens Bioelectron* 25, 76-81.
- Tang, H., Wang, Y., Nicholson, J.K., and Lindon, J.C. (2004). Use of relaxation-edited one-dimensional and two dimensional nuclear magnetic resonance spectroscopy to improve detection of small metabolites in blood plasma. *Anal Biochem* 325, 260-272.

- Thomason, M.K., and Storz, G. (2010). Bacterial antisense RNAs: how many are there, and what are they doing? *Annu Rev Genet* 44, 167-188.
- Topp, S., and Gallivan, J.P. (2008). Random walks to synthetic riboswitches--a high-throughput selection based on cell motility. *Chembiochem* 9, 210-213.
- Topp, S., and Gallivan, J.P. (2010). Emerging applications of riboswitches in chemical biology. *ACS Chem Biol* 5, 139-148.
- Topp, S., Reynoso, C.M., Seeliger, J.C., Goldlust, I.S., Desai, S.K., Murat, D., Shen, A., Puri, A.W., Komeili, A., Bertozzi, C.R., *et al.* (2010). Synthetic riboswitches that induce gene expression in diverse bacterial species. *Appl Environ Microbiol* 76, 7881-7884.
- Tu, S.C., and Mager, H.I. (1995). Biochemistry of bacterial bioluminescence. *Photochem Photobiol* 62, 615-624.
- Tucker, B.J., and Breaker, R.R. (2005). Riboswitches as versatile gene control elements. *Curr Opin Struct Biol* 15, 342-348.
- Tuerk, C., and Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249, 505-510.
- Vitreschak, A.G., Rodionov, D.A., Mironov, A.A., and Gelfand, M.S. (2004). Riboswitches: the oldest mechanism for the regulation of gene expression? *Trends Genet* 20, 44-50.
- Wachter, A. (2010). Riboswitch-mediated control of gene expression in eukaryotes. *RNA Biol* 7, 67-76.
- Wachter, A., Tunc-Ozdemir, M., Grove, B.C., Green, P.J., Shintani, D.K., and Breaker, R.R. (2007). Riboswitch control of gene expression in plants by splicing and alternative 3' end processing of mRNAs. *Plant Cell* 19, 3437-3450.
- Weinberg, Z., Barrick, J.E., Yao, Z., Roth, A., Kim, J.N., Gore, J., Wang, J.X., Lee, E.R., Block, K.F., Sudarsan, N., *et al.* (2007). Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. *Nucleic Acids Res* 35, 4809-4819.
- Weinberg, Z., Wang, J.X., Bogue, J., Yang, J., Corbino, K., Moy, R.H., and Breaker, R.R. (2010). Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol* 11, R31.
- Weissbach, H., and Brot, N. (1991). Regulation of methionine synthesis in *Escherichia coli*. *Mol Microbiol* 5, 1593-1597.
- Welz, R., and Breaker, R.R. (2007). Ligand binding and gene control characteristics of tandem riboswitches in *Bacillus anthracis*. *RNA* 13, 573-582.
- White, J.C., DiGirolamo, P.M., Fu, M.L., Preston, Y.A., and Bradbeer, C. (1973). Transport of vitamin B 12 in *Escherichia coli*. Location and properties of the initial B 12 -binding site. *J Biol Chem* 248, 3978-3986.
- Wickiser, J.K., Winkler, W.C., Breaker, R.R., and Crothers, D.M. (2005). The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. *Mol Cell* 18, 49-60.
- Wilson, D.S., and Szostak, J.W. (1999). In vitro selection of functional nucleic acids. *Annu Rev Biochem* 68, 611-647.
- Winkler, W., Nahvi, A., and Breaker, R.R. (2002a). Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* 419, 952-956.
- Winkler, W.C., and Breaker, R.R. (2003). Genetic control by metabolite-binding riboswitches. *Chembiochem* 4, 1024-1032.
- Winkler, W.C., and Breaker, R.R. (2005). Regulation of bacterial gene expression by riboswitches. *Annu Rev Microbiol* 59, 487-517.

- Winkler, W.C., Cohen-Chalamish, S., and Breaker, R.R. (2002b). An mRNA structure that controls gene expression by binding FMN. *Proc Natl Acad Sci U S A* *99*, 15908-15913.
- Winkler, W.C., Nahvi, A., Roth, A., Collins, J.A., and Breaker, R.R. (2004). Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* *428*, 281-286.
- Wittmann, C., Kromer, J.O., Kiefer, P., Binz, T., and Heinzle, E. (2004). Impact of the cold shock phenomenon on quantification of intracellular metabolites in bacteria. *Anal Biochem* *327*, 135-139.
- Wood, K.V., Lam, Y.A., and McElroy, W.D. (1989). Introduction to beetle luciferases and their applications. *J Biolumin Chemilumin* *4*, 289-301.
- Yan, H., and Chen, W. (2010). 3',5'-Cyclic diguanylic acid: a small nucleotide that makes big impacts. *Chem Soc Rev* *39*, 2914-2924.
- Zhang, J., Campbell, R.E., Ting, A.Y., and Tsien, R.Y. (2002). Creating new fluorescent probes for cell biology. *Nat Rev Mol Cell Biol* *3*, 906-918.
- Zhang, Y.Z., Naleway, J.J., Larison, K.D., Huang, Z.J., and Haugland, R.P. (1991). Detecting lacZ gene expression in living cells with new lipophilic, fluorogenic beta-galactosidase substrates. *FASEB J* *5*, 3108-3113.
- Zhao, W., Chiuman, W., Lam, J.C., McManus, S.A., Chen, W., Cui, Y., Pelton, R., Brook, M.A., and Li, Y. (2008). DNA aptamer folding on gold nanoparticles: from colloid chemistry to biosensors. *J Am Chem Soc* *130*, 3610-3618.
- Zhou, B., and Wang, B. (2006). Pegaptanib for the treatment of age-related macular degeneration. *Exp Eye Res* *83*, 615-619.
- Zimmermann, G.R., Jenison, R.D., Wick, C.L., Simorre, J.P., and Pardi, A. (1997). Interlocking structural motifs mediate molecular discrimination by a theophylline-binding RNA. *Nat Struct Biol* *4*, 644-649.