ABSTRACT

Title of Dissertation: ANALYSIS OF GENETIC REGULATORY MECHANISMS THAT CONTROL ETHANOLAMINE UTILIZATION IN *ENTEROCOCCUS FAECALIS*

Margo Page Gebbie, Doctor of Philosophy, 2017

Dissertation directed by: Associate Professor Wade C. Winkler, Department of Cell Biology and Molecular Genetics

In this project, we studied the genetic regulatory mechanisms that affect utilization of ethanolamine, an abundant compound in the gastrointestinal environment. In *Enterococcus faecalis*, the ethanolamine utilization (*eut*) gene cluster encodes for a two-component regulatory system (TCS), comprised of a histidine kinase, EutW, which autophosphorylates upon sensing EA, and a cognate response regulator, EutV, which dimerizes upon receiving the phosphoryl group from EutW and binds the nascent transcript to prevent premature transcription termination. This TCS is responsible for coupling sensing of ethanolamine to production of *eut* transcripts. However, clues from other organisms had previously suggested that adenosylcobalamin (AdoCbl) might also be an important genetic regulatory signal for the *E. faecalis eut* genes. Indeed, we discovered a novel trans-acting noncoding RNA (EutX) that contained an AdoCbl-responsive riboswitch. Our data demonstrated that

the riboswitch promotes a shortened form when cellular AdoCbl levels are replete. In contrast, a longer form of EutX is synthesized when AdoCbl levels are depleted. We demonstrated that structural motifs contained in the longer form of EutX act to sequester the EutV protein, preventing it from promoting transcription elongation of *eut* transcripts. These unexpected data revealed an important new type of regulatory mechanism for riboswitch RNAs. In support of this overall genetic regulatory model, we recapitulated the full genetic circuitry in a heterologous host. Using this system, we employed extensive site-directed mutagenesis to examine the functional importance of highly conserved EutV residues. This led to the identification of a cluster of positively charged residues, which we speculated are important determinants for RNA-binding activity. Consistent with this hypothesis, mutations of these residues resulted in loss of RNA-binding activity. Furthermore, we also explored whether the *eut* gene cluster was affected by additional genetic regulatory mechanisms. From these efforts, we concluded that oxygen is not a genetic regulatory feature of *eut* genes, in contrast to previously published speculation. However, we did find that it is likely to be repressed under conditions of high glucose. Therefore, these aggregate studies revealed new mechanisms of post-initiation genetic regulation, and showed how *E. faecalis* specifically controls expression of ethanolamine catabolism genes.

ANALYSIS OF GENETIC REGULATORY MECHANISMS THAT CONTROL ETHANOLAMINE UTILIZATION IN *ENTEROCOCCUS FAECALIS*

by

Margo Page Gebbie

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2017

Advisory Committee:
Associate Professor Wade C. Winkler, Chair
Associate Professor Jason D. Kahn
Associate Professor Paul Paukstelis
Associate Professor Nicole LaRonde
Professor Kevin S. McIver
Associate Professor Daniel Nelson

Publications

Gebbie, M., Kaval, K., Garsin, D., Winkler, W. *Enterococcus faecalis* employs microcompartments to utilize ethanolamine under aerobic and anaerobic conditions. *In preparation*

Gebbie, M., Garsin, D., Goodson J., Winkler, W. Antitermination determinants of the *Enterococcus faecalis* ethanolamine-responsive two component regulatory system. *In preparation*

DebRoy, S*., **Gebbie, M*.,** Ramesh, A., Goodson, J., Cruz, M., van Hoof, A., Winkler, W., and Garsin, D. (2014) A Riboswitch-containing sRNA Controls Gene Expression by Sequestration of a Response Regulator. ***Science.*** August 2014: **345** (6199), 937-940 (* Equal contribution)

Patel, D. K. **Gebbie, M. P.,** and Lee, V. T. (2014) Assesing RNA interactions with proteins by DRaCALA. ***Methods Enzymology.*** 549, 489-512.

Dedication

This dissertation is dedicated to my family: Sandra Gebbie, Peter Gebbie, Sarah Gebbie-Measeck, Kathy Gebbie, and Ian Measeck. Without their encouragement and love, and hint of competition, I would not be where I am today.

Acknowledgements

First, I thank my advisor Dr. Wade C. Winkler. His guidance has been invaluable to my development as a scientist. I am grateful to have had the opportunity to grow both as a scientist and a person under his guidance.

I am grateful to my committee members Drs. Kevin S. McIver, Jason D. Kahn, Paul Paukstelis, Nicole LaRonde, and Daniel Nelson. Their advice and encouragement have been crucial for directing my projects. I would also like to thank my "unofficial" committee member, Dr. Vincent T. Lee for fruitful scientific discussions, and some tough talks.

I thank the people whose work has contributed to this dissertation: Jonathan Goodson for RNA-Seq work, Drs. Danielle Garsin, Sruti Debroy, and Karan Kaval for TEM imaging, Steven L. Klupt for helpful FACS discussions, and Cordelia A. Weiss for microscopy imaging of bacterial cells.

I thank the many members of the Winkler and Lee laboratories who have provided useful tips and a sympathetic ear. The "tea time", snacks, and the occasional conical tubes of alcohol we've shared helped make this difficult journey that much easier.

My wonderful friends, Carey Stuart, Cordelia A. Weiss, Rachel Shahan, Sarah Ahlbrand, and Susan Park-Ochsner for our innumerable spontaneous dinners and lunches. You all have made my time in Maryland something I will never forget.

Finally, I thank Patrick T. Mills for toughing it out the last few years, especially the last 18 plus months of long-distance, and me venting constantly about how "science" hates me. This dissertation would not have happened without you.

Table of Content

Publications ii
Dedication iii
Acknowledgements iv
List of Tables vii
List of Figures viii
List of Abbreviations x
Chapter 1: Introduction and literature review 1
1.1 Regulatory RNA and post-initiation gene regulation 1
1.2 Genetic regulation of ethanolamine utilization in Enterococcus faecalis 9
1.2.1 Ethanolamine Utilization in Bacteria 10
Chapter 2: Discovery of a riboswitch-containing sRNA that controls gene expression through sequestration of a response regulator 25
2.1 Copyright notice 25
2.2 Introduction 25
2.3 Results and Discussion 31
2.3.3 Detection of transcripts produced from the *E. faecalis eutT-G* intergenic region 31
2.3.3 Sequence analysis of the *eutT-G* intergenic region reveals a new putative ANTAR substrate 36
2.3.3 Determination of termini of EutX RNAs 37
2.3.4 Deletion analysis of EutX 40
2.3.5 The EutX AdoCbl riboswitch can terminate transcription *in vitro* 41
2.3.6 The EutX P3P4 site is required for genetic regulation of the *eut* gene cluster 46
2.4 Conclusion 51
Chapter 3: Antitermination determinants of the *Enterococcus faecalis* ethanolamine-responsive two component regulatory system 56
3.1 Author contribution 56
3.2 Introduction 56
3.3 Results and Discussion 59
3.3.1 *E. faecalis* EutV/W TCS promotes antitermination in a heterologous host 59
3.3.2 EutX antagonizes EutV antitermination 64
3.3.3 Charged amino acids in ANTAR domain that may be important for antitermination 71
3.3.4 Putative positively charged patch on surface of the ANTAR domain 77
3.3.5 *nasF* ANTAR substrate and NasR cannot function in a heterologous host 80
3.3.6 *In vitro* analysis of the ANTAR domain 84
3.4 Conclusion 89
Chapter 4: *Enterococcus faecalis* can utilize ethanolamine under anaerobic and aerobic conditions 91
4.1 Author contribution 91
4.2 Introduction 91
4.3 Results and Discussion 92

4.3.1 *E. faecalis* utilizes EA aerobically 92
4.3.2 Complex media components inhibit aerobic EA utilization 98
4.3.3 Visualization of Eut microcompartments 100
4.3.4 Decoupling EutX from the *eut* pathway enhances growth and increases Eut microcompartment formation 110
4.4 Conclusion 114
Chapter 5: Summary of primary findings and future directions 122
5.1 Summary of Primary Research findings 122
5.2 Sensing AdoCbl and EA, and eut Gene Cluster regulation 124
5.3 Biological Relevance of Ethanolamine Utilization in E. faecalis 126
5.4 Conclusion 128
Chapter 6: Materials and Methods 130
6.1 Chemicals and oligonucleotides 130
6.2 Bacterial strains and media 130
6.2.1 Growth conditions 130
6.2.2 Antibiotic concentrations 131
6.2.3 Amino acid concentrations for auxotrophs 131
6.2.4 Inducer concentrations 131
6.3 Strain construction 132
6.3.1 Construction of *amyE*:: P*const* P1P2T-YFP 132
6.3.2 Construction of *thrC*:: P_{lial}*I* EutVW 132
6.3.3 Construction of inducible EutX 132
6.3.4 Construction of inducible GlmS-P3P4 133
6.3.5 Construction of *thrC*:: P_{lial} NasR 133
6.3.6 Construction of NasR expression vector 133
6.3.7 Genetic transformation of *E. coli* 134
6.3.8 Genetic transformation of *B. subtilis* 134
6.4 RNA extraction 135
6.5 Northern blot analyses 135
6.6 Rapid Amplification of cDNA Ends (RACE) 136
6.6.1 With 5' adenylated linkers 136
6.6.2 From circularized RNA transcripts 137
6.7 In vitro transcription assays 138
6.8 Quantitative real-time RT-PCR 139
6.9 RNA–binding studies using electrophoretic mobility shift assays (EMSA) or differential radial capillary action of ligand assay (DRaCALA) 139
6.10 Protein expression and purification 140
6.11 Growth curves 141
6.11.1 E. faecalis 141
6.11.2 Aerobic 141
6.11.3 Anaerobic 141
6.12 Fluorescence microscopy 141
6.13 Flow cytometry 142
Appendices 143
Bibliography 166

List of Tables

Table 1. Composition of chemically defined media (CDM) 93
Table 2. Analysis of EFKK6, OG1RF, and ΔEutX growth in CDM..... 98
Table 3. Analysis of EFKK6 growth in CDM supplemented with yeast extract..... 100
Table 4. Quantification of microcompartments in EFKK6, OG1RF, and ΔEutX.... 112
Table 5. List of oligonucleotides used in these studies..... 143
Table 6. List of gBlocks used in these studies..... 154
Table 7. List of plasmids used in these studies..... 159
Table 8. List of strains used in these studies..... 161

List of Figures

Figure 1-1. Different modes of action for *trans*-acting regulatory RNAs 3
Figure 1-2. Different modes of action for *cis*-acting RNAs 8
Figure 1-3. *S. enterica* subsp. Typhimurium *eut* regulation 12
Figure 1-4. Schematic overview of characterized bacterial microcompartments..... 14
Figure 1-5. Model for ethanolamine catabolism in *E. faecalis* 17
Figure 1-6. Two-hairpin motif ANTAR substrates in *E faecalis eut* gene cluster 21
Figure 1-7. Model for post-initiation control of gene expression in the *eut* gene cluster of *E. faecalis* 24
Figure 2-1. EA and AdoCbl induce *eut* gene expression and microcompartment formation..... 29
Figure 2-2. Northern blot analysis of the *eutT-eutG* intergenic region show two RNA species 32
Figure 2-3. Northern blot analysis of the *eutT-eutG*..... 34
Figure 2-4. Analysis of EutX termini 38
Figure 2-5. Northern blot analysis of deletion strains..... 42
Figure 2-6. AdoCbl binding to the riboswitch causes premature 44
Figure 2-7. EutV specifically binds EutX..... 47
Figure 2-8. EutX acts in *trans* to sequester EutV 48
Figure 2-9. Model of the regulation of the *eut* gene cluster by EutX 52
Figure 2-10. Updated summary of different modes of action for *cis*-acting RNAs ... 55
Figure 3-1. Schematic for expression of an *E. faecalis* EutV antitermination circuit in a heterologous host 61
Figure 3-2. High EA inhibits growth of *B. subtilis*..... 63
Figure 3-3. *E. faecalis* EutV/W TCS promotes antitermination in a heterologous host 65
Figure 3-4. EutX antagonizes EutV antitermination in a heterologous host 67
Figure 3-5. Northern blot analysis of EutX expression in *B. subtilis* 68
Figure 3-6. Heterologous expression of EutX antagonizes EutV 70
Figure 3-7. Alignment of putative EutV homologues 72
Figure 3-8. *In vivo* antitermination activity of mutant EutV proteins..... 74
Figure 3-9. Electrostatic surface potential maps of AmiR, NasR, and Rv1626 X-ray crystal structures 76
Figure 3-10. Electrostatic surface potential of AmiR, NasR, and Rv1626 X-ray crystal structures..... 78
Figure 3-11. Electrostatic surface potential maps of putative EutV homologues..... 79
Figure 3-12. R142K mutation does not restore antitermination 81
Figure 3-13. *nasF* RNA and NasR ANTAR protein do not support antitermination in a heterologous host 83
Figure 3-14. *nasF* RNA binds specifically to NasR 86
Figure 3-15. Mutation of NasR disrupts binding to *nasF* RNA 88
Figure 4-1. Growth of *E. faecalis* in CDM 95
Figure 4-2. *E. faecalis* utilizes ethanolamine aerobically in CDM..... 97
Figure 4-3. Yeast extract inhibits *E. faecalis*' ability to utilize EA..... 99

Figure 4-4. Endogenously integrated *eutM:gfp* hinders Eut microcompartment formation 102
Figure 4-5. TEM of the endogenously integrated *eutM:gfp* shows septal defects 103
Figure 4-6. Fluorescence microscopy of EFKK6 during anaerobic growth 105
Figure 4-7. TEM images of EFKK6 107
Figure 4-8. WT *E. faecalis* supports growth in the presence of EA aerobically 109
Figure 4-9. TEM images of OG1RF 111
Figure 4-10. EutX deletion supports enhanced growth in EA 113
Figure 4-11. TEM images of *eutX* mutant 115
Figure 4-12. EA alone induces microcompartment formation in a *eutX* mutant 116
Figure 4-13. Glucose inhibits EA utilization in OG1RF 118
Figure 5-1. Strategies of *eut* gene regulation in multiple organisms 125
Figure 5-2. Final model for *E. faecalis eut* gene regulation 127

List of Abbreviations

A	Adenosine
AdoCbl	Adenosylcobalamin
ANTAR	AmiR and NasR transcription antitermination regulator domain
AsRNA	antisense small regulatory RNA
AT	Antiterminator
bp	basepair
BME	β-mercaptoethanol
BMC	Bacterial Microcompartment protein domain
BMC-H	BMC-Hexamer
BMC-P	BMC-Pentamer
BMC-T	BMC-Trimer
C	Cytosine
CA	Carbonic Anhydrase
CBB	Calvin-Benson-Bassham Cycle
CCR	Carbon Catabolite Repression
CDM	Chemically Defined Media
CoA	Coenzyme A
DH	Dehydrogenase
DNA	Deoxyribose Nucleic Acid
DNA-seq	DNA sequencing
DRaCALA	Differential Radical Capillary of Action Ligand Assay

EA	Ethanolamine
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic Mobility Shift Assay
EtBr	Ethidium Bromide
fM	Femtomolar
G	Guanosine
GI	Gastrointestinal
gBlock	Gene Block, IDT
IPTG	Isopropyl- β-thiogalactopyranoside
kDa	Kilo Dalton
LB	Luria-Bertani, Luria Broth, Lennox Broth, or Lysogeny Broth
MCP	Bacterial Microcompartment
mM	Millimolar
mRNA	Messenger RNA
NAD^+	Nicotinamide Adenine Dinucleotide, oxidized
NADH	Nicotinamide Adenine Dinucleotide, reduced
nM	Nanomolar
nt	Nucleotide
OD	Optical Density
oligo	Oligonucleotide
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction

PE	Phosphatidylethanolamine
PGA	Phosphoglycerate
pM	Picomolar
qRT-PCR	Quantitative Reverse Transcriptase PCR
PTAC	Phosphotransacetylase
RACE	Rapid Amplification of cDNA Ends
RBS	Ribosome Binding Site
RNA	Ribonucleic Acid
RNAP	RNA Polymerase
RNase	Ribonuclease
RNA-seq	RNA sequencing
rRNA	Ribosomal RNA
RT	Room Temperature or Reverse Transcriptase
RuBP	Ribulose 1,5-bisphosphate
SDS	Sodium Dodecyl Sulfate
sRNA	Small RNA
T	Thymidine
TCS	Two-Component System
TEM	Transmission Electron Microscopy
TBAB	Tryptose Blood Agar Base
TSS	Transcription Start Site
U	Uridine
UTR	Untranslated region

WT	Wild-type
YFP	Yellow Fluorescent Protein

Chapter 1: Introduction and literature review

The ability to sense and respond to sudden changes within an environmental niche is essential for all microorganisms. Therefore, microorganisms have evolved the ability to respond to environmental changes by specifically altering expression of certain genes. The majority of these strategies occur at the level of transcription initiation (1); however, recent studies have found many examples of RNA-mediated mechanisms that affect gene expression at the post-initiation level (2, 3). In this introduction, I will review previous work on representative regulatory RNAs and the mechanisms by which they regulate gene expression. I will focus in particular on how RNA-mediated control relates to the *e*thanolamine *ut*ilization (*eut*) gene cluster in the Gram-positive bacterium *Enterococcus faecalis.*

1.1 Regulatory RNA and post-initiation gene regulation

In the past 15 years, many small (~70-500 nt) *trans*-acting noncoding RNAs ('sRNA'), have been discovered in bacteria (2). Regulatory RNAs employ diverse mechanisms to moderate expression of their targets, functioning primarily at the post-transcriptional level. By using mRNAs as targets, they can affect transcription, translation, RNA stability, maturation, and processing (2). In certain instances, proteins can be sequestered away from their biological function by sRNAs, consequently causing numerous downstream effects (4–8).

It has been estimated that the number of different sRNAs in a given bacterium is likely to be approximately equal to the total number of transcription factors, numbering between ~100-300 (9). Many of these sRNAs are individually produced in

response to different cellular stress signals. Upon induction, some associate with a single target mRNA while others target a regulon of different transcripts, oftentimes through common seed sequence interactions (10). Although sRNAs have been discovered in diverse bacterial species (11) their molecular mechanisms have been best characterized in the Gram-negative bacteria *Escherichia coli* and *Salmonella* (12, 13).

1.1.1 *Trans*-acting Regulatory RNAs: antisense RNAs

sRNAs can in principle utilize different types of mechanisms to regulate gene expression. Antisense RNAs (asRNAs) can be classified as *cis*- or *trans*- effectors, depending on whether they are transcribed from the opposite DNA strand of the gene that they regulate or from an independent, distal locus in the genome, respectively (Figure 1-1). In the case of the *cis*-asRNAs, complementarity with the target mRNA is extensive, while *trans*-asRNAs tend to exhibit only short stretches of complementarity with their targets. In both cases the interactions between the sRNAs and their target mRNAs leads to either inhibition or enhancement of mRNA translation initiation (*e.g., Bacillus subtilis* SR1, *Staphyloccus aureus* RNAIII) and/or mRNA degradation (*S. aureus* RNAIII) (14–21). In contrast to *cis*-asRNAs that usually target one mRNA, *trans*-asRNAs can target multiple mRNAs. *Cis*-asRNAs are usually constitutively expressed and are primarily associated with mobile genetic elements such as plasmids, phages, and transposons. It is here that *cis*-asRNAs regulate plasmid copy number, conjugation, phage life cycle, and transposition (22–24). Several *cis*-asRNAs also act as antitoxins (25–27). Expression of the majority of

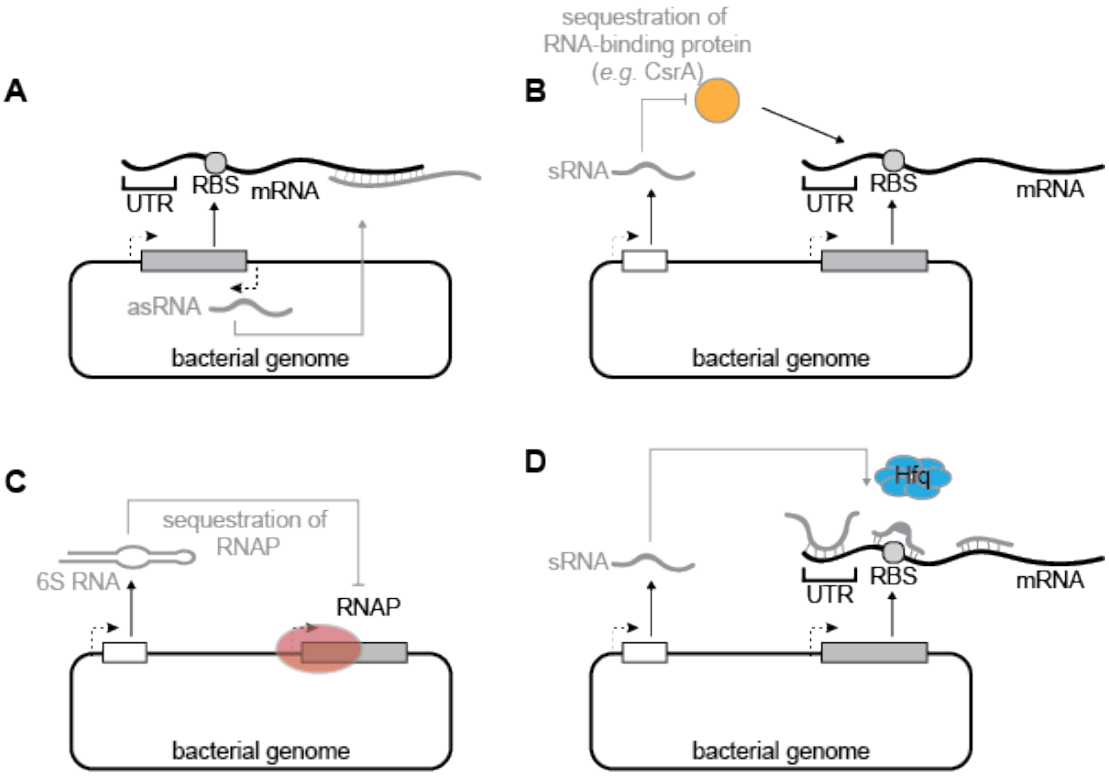

Figure 1-1. Different modes of action for *trans*-acting regulatory RNAs. Shown in (A) is an example of 'antisense RNA' which is typically encoded at the same genomic locus as its target, but is transcribed from the opposite DNA strand. Antisense RNAs form longer and more extensive base-pairing than *trans*-encoded RNAs with their target mRNAs; they also do not usually require Hfq. Examples of protein-sequestering *trans*-acting RNAs in (B-C). This particular class of RNAs can indirectly modulate gene expression by (B) sequestering specific RNA-binding regulatory proteins normally required for genetic regulation (*e.g.*, CsrA and RsmA) or (C) direct binding to RNA polymerase (*e.g.,* under stationary phase condition), thereby preventing it from initiating transcription. Shown in (D) is a more common example of *trans*-acting RNAs that form imperfect base-pairing with sequences within the untranslated region, near the ribosomal binding site, or even within the coding region of transcripts. In most cases, a chaperone protein, Hfq, is required for stability and activity of these RNAs *in vivo*.

the chromosomally-encoded asRNAs are usually induced in response to stress or growth-specific conditions (2, 28, 29). Many of these asRNAs have been shown to regulate metabolic pathways (29), sugar and carbon source utilization (5), and the composition of the bacterial membranes (30–32). Several asRNAs have also been found expressed in pathogenic bacteria and their specific roles have been recently elucidated (33–35).

1.1.2 *Trans*-acting Regulatory RNAs: protein-sequestering RNAs

Regulatory RNAs also have the ability to bind to proteins and thus regulate their activities (Figure 1-1A). This protein sequestration capability has been demonstrated for the *E. coli* CsrB sRNA (targeting the global carbon storage regulatory protein, CsrA), and 6S RNA (targeting σ^{70} RNA polymerase) (4, 5, 36, 37). These sRNAs are both able to mimic structures of the natural substrates for the target proteins and, by doing so, counteract their activity. For example, the *E. coli* 6S RNA mimics an open promoter for association with Sigma 70-RNA polymerase holoenzyme, and therefore can titrate transcription complexes away from promoters, resulting in a global decrease in transcription. Control of gene expression by 6S RNA has been shown to occur in response to the shift from exponential to stationary phase growth, suggesting that 6S RNA is an important factor in bacterial adaptation processes (37). While CsrB/C and 6S RNAs have been best characterized in *E. coli*, the model Gram-positive bacterium *Bacillus subtilis* appears to exhibit differences. For example, no protein sequestration mechanism like CsrA/B has been identified. Also, two separate 6S RNA transcripts have been identified, which are differentially

regulated by growth phase, although their regulatory effects have not yet been studied (Figure 1-1B) (38, 39).

In *E. coli* and many Gram-negative bacteria, most of the sRNAs that have been mechanistically studied require an RNA-binding protein, Hfq, for their regulatory functions (Figure 1-1C). *E. coli* Hfq is a small (8-15 kDa) protein that self-assembles into a doughnut-shaped homohexameric complex (40, 41). Multiple RNA-binding regions are likely to be important for Hfq function as it associates simultaneously with an sRNA and its mRNA target (43, 44). Additionally, the C-terminal extension of the *E. coli* Hfq constitutes an RNA interaction surface with specificity for mRNAs (45). Interestingly, this C-terminal extension is lacking in *S. aureus*, therefore suggesting the protein has lost some RNA binding properties. Hfq extends the intracellular half-life of a majority of *E. coli* sRNAs, but can result in shortening of the half-life of target mRNAs after sRNA-bound Hfq associates with them (46). However, the role of Hfq in many Gram-positive bacteria is still not well understood. Indeed, Hfq has been shown to be dispensable in *B. subtilis* (15), as well as other Gram-positive bacteria (47, 48). Also the relatively few sRNAs characterized in these microbes do not require Hfq for stability or activity (49–51).

In general, the aggregate data on Hfq and CsrA proteins, and the RNAs that bind to them, have revealed important information on how RNA binding proteins can be generally utilized in bacteria.

1.1.3 *Cis*-acting Regulatory RNAs

In addition to *trans*-acting sRNAs, *cis*-acting regulatory RNAs are prevalent in bacteria and are most commonly found in the 5' leader region of mRNAs; however, they can also be found within intercistronic regions or in the 3' UTR of a transcript (52, 53). Functionally important *cis*-acting regulatory RNAs have been identified most effectively by comparative sequence analyses. This approach can be used to identify RNA primary sequences and secondary structures that exhibit significant evolutionary conservation (54–56).

Some *cis*-acting regulatory RNAs alter downstream gene expression in response to interaction with a particular signaling ligand. These signal-responsive, *cis*-acting regulatory RNAs are called riboswitches (3, 7, 57–59). Many different riboswitches have been discovered in bacteria. The signaling ligands that trigger regulation by riboswitches range from RNA-binding proteins (3, 58), to *trans*-RNAs (60), to small molecule metabolites or metal ions (7, 61). There are also 'orphan' riboswitches, which exhibit conservation and phylogenetic patterns that closely resemble proven riboswitches, but for whom signaling ligands have not yet been identified (3). Riboswitch regulatory RNAs utilize conserved features to fold into elaborate three-dimensional structures that bind the cognate ligands with high specificity and affinity. Riboswitch RNAs consist of two conceptually distinct portions: a ligand-binding domain, commonly referred to as the aptamer domain, and an expression platform that is involved in coupling ligand-induced conformational changes to genetic control (62).

Riboswitches have been discovered that use a range of different molecular mechanisms for genetic regulation. The most commonly observed examples couple

binding of the appropriate signaling ligand to control of premature transcription termination or the efficiency of translation initiation (Figure 1-2A and 1-2B) (3, 63). For the former, binding of the ligand to the aptamer domain controls formation of a termination site, usually by regulating the formation of a competing, mutually exclusive secondary structure (*i.e.,* antiterminator) (3). Analogously, mutually-exclusive base paired structures are exploited by riboswitches to control ribosome access to the ribosome binding site (RBS) or Shine-Dalgarno (SD) sequence, thus regulating translation initiation (Figure 1-2C) (58). Some riboswitches affect access to an RNase enzyme for signal-dependent control of mRNA stability (64). Finally, one particular riboswitch class (*glmS*) functions as a metabolite-dependent, self-cleaving ribozyme, which, upon self-cleavage, targets the downstream transcript for destruction (65, 66). For most of these mechanisms, the riboswitch RNAs couple a swap in Watson-Crick base-pairing schemes to genetic elements that affect downstream gene expression.

Many riboswitches promote transcription attenuation (Figure 1-2A) (67), wherein a metabolic signal is received by the aptamer domain, which in turn stimulates the formation of a transcription termination signal. Oftentimes, in the termination signal is comprised of an intrinsic terminator helix, which upon formation turns downstream gene expression "OFF". However, in the absence of the metabolic signal an alternate 'antiterminator' conformation is formed, which actively prevents formation of the termination site. It is the ligand-controlled interchange between terminator and antiterminator helices that ultimately dictates downstream transcription levels. For some riboswitches, the default conformation is one that

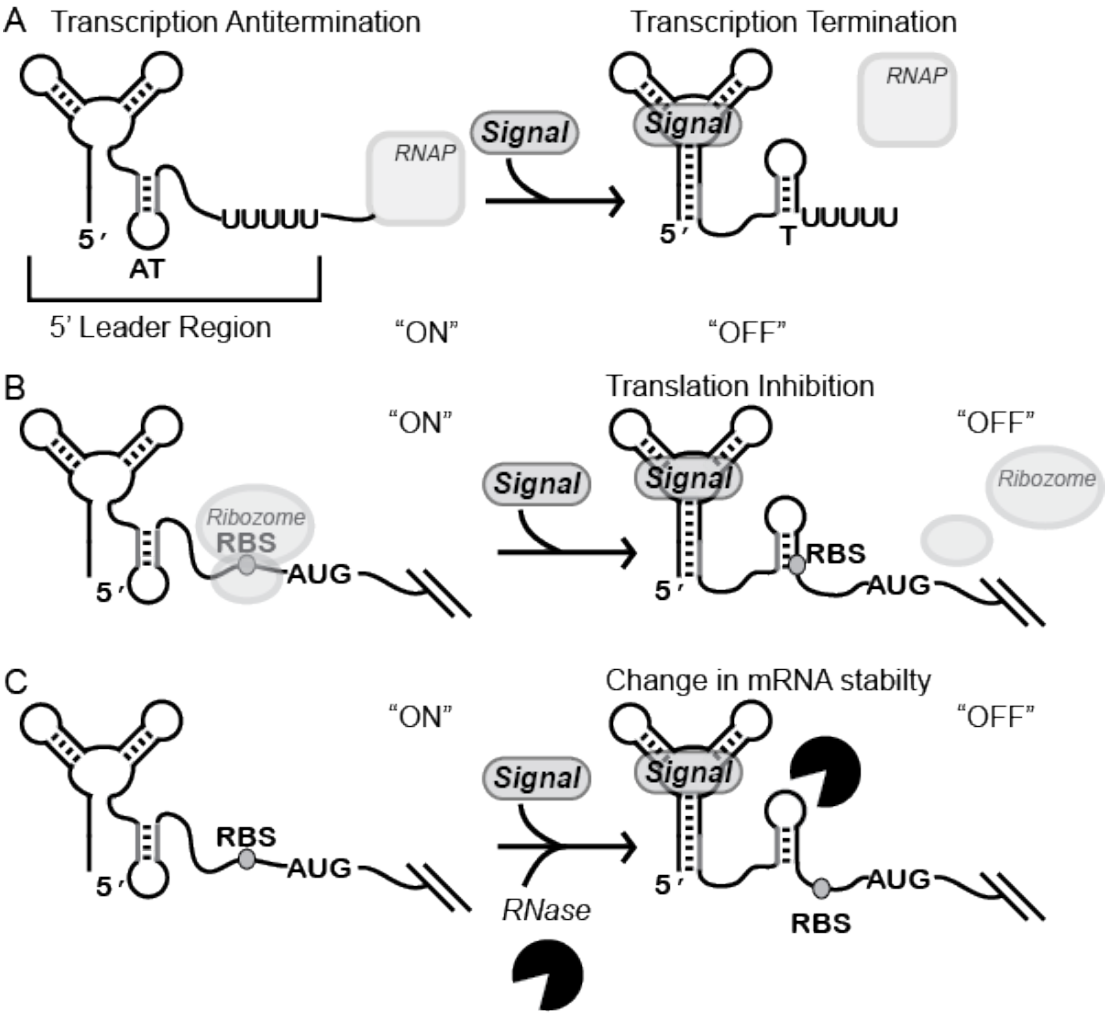

Figure 1-2. Different modes of action for *cis*-acting RNAs. (A) Control of formation of an intrinsic transcription terminator by a metabolite-binding riboswitch. (B) Control of translation initiation efficiency by a metabolite-sensing riboswitch. (C) Control of mRNA stability by conformational change of a riboswitch upon binding its ligand. AT, Antiterminator; T, terminator; RBS, ribosome binding site.

promotes a transcription terminator site and, for those riboswithces, an antiterminator configuration is promoted by binding of the metabolic signal to the aptamer domain, thereby turning gene expression "ON". The outcome of these mechanisms is a single choice between terminator and antiterminator elements during the process of transcription. The outcome of this "choice" between terminator/antiterminator helices oftentimes depends on the precise coordination of multiple processes, including the nascent RNA folding pathway, ligand-binding kinetics, and the kinetics of transcription. Indeed, it has been shown in at least one example that coordination between transcription and ligand-binding kinetics is important for gene regulation by a number of different regulatory RNAs (68–70).

The remaining modes of regulation by which riboswitches regulate genes are through translation inhibition and mRNA stability (Figure 1-2B and 1-2C). Similar to transcription termination, ligand association influences formation of nucleobase pairing interactions between alternate helical pairings. However, instead of controlling terminator formation, the RBS is either obstructed from ribosomal access or rendered more accessible once the riboswitch is bound by ligand (3, 7). Lastly, some riboswitches have also been shown to regulate mRNA stability. For these riboswitches, binding of cognate ligand promotes an alternate structure that allows access to an RNase enzyme, triggering degradation of the mRNA transcript (71).

1.2 Genetic regulation of ethanolamine utilization in Enterococcus faecalis

The ability to survive in an otherwise hostile environment, such as the gastrointestinal (GI) tract, requires metabolic flexibility. Thus, many gut

microorganisms, such as *Enterococcus faecalis*, have evolved the ability to utilize a number of energy sources including ethanolamine (EA), which is richly prevalent in the GI tract environment (72–74). In this project, we analyzed several types of regulatory RNAs that are required for genetic regulation of *E. faecalis* ethanolamine utilization genes.

1.2.1 Ethanolamine Utilization in Bacteria

Enterococcus faecalis is a nosocomial pathogen, prevalent in the hospital environment. However, its most common lifestyle is as a commensal in the GI tract of mammals. Conversely, in immune-compromised patients this close bacteria-host relationship, coupled with *E. faecalis*' ability to acquire multiple antibiotic resistances, can promote pathogenesis (75). To thrive in the GI tract environment, enterococci have evolved the ability to metabolize various energy sources, including EA. EA is the byproduct of the catabolism of phosphatidylethanolamine (PE), an abundant phospholipid found in both mammalian and bacterial membranes (76). Therefore, the combination of both types of cells within the intestine (bacterial and epithelial) are thought to provide a rich reservoir of EA (77). Interestingly, several bacterial species found within the GI tract, including but not limited to *E. faecalis*, *E. coli*, *Listeria*, *Salmonella*, and *Clostridium* species, encode the genes necessary for EA catabolism.

For those organisms that can utilize EA, the central genes in the process of EA catabolism are *eutB* and *eutC*, the protein products of which form the enzyme ethanolamine ammonia lyase. Additional supplemental genes, as many as sixteen, are

often found associated with *eutBC*. They also often include genes that encode for alcohol and aldehyde dehydrogenase, and phosphotransacetylase (PTAC) (78). The gene clusters for EA utilization have been phylogenetically defined into one cluster with three sub-clusters: EUT1, EUT2, and EUT3 (79). EUT1 loci are found in enterobacteria and are represented by the experimentally characterized EA utilization operon in *Salmonella enterica* (*S. enterica*). The EUT2 loci, however, are not as clearly defined, due to more variety between the loci. EUT2 sub-types are most commonly found in Firmicutes. Finally, EUT3 are found only in two strains of *Desulfitobacterium hafniense* (79). For the EUT1 subfamilies, a DNA-binding transcription factor called EutR is routinely associated with the *eut* gene cluster. However, the *eutR* gene is absent in the EUT2 subfamily and instead the gene cluster appears to encode for a two-component regulatory system. Otherwise, all of the *eut* subfamilies largely encode for a similar set of proteins, with only a few few exceptions (79). EUT2 loci usually encode for a PduL-like PTAC, while EUT1 loci encode the *pta*-like PTAC, EutD. Additionally, nearly half of the EUT2 loci lack a complete cores set of proteins; only sub-type EUT2A encodes for the EutG alcohol dehydrogenase, which is responsible for converting an aldehyde to an alcohol. *E. faecalis* falls into this last sub-class of EA utilization microcompartment gene clusters, EUT2A.

EA utilization has been most extensively studied in *S. typhimurium*. In this organism, 17 genes encode the proteins involved in EA catabolism and are located together on the chromosome in the *eut* operon (77, 80–82) (Figure 1-3). EutR is the positive regulator of the *eut* operon. Prior results, primarily from the J. Roth

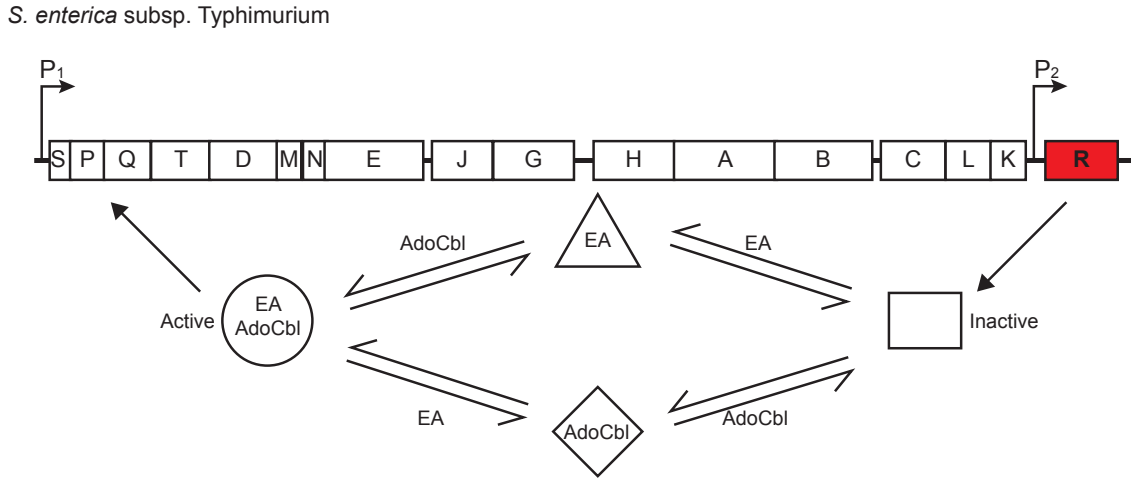

Figure 1-3. *S. enterica* subsp. Typhimurium *eut* regulation. Top shows the *eut* operon with two promoters, P_1 and P_2, and the gene for transcription factor EutR highlighted in red. This model, adapted from Roof and Roth 1992, proposes that each effector binds independently. Generally, both - effectors must be present to stabilize the active conformation and ensure sufficient activator for induction of the operon. In cells with a high level of EutR, a single effector may generate a large enough pool of an intermediate conformation to allow for a significant level of activation.

laboratory, established that the primary promoter at the beginning of the *eut* operon is induced in the presence of both EA and adenosylcobalamin (AdoCbl) (83). However, a second promoter, found directly in front of the *eutR* gene, providesconstitutive low-level expression. While induction of the operon by EutR requires both EA and AdoCbl, high-levels of EutR have been shown to cause partial induction with only one of the inducers, and an increased maximal expression overall with both EA and AdoCbl (83). The exact mechanisms by which AdoCbl and EA activate EutR have yet to be elucidated. However, prior experimentation suggested that EutR binds directly and simultaneously to both of these compounds, presumably causing a conformational change that allows for DNA binding and transcriptional activation by EutR (84, 85).

1.2.2 Bacterial Microcompartments and Ethanolamine Catabolism

One characteristic of EA utilization is that all the essential enzymes reside in a multi-protein complex termed the bacterial microcompartment. In EA catabolism, the formation of a microcompartment is believed to help in retaining the volatile intermediate, acetaldehyde, to prevent the loss of this source of carbon and/or protect the cell from its potential toxic effects (86, 87).

In general, bacterial microcompartments are proteinaceous organelle-like structures, involved in autotrophic or heterotrophic metabolism (78, 88). There are two subtypes of microcompartments: carboxysomes (Figure 1-4), which are involved in carbon fixation, and catabolic metabolosomes (Figure 1-4), which share a common catalytic core of enzymes while processing diverse substrates using subtype-specific

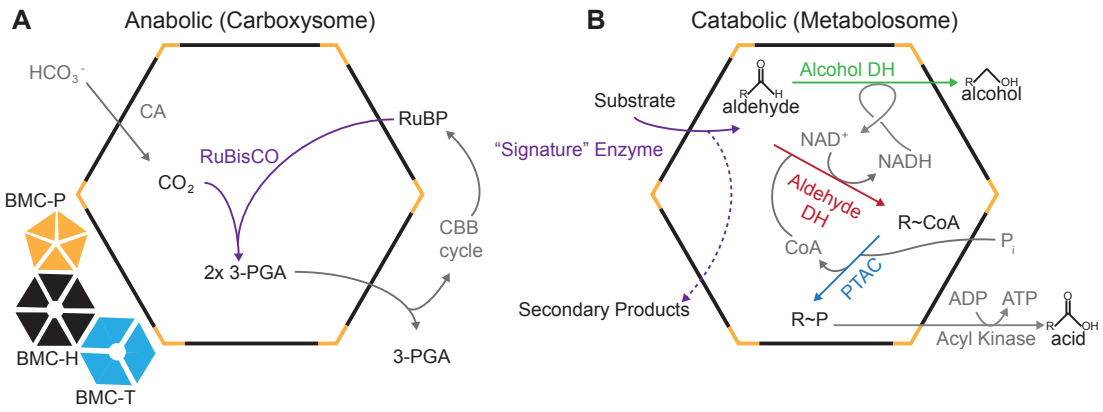

Figure 1-4. Schematic overview of bacterial microcompartments classes. (A) Carboxysome. (B) Metabolosome. An example substrate is EA and the signature enzyme, ethanolamine ammonia lyase, produces acetaldehyde and ammonia, a secondary product. Reactions in grey are peripheral to the core microcompartment chemistry. Microcompartment shell protein oligomers are depicted on the left: black, BMC-H; blue, BMC-T; yellow, BMC-P. CA, carbonic anhydrase; 3-PGA, 3-phosphoglycerate; RuBP, ribulose 1.5-bisphosphate; DH, dehydrogenase; PTAC, phosphotransacetylase.

signature enzymes. The microcompartment shell creates an environment ideal for metabolite channeling, sequestering toxic or volatile chemical intermediates, and possibly maintaining anaerobic conditions within the microcompartment (89–92). Microcompartments are comprised of thousands of subunits including hexameric, trimeric, and pentameric structural proteins. These proteins form a shell that encapsulates a catalytic core. The structural proteins of the microcompartment are paralogous to the shell proteins of the carboxysome, a bacterial organelle-like structure that concentrates CO_2 for fixation by ribulose-bisphosphate carboxylase in cyanobacteria and many other bacteria (93, 94). The proteinaceous shells of microcompartments are similar in that they contain similar bacterial microcompartment protein domains (BMC). Indeed, the microcompartments simply differ from carboxysomes in the identities of their core enzymes, which are encoded from genes that are typically adjacent to shell protein genes within the same gene cluster. Bacterial microcompartments are predicted to be present across 23 bacterial phyla, and are represented by as many as 23 different types (79, 95). Subtypes of metabolosomes contain different core enzymes that perform different metabolic functions. There are as many as eight subtypes of catabolic metabolosomes: 1,2-propanediol utilizing (PDU), Ethanolamine utilizing (EUT), Glycyl radical enzyme-containing microcompartment (GRM), Planctomycetes and Verrucomicrobia-Type (PVM), PVM-like, Ethanol-utilizing (ETU), Metabolosomes of Unknown Function (MUF), and *Rhodococcus* and *Mycobacterium* microcompartment (RMM) (79). Some individual subtypes may be found in multiple organisms. Alternatively, one organism can have more than one microcompartment gene cluster, thereby potentially encoding

for multiple subtypes, such as *Clostridium saccharolyticum*, which encodes for GRM, EUT, and PVM-like microcompartments (79). In certain instances, the functions of microcompartments have been associated with a variety of microbial activities, including biological nutrient cycling, pathogenesis, and symbiosis (72, 96–99).

The EA microcompartment shell is constructured from five different Eut proteins, encoded by *eutS*, *eutL*, *eutK*, *eutM*, and *eutN* genes (82). The ethanolamine ammonia lyase, EutBC, cleaves EA to acetaldehyde and ammonia; however, it is not yet known whether EutBC is packaged inside the microcompartment or if it resides outside, shuttling its byproducts into the lumen of the microcompartment. Some recent data suggested the latter (100). This initial breakdown of EA requires AdoCbl as a cofactor, which is salvaged by a corrinoid cobalamin adenosyltransferase encoded by *eutT* (101, 102). *EutBC*'s requirement for AdoCbl may explain why *Salmonella* EutR senses AdoCbl as an important signal for genetic regulation of the *eut* cluster (83). The acetaldehyde produced in the first enzymatic reaction is then converted to acetyl-CoA by the acetaldehyde dehydrogenase, EutE (72). Acetyl-CoA is subsequently used in several metabolic processes including lipid biosynthesis. Conversely, acetyl-CoA can also be converted into acetylphosphate by the phosphotransacetylase, EutD (72). The housekeeping acetate kinase, AckA, then carries out substrate level phosphorylation to make acetate from acetylphosphate, with the production of an ATP molecule. Alternatively, acetaldehyde can be converted to ethanol by the alcohol dehydrogenase, EutG, thus preventing the buildup of toxic levels of acetaldehyde (72) (Figure 1-5).

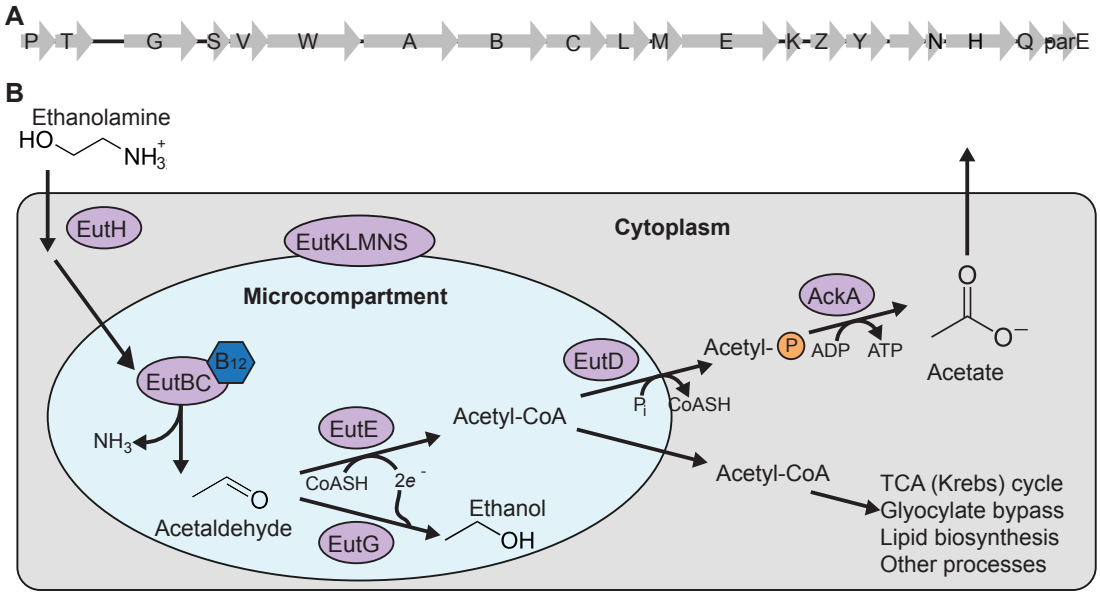

Figure 1-5. Model for EA catabolism in *E. faecalis*. (A) *E. faecalis eut* gene cluster. (B) EA enters the cell through diffusion or with the help of EA utilization protein H (EutH). In the EA-specific microcompartment (Eut), the structural components of which are EutK, EutL, EutM, EutN, and EutS (EutKLMNS), EA is degraded to acetaldehyde and ammonia (NH_3) by the EA ammonia lyase (EutBC), which requires the cofactor AdoCbl. Acetaldehyde can be catabolized to ethanol by EutG (an alcohol dehydrogenase), or to the metabolically useful compound, acetyl-CoA, by EutE (an acetaldehyde dehydrogenase). Acetyl-CoA can be used in a range of metabolic processes or converted to acetylphosphate by EutD (a phosphotransacetylase). The housekeeping acetate kinase (AckA) can generate ATP and acetate from acetylphosphate through substrate-level phosphorylation of ADP.

1.2.3 Eut Gene Organization and Content

A study by the Mushegian group on the comparative genomics of EA utilization genes found that nearly 100 fully sequenced bacterial genomes contain *eut* gene clusters (103). They determined the content and organizational differences among the different phylogenetic groups of those bacteria that contain these genes. Organisms with *eut* gene clusters usually fall into one of two categories: long gene cluster or short gene cluster. Actinobacteria, and most Proteobacteria are part of the short gene cluster group; consisting only of *eutBC* and usually the EA transporter, *eat*. However, some of the Proteobacteria also contain an ortholog of the transcription regulator *eutR* at another genomic location, but no apparent orthologs of other *eut* genes.

On the other hand, the long *eut* gene clusters are found in Firmicutes and Enterobacteriaceae, which include *S. typhimurium, E. coli,* and *E. faecalis* (103). In the long *eut* gene cluster there is an arrangement of as many as 17 genes. Intriguingly, among the species containing long *eut* gene clusters, there are substantial differences in the exact gene content and organization. *Klebsiella pneumoniae*, *Pseudomonas fluorescens*, and *Marinobacter aquaeolei* contain two types of *eut* gene clusters. *K. pneumoniae* contains a short operon containing *eutBC* and *eat*. *P. fluorescens* and *M. aquaeolei* contain only *eutBC* in their short operon. *K. pneumoniae* also has a 17-gene long gene cluster, while *M. aquaeolei* contains an additional 16-gene long gene cluster. *P. fluorescens* has a second operon containing just *eutABC*. Of note is that in *P. fluorescens* the two sets of *eutBC* genes belong to the same gene cluster, but their phylogenies differ (103).

Nearly all the species that contain long *eut* gene clusters are facultative anaerobes that inhabit the gut and/or mouth as either commensals and/or pathogens. In comparison, those organisms that have short *eut* gene clusters tend to be obligate aerobes. Based on these differences, it is feasible that the microcompartment is important for EA utilization in the GI tract.

1.2.4 Genetic Regulation of the eut Gene Cluster in E. faecalis

The gene content and organization of the *eut* operon in *E. faecalis* is nearly identical to the genes found in *S. typhimurium*, with a few exceptions. Most importantly, *E. faecalis* lacks the gene encoding for the EutR regulator and contains a two-component system, EutW (sensor histidine kinase) and EutV (response regulator). The Mushegian laboratory also found that, like *E. faecalis*, all of the other Firmicutes that contain *eut* gene clusters include the EutV/W two-component system, while all the Enterobacteriaceae encode for a EutR regulator.

Sensor histidine kinases of two-component systems usually autophosphorylate upon sensing a specific signal, and subsequently transfer the phosphoryl group to a dedicated response regulator. Phosphorylation of the response regulator usually occurs at a conserved aspartate residue on the receiver domain. Upon phosphorylation of the response regulator, a conformational change typically takes place, which sometimes promotes dimerization, influencing the activity of a second domain, which conducts the regulatory response (104). The majority of response regulator output domains modulate the activity of various protein-protein interactions, enzymatic activity, or DNA-protein interactions (105). However, a small fraction have been predicted to regulate gene expression through RNA-binding activity, as mediated by

ANTAR (AmiR and NasR Transcriptional Antiterminator Regulators) domains (106). Work on two ANTAR domain-containing proteins, AmiR and NasR, show that these proteins interact with their target mRNA transcripts, preventing formation of an intrinsic terminator and thereby promoting downstream gene expression (105–109).

EutV, much like AmiR and NasR, contains an ANTAR domain at its C-terminus, and it has been shown that the ANTAR domain disrupts specific intrinsic terminators to allow for downstream gene expression (110–114). Additionally, *eutV/W*-associated gene clusters have been identified in several different bacteria, including but not limited to *Listeria*, *Clostridium*, and *Enterococcus* species (113). Our analysis of these gene clusters also discovered that they were interspersed with multiple examples of a putative terminator/antiterminator site. Each of these sites was found to be comprised of two, tandem hexanucleotide stem-loops, with the second hairpin directly overlapping an intrinsic transcription terminator site (114) (Figure 1-6A). In *E. faecalis*, there are four of these two-hairpin motifs, located immediately upstream of *eutP*, *eutG*, *eutS*, and *eutA* (Figure 1-6B), termed P1P2, P5P6, P7P8, and P9P10, respectively. Interestingly, the RNA substrates of the two previously characterized ANTARs, AmiR and NasR, also have two-hairpin motifs upstream of an intrinsic terminator, suggesting that ANTAR domains bind to a common RNA element. In support of this, certain nucleotides in the hexanucleotide loops have been shown to be important for both NasR and EutV (109, 114–116). Specifically, an A in the first position, and a G at the fourth position within the loop were found to be required for antitermination *in vivo* (114). These features of the ANTAR substrates

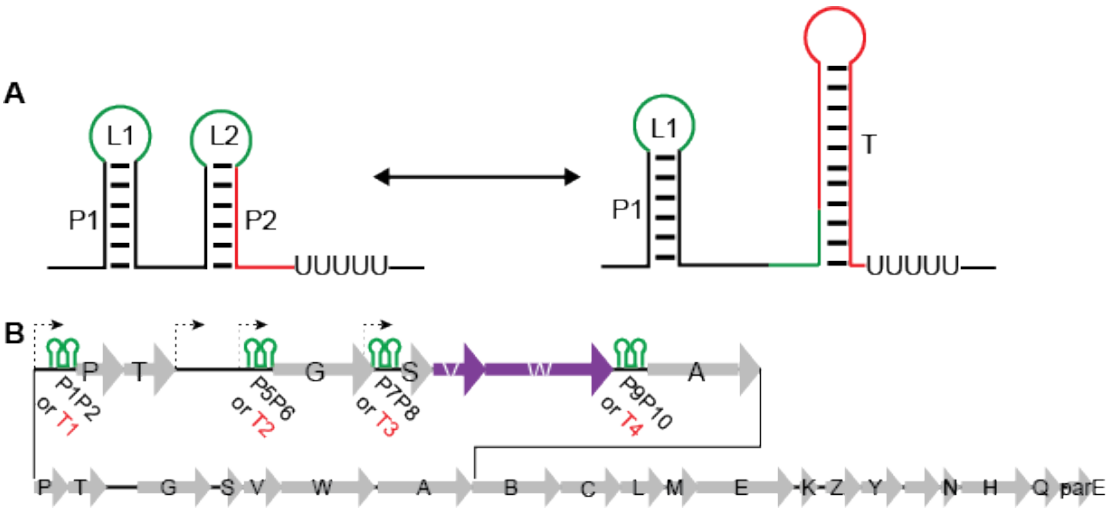

Figure 1-6. Two-hairpin motif ANTAR substrates in *E. faecalis eut* gene cluster. (A) Alternative structures formed by the conserved RNA motif are shown. L1 and L2 denote the two terminal loops involved in ANTAR recognition (green). Altered base pairing allows formation of an alternate RNA structure that includes a terminator stem-loop (red). (B) Location of the four antitermination sites within the *eut* gene cluster. L, loop; P, pairing; T, terminator.

are found in both the Gram-positive bacteria, such as the *eut* genes of *E. faecalis*, and also the Gram-negative bacteria, such as the *ami* and *nas* operons (106, 108, 115).

The EutV ANTAR domain alone is sufficient for dimerization (114), although an adjacent α-helical region is likely to alter the dimer to improve its RNA-binding ability. However, the receiver domain of the EutV response regulator inhibits dimerization, and exists as a monomer, when it is in an unphosphoryated state. Furthermore, phosphorylation-induced dimerization improved EutV RNA-binding affinity (114). This corroborates the previous findings by Baker and Perego where they saw a 3-fold increase in transcriptional readthrough of an intrinsic terminator upstream of *eutG* only under phosphorylating conditions of EutV (112). The functional importance for dimerization of ANTAR domain-containing proteins is further supported by the observation that other classes of ANTAR-containing proteins are also likely to form dimers or higher ordered oligomers (114).

Deletion of the two-component system prevents expression from a *lacZ* gene when it has been transcriptionally fused to the two-hairpin P1P2 region, highlighting the connection between sensing of EA and *eut* antitermination (113). However, while identification of the EutV/W pathway accounts for how *eut* gene expression responds to EA, it was not understood how or if the genes also responded to AdoCbl. Yet, a putative AdoCbl riboswitch was discovered in an intergenic region between *eutT* and *eutG*, upstream of a terminator in front of *eutG* (112, 113) (Figure 1-7). Previously, our lab proposed that this AdoCbl riboswitch served as a 5' UTR for *eutG*, wherein AdoCbl activated antitermination (113). In contrast, the Perego group proposed that this AdoCbl riboswitch served as a 5' UTR of *eutG* or a 3' UTR of *eutT*, but that

binding of AdoCbl promoted termination. Together, these data suggested that in the Firmicutes, a combination of the two-component system, EA, and AdoCbl participate in regulation of *eut* gene expression. As part of my dissertation research, I chose to investigate how *E. faecalis* utilized the AdoCbl riboswitch to genetically regulate the *eut* gene cluster, and to explore how this regulation relates to other ANTAR systems.

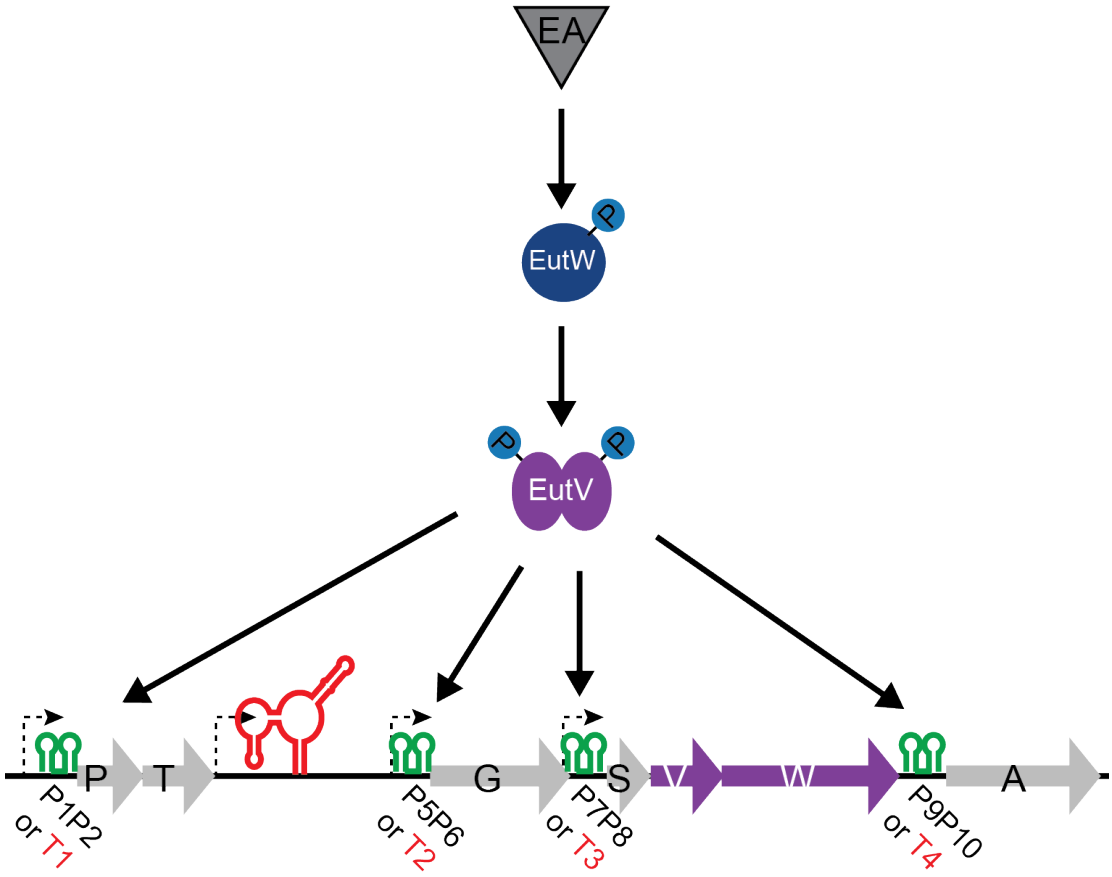

Figure 1-7. Model for post-initiation control of gene expression in the *eut* gene cluster of *E. faecalis*. A portion of genes present in the *eut* gene cluster of *E. faecalis* are shown (grey boxes) and the two-component system, EutVW (purple arrows). A putative AdoCbl-riboswitch is indicated (red) within the intergenic region of *eutT-eutG*. Conserved ANTAR substrates are shown (green). The conformation of these ANTAR substrates is affected by the presence of EA via the EutVW two-component system.

Chapter 2: Discovery of a riboswitch-containing sRNA that controls gene expression through sequestration of a response regulator

6.1 Copyright notice

Portions of chapter 2 were originally published by *Science* (345(6199):937–940) as: Debroy, S., Gebbie, M., Ramesh, Goodson, J. R., A., Cruz, M. R., van Hoof, A., Winkler, W. C., Garsin, D. A. (2014). A riboswitch-containing sRNA controls gene expression by sequestration of a response regulator.

Author Contributions: Gebbie, M., Dr. Debroy, S., Dr. Ramesh, A., Goodson, J. R., Cruz, M. R. performed research; Gebbie, M., Dr. Debroy, S., Dr. Ramesh, A., Goodson, J. R., Dr. van Hoof, A., Dr. Garsin, D. A., Dr. Winkler, W. C. analyzed data; and Garsin, D. A., and Dr. Winkler, W. C. wrote the paper.

Data gathered by Gebbie, M. are in Figs. 2-1A, 2-2, 2-3, 2-4, 2-5, 2-6, 2-7, and 2-9, Debroy, S. and Cruz, M. R. are in Figure 2-1B, 2-8 and 2-9, Goodson, J. R. are in Figure 2-1A.

2.2 Introduction

During the last decade there has been dramatically increased recognition that RNAs can serve many types of regulatory roles, and not simply as a passive intermediate between DNA and protein. Regulation at the RNA level is important in all three kingdoms of life, and a variety of regulatory RNAs and their mechanisms of action have been identified. Riboswitches are one such important class of RNA-based regulation (2, 58, 66, 68). Traditionally, riboswitches are described as *cis*-acting

regulatory components located primarily in the 5' untranslated regions of mRNAs; however, they can also be found in the 3' untranslated regions of mRNAs or within intercistronic regions of transcripts. They comprise an aptamer domain that directly interacts with a target ligand, and a downstream expression platform, which undergoes structural changes upon ligand binding to influence gene expression. The earliest studies on riboswitches described expression platforms controlling transcription termination via rho-independent terminator/antiterminator structures, or the start of translation by controlling access to a RBS (71). However, more recently, additional mechanisms have been documented, such as the signal-responsive control of mRNA cleavage and degradation by riboswitches (64, 65). Additionally, the role of riboswitches in termination was recently expanded to include rho-dependent termination (117).

In addition to riboswitches, bacteria also frequently employ *trans*-acting noncoding RNAs (sRNAs) for posttranscriptional genetic regulation (9, 118, 119). sRNAs typically regulate gene expression by forming short, interrupted base-pairing interactions with target mRNAs, to affect translation or mRNA stability (2). However, a few sRNAs have been shown to affect gene expression patterns by instead sequestering RNA-binding proteins. For example, CsrB and CsrC are *E. coli* sRNAs that feature a tandem collection of binding sites to sequester the important regulatory protein CsrA, which is involved in catabolite repression (4). CsrA is an RNA-binding protein that affects expression of many mRNAs; consequently, the CsrB/C sRNAs affect global gene expression by titrating CsrA from its mRNA targets. *E. coli* CsrB/C are not the only protein-sequestering sRNAs. Indeed, many

bacteria appear to encode for CsrA or proteins that are reasonably closely related to CsrA (4, 34). For example, in *Pseudomonas* species, RsmA, a translational repressor protein, belonging to the CsrA family, binds GGA motifs present in the leader of target mRNAs (120). The location of the GGA motif often coincides with the Shine Dalgarno sequence, thus RsmA can obstruct translation initiation. In *Pseudomonas aeruginosa* the function of RsmA is regulated by two functionally redundant sRNAs, RsmY and RsmZ. Lastly, in several bacterial plasmids regulation of RepA synthesis is controlled by an asRNA, CopA. CopA interacts with its target, CopT, an asRNA in the RepA mRNA transcript, which results in the formation of an RNA-RNA duplex, demonstrating an additional form of posttranscriptional regulation by RNA-mediated sequestration (121).

Interestingly, some recently discovered genetic regulatory mechanisms challenge the traditional perceptions on riboswitch regulation, and blur the lines between simple classification of "*cis"* and "*trans*" regulatory RNAs. For instance, two examples were discovered where riboswitches affected gene expression in *trans*, rather than, or in addition to, the *cis*-acting regulatory functions that are generally presumed to occur (122). This was previously demonstrated for a *Listeria monocytogenes S*-adenosylmethionine (SAM)-responsive riboswitch, which, although it functions as a classical *cis*-acting riboswitch, was also discovered to act in *trans* to form Watson-Crick base pairing interactions with the 5' leader region of a target mRNA. Upon associating with the target mRNA, it was found to decrease translation, analogous to the function of traditional sRNAs (122).

The Winkler and Garsin laboratories have been working together to study post-

initiation mechanisms of gene regulation in the EA utilization (*eut*) locus of *E. faecalis* for several years. This locus encodes the genes required for the breakdown of EA. In *Salmonella* species, the *eut* locus is activated by a DNA-binding transcription factor, EutR, in response to two signals – EA and AdoCbl (77, 83, 84, 100). In *Enterococcus*, *Clostridium* and *Listeria* species, the *eut* locus contains in place of EutR a two-component regulatory system: EutW histidine kinase and EutV response regulator. EutVW is important for promoting microcompartment formation in *E. faecalis*, as can be seen in figure 2-1B when deletion of EutVW, even in the presence of AdoCbl and EA, prevents microcompartment formation.

Discovery of the EutVW TCS accounts for how *E. faecalis* couples detection of EA to genetic regulation of the large *eut* gene cluster. However, if *E. faecalis* relies upon the same signals for genetic regulation of EA utilization microcompartments as *Salmonella*, it should also incorporate a mechanism for sensing AdoCbl (83). Indeed, our prior data suggested that *in vivo* expression of the *E. faecalis eut* genes requires AdoCbl in addition to EA (Figure 2-1A) (113). This is also true for the formation of Eut microcompartments (Figure 2-1B). Therefore, we hypothesized that *E. faecalis*, and by extension other *eut*-containing bacteria, regulate EA catabolism in a manner that is sensitive to the presence of AdoCbl.

Riboswitches that sense AdoCbl have been previously described (62, 123). To examine whether a riboswitch could account for AdoCbl regulation of *E. faecalis eut* genes, the Winkler and Garsin laboratories specifically examined the intergenic regions within the *eut* gene cluster for evidence of secondary structure and primary sequence conservation related to AdoCbl-riboswitches. A putative regulatory RNA

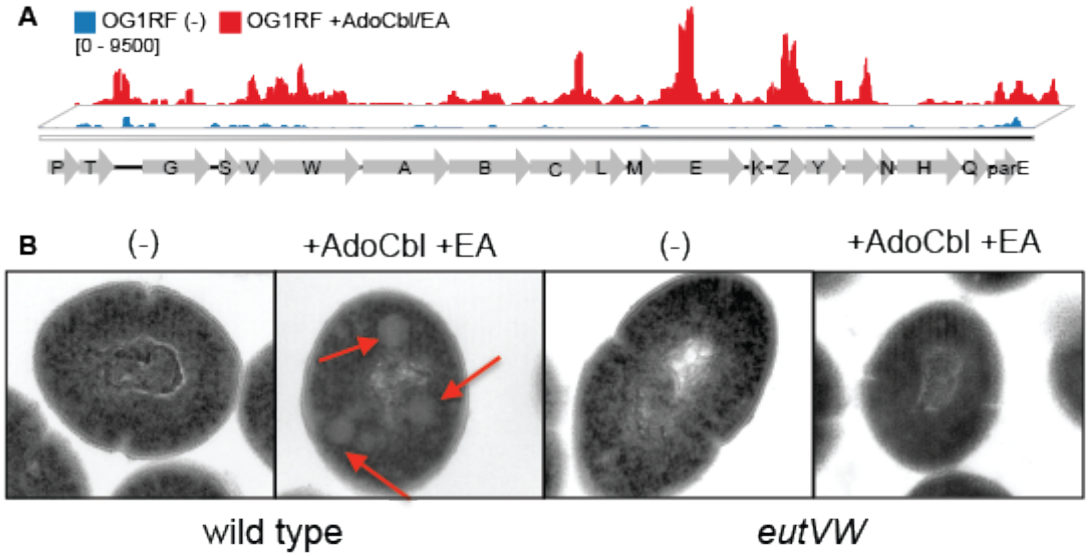

Figure 2-1. EA and AdoCbl induce *eut* gene expression and microcompartment formation. (A) RNA-seq showed that *eut* genes were expressed in the presence of EA and AdoCbl, whereas little or no expression was observed in modified minimal media. (B) Shown are representative cells observed by TEM. The wild type cells (OG1RF) grown in EA and AdoCbl contained microcompartments, while the *eutVW* cells, containing a deletion of the two-component system, and/or without both inducing compounds, had none. The light grey, icosahedral-like structures within the cells are indicative of these structures, and are marked with red arrows.

located between *eutT* and *eutG* was identified through these efforts based on secondary structure conservation. Of specific interest, this RNA element exhibited high similarity to an AdoCbl-sensing riboswitch class described previously (83, 124–126). Also, a sequence motif closely resembling the putative *E. facaelis eutT-G* AdoCbl riboswitch could also be identified in the *eut* operon of *Listeria* species.

The aptamer portion of the putative *E. facaelis eutT-G* AdoCbl riboswitch was then investigated by in-line probing (113), which is a secondary structure probing method (127). This revealed that only a small portion of the *eutT-G* intergenic region appeared to exhibit a significant degree of secondary structure content. Analysis of truncated sequences then narrowed this region to ~200 nt, which included a putative AdoCbl riboswitch. In-line probing was again used to investigate the secondary structural arrangement of the *eutT-G* riboswitch upon varying AdoCbl concentrations. The addition of AdoCbl resulted in decreased band intensity at a number of different RNA positions, suggesting that AdoCbl bound to the RNA. These data suggested that AdoCbl induces a global conformational change for the *eutT-G* riboswitch with an apparent K_d of ~ 540 nM, which is comparable with the dissociate constants for the canonical AdoCbl riboswitch (126). However, the apparent K_d for an AdoCbl analog, cyanocobalamin, was significantly poorer, demonstrating that the *eutT-G* riboswitch is selective for AdoCbl.

Classical AdoCbl riboswitches promote premature transcription termination in 5' leader regions in response to binding of AdoCbl. Such a role would not be logical for the *eutT-G* AdoCbl riboswitch; it would not make sense for the riboswitch to decrease *eut* transcription in response to an increase in AdoCbl. Indeed, we predicted

the opposite. Therefore, at the onset of this project it was unclear how the putative AdoCbl riboswitch affected *eut* gene expression in *E. faecalis*, as well as other *eut*-expressing Gram-positive bacteria. In this chapter, I sought to answer this challenging question.

2.3 Results and Discussion

2.3.3 Detection of transcripts produced within the *E. faecalis eutT-G* intergenic region

Previously a constitutive promoter was detected upstream of the AdoCbl riboswitch located in the upstream portion of the *eutT-G* intergenic region (112). However, a second promoter was also predicted and experimentally verified upstream of *eutG* (112, 113). The discovery of two active promoter DNA sequences within a single intergenic region (*eutT-eutG*) was unexpected, and suggested strongly that the AdoCbl riboswitch may be involved in a regulatory mechanism other than simple attenuation control of *eutG*, as would be predicted for traditional riboswitches. Therefore, we sought to investigate the transcriptional landscape of this region and to begin to explore the regulatory function of the putative *eutT-eutG* AdoCbl riboswitch.

The first question to explore was whether a transcript encompassing the putative AdoCbl riboswitch could be detected, either as the result of premature termination within a long *eutG* leader region or as part of an independent noncoding RNA transcript. To that end, total RNA from *E. faecalis* OG1RF was extracted from cells grown in modified minimal medium in the presence or absence of EA and/or AdoCbl. The RNA was resolved by denaturing polyacrylamide gel electrophoresis and transferred to positively charged membranes for northern blot analyses. The

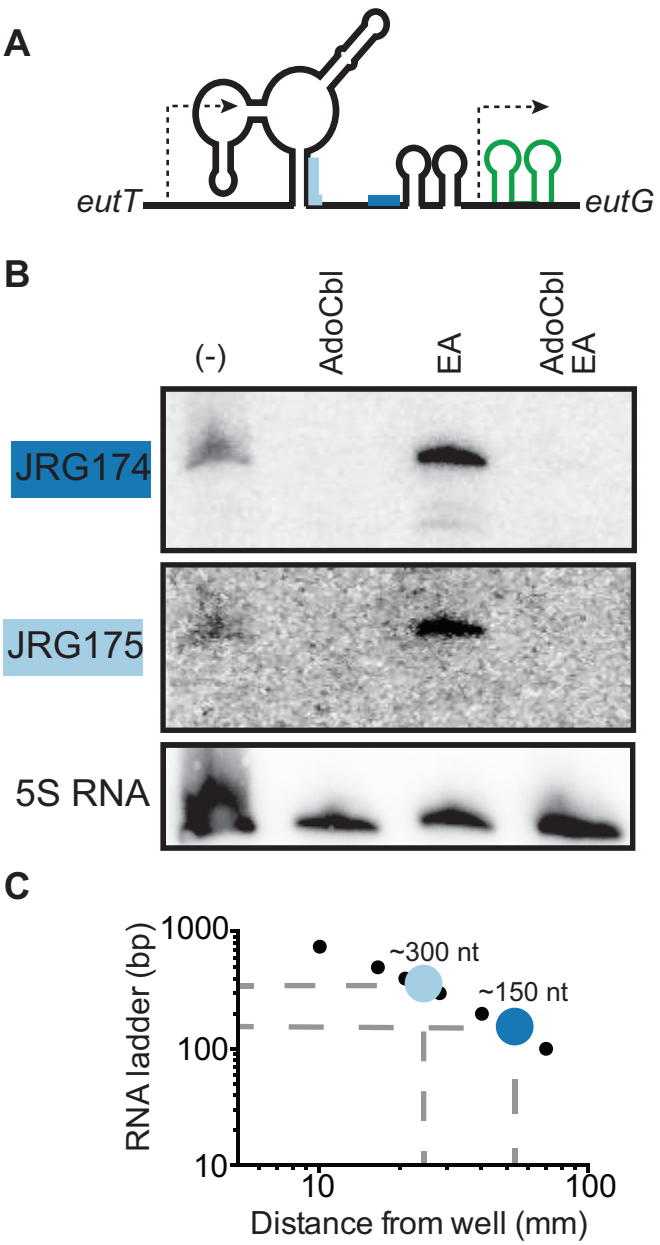

Figure 2-2. Northern blot analysis of the *eutT-eutG* intergenic region shows two RNA species. (A) Location of the ^{32}P-radiolabeled DNA oligonucleotide probes used to analyze the *eutT-eutG* intergenic region. (B) Total RNA was isolated from cells grown in the indicated medium conditions, resolved by denaturing polyacrylamide gel electrophoresis, transferred to a membrane, and hybridized with different DNA probes. (C) Size standard curve showing the sizes of the RNA species found in (B).

membranes were then probed with 5'-radiolabeled DNA oligonucleotides. Specifically, the JRG174 oligonucleotide corresponded to the 3' end of the intergenic region, and JRG175 corresponded to the 3' portion of the AdoCbl riboswitch (Figure 2-2A).

DNA probe JRG174 indicated a faint band present for cells grown in the absence of EA and AdoCbl, but a slightly more intense band of the same size when cells were grown in the presence of only EA (Figure 2-2B). Upon reprobing the northern blot with DNA probe JRG175, the two previously mentioned bands exhibited greater signal intensity, but the blots also showed multiple faint bands of smaller sizes, found only for cells cultured in the presence of AdoCbl, regardless of the presence of absence or EA (Figure 2-2B). The simplest interpretation of these data was that two RNA species were being produced: one was present in low levels in minimal media and induced in the presence of EA, and the second, smaller species, was present only when AdoCbl was supplemented to the media. However, these northern blots were not entirely convincing by themselves, due to the lack of overall intensity of each band. Due to the lack of sensitivity, we turned to internally radiolabeled antisense RNA probes in the hopes of increased sensitivity to these two RNA species. Indeed, probing northern blot membranes with a radiolabeled RNA that corresponded to the antisense sequence of the ~370 nucleotide *eutT-eutG* intergenic region revealed much higher band intensity signals (Figure 2-3). However, the overall pattern of bands remained the same. Therefore, the combination of RNA probes revealed that *E. faecalis* was likely to produce two RNA species of differing lengths, depending on the presence or absence of AdoCbl and/or EA.

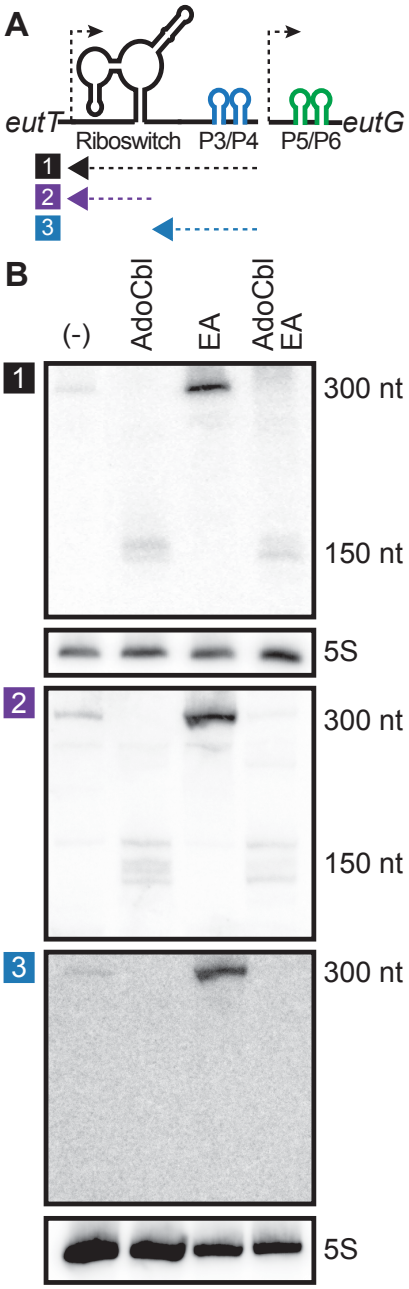

Figure 2-3. Northern blot analysis of the *eutT-eutG* IGR. (A) Location of the ^{32}P-radiolabeled RNA antisense probes used to analyze the *eutT-eutG* intergenic region. (B) Total RNA was isolated from cells grown in the indicated medium conditions, resolved by denaturing polyacrylamide gel electrophoresis, transferred to a membrane, and hybridized with different antisense RNA probes.

In order to determine the approximate sizes of these RNA species, an RNA size standard was resolved alongside total RNA aliquots by denaturing polyacrylamide gel electrophoresis and subjected to northern blot analyses. Total RNA from the four primary growth conditions (minimal media, plus AdoCbl, plus EA, and plus AdoCbl/EA) were analyzed thusly. The distance that each RNA species of the RNA size ladder traveled relative to the well was plotted as part of a size standard curve (Figure 2-2C). By comparing the RNA species that were identified for cells cultured in the presence of AdoCbl to the size standard curve, they were estimated to correspond to a length of approximately 150 nucleotides. However, the larger species, present at low levels in minimal media, and induced by cells grown in the presence of EA, was approximately 300 nucleotides in length.

The discovery of two active promoter DNA sequences within the single intergenic region of *eutT-eutG* was generally unexpected. Nevertheless, we initially hypothesized that the AdoCbl riboswitch found within the *eutT-eutG* intergenic region may participate in a regulatory mechanism, such as attenuation control of *eutG*, as would be expected for traditional riboswitches. However, northern blot analyses of this intergenic region, first using DNA oligonucleotide probes, and second, using higher sensitivity RNA antisense probes, suggested instead that there were two RNA species produced from this intergenic region. This intriguing observation suggested that the AdoCbl riboswitch was not simply involved in attenuation control of *eutG* through use of a classical riboswitch regulatory mechanism; we suspected the riboswitch was participating in something more

complex or unusual.

2.3.3 Sequence analysis of the *eutT-G* intergenic region reveals a new putative ANTAR substrate

Our northern blot analyses had revealed the presence of two transcripts of different sizes; however, it did not determine to which portions of the *eutT-eutG* intergenic region they corresponded. To help resolve this question, the membranes were hybridized with internally radiolabeled RNA probes, which were designed to anneal to different regions (Figure 2-3A). Probes that were designed to specifically anneal to the AdoCbl riboswitch resulted in bands corresponding to both the ~150mer and ~300mer RNAs (Figure 2-3B, purple). This suggested that the riboswitch was included in both of the *eutT-eutG* intergenic transcripts. From these aggregate data, we concluded that an sRNA of approximately ~300 nucleotides and that includes the AdoCbl riboswitch is produced from the *eutT-eutG* intergenic region. We named the noncoding RNA EutX. These data also suggested that a shortened version of EutX, which also includes the AdoCbl riboswitch, is produced when cells are cultured in media that includes AdoCbl.

Manual inspection of the sequences downstream of the AdoCbl riboswitch surprisingly revealed the presence of a putative two-hairpin motif, which is remarkably similar to the EutV transcription attenuation substrates found elsewhere in the *eut* gene cluster. Specifically, it was found that there are two hairpins approximately 114 nucleotides downstream of the riboswitch, which are separated by 7 nucleotides and contain the core conserved residues found in other ANTAR substrates (Figure 2-3A, blue versus green hairpins). However, unlike the EutV

attenuation sites, this particular two-hairpin motif, which was termed P3P4, did not appear to overlap with an adjacent intrinsic terminator site, suggesting it might have a function other than classical transcription attenuation. To investigate whether these hairpins were included as part of the newly discovered EutX RNA, an antisense probe that included this sequence was utilized in northern blot analyses. The short form of EutX (~150 nt) was not observed on these blots, whereas the long form (~300 nt) could be readily observed (Figure 2-3B, blue). Therefore, we conclude from these data that the shortened form of the EutX transcript is likely to be comprised largely of the AdoCbl riboswitch, whereas the longer form of EutX also includes the riboswitch but extends beyond it to include the newly identified P3P4 hairpin motifs.

2.3.3 Determination of termini of EutX RNAs

To identify the termini of the short and long forms of EutX, we employed rapid amplification of cDNA ends (RACE) with circularization. We isolated RNA from cells grown in the presence of EA without AdoCbl, since these cells lacked the shortened form of EutX. Total RNA was treated with DNase I to remove contaminating DNA and with Terminator exonuclease, which exhibits specificity for 5' monophosphorylated transcripts and is oftentimes used to remove ribosomal RNA. The Terminator exonuclease-treated RNA was then incubated with Tobacco Acid Pyrophosphatase to convert 5'-triphosphates to 5'-monophosphates. The RNA was then incubated under dilute conditions with T4 RNA ligase, in order to encourage intramolecular circularization. After cDNA synthesis and PCR amplification, the amplified DNA was sub-cloned into an appropriate vector. Sequencing analyses were performed on a subset of these plasmids.

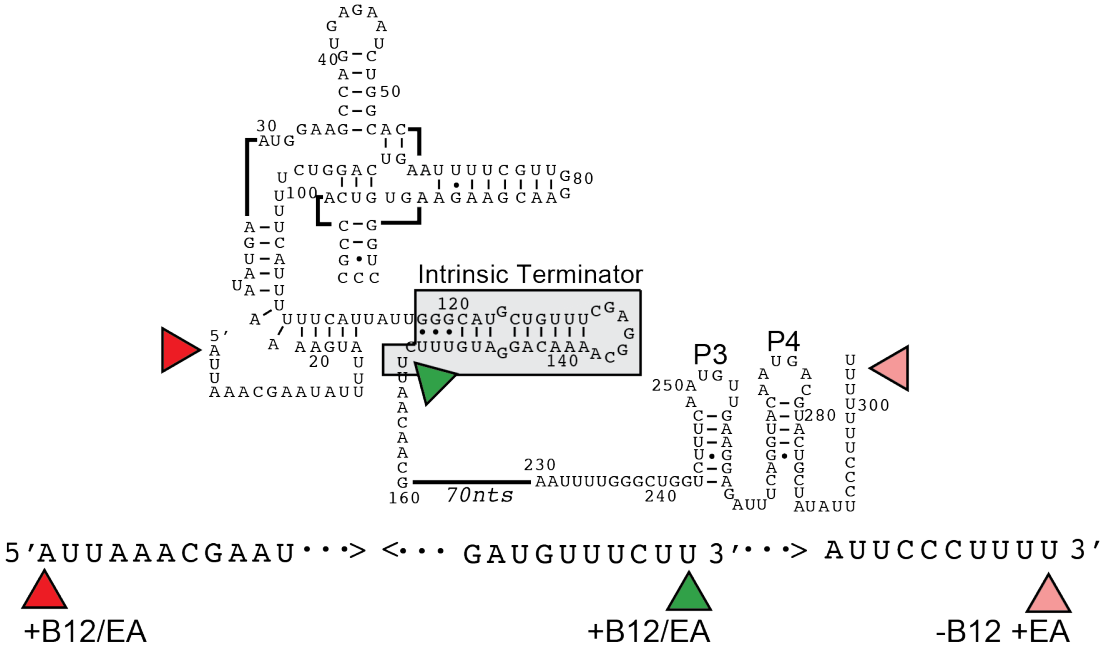

Figure 2-4. Analysis of EutX termini. 5' RACE (circularized) mapped 8 sequences to the 5' site AAUUA in the presence of AdoCbl and EA (red triangle). 3' RACE (circularized and linear) mapped 8 sequences to the 3' termini at U153, in the presence of AdoCbl alone, the same size indicated by northern blot (green triangle). 3' RACE (linear) mapped 8 sequences to the oligouridate tract 3' of the P4 stem-loop, in the presence of EA alone, the same size indicated by northern blot (pink triangle). The putative intrinsic terminator is highlighted by a grey-shaded box.

Of eight sequences that were obtained from this analysis, all exhibited the same apparent 5' terminus. Specifically, the sequences all began at the same A residue, located shortly after the promoter upstream of the riboswitch (Figure 2-4, red triangle). This method also allows for simultaneous identification of the 3' terminus. Of these same eight sequences, all exhibited a 3' terminus located at the end of an oligouridylate tract that occurs just upstream of the -35 of the downstream *eutG* promoter (Figure 2-4, pink triangle) and just downstream of P3P4. Since oligouridylate tracts are usually positioned at the ends of intrinsic terminator sites, we examined the sequence surrounding P3P4, but could find no convincing helical element that could be construed as being part of an intrinsic terminator site. Also, the distance between the 5' and 3' termini, as mapped by this study, was approximately 300 nucleotides, which corroborates well with the results of the size estimation of EutX described above.

We also attempted a complementary approach identifying the termini of EutX RNAs for cells cultured in the absence of AdoCbl and presence of EA. We employed an altered 3' RACE procedure where RNA molecules were ligated to a 5' pre-adenylated linker (NEB, Universal miRNA cloning linker), which could be used for subsequent cDNA amplification, PCR amplification, sub-cloning and DNA sequencing. This treatment and analysis revealed that culture conditions that include high amounts of EA, but low amounts of AdoCbl, resulted in 3' termini within an oligouridylate tract after P3P4 (Figure 2-4, pink triangle). The sequencing data again demonstrated that 3' termini form at one of several sites at the end of the polyuridylate tract. Specifically, we found two sequences corresponding to termini in

the second U, two sequences corresponding to termini in the third U, and finally four sequences corresponding to the fourth U of the polyU tract found just downstream of the second hairpin, P4, in the two-hairpin motif. This suggests that EutX overall is between 296 and 298 nucleotides in length. Again, this confirmed our predicted size of ~300 nucleotides of the RNA induced in the presence of EA.

The same method was used to determine the 3' termini of the shortened form of EutX, which was observed for cells grown in the presence of AdoCbl but absence of EA. This treatment and analysis revealed a 3' terminus at the end of the riboswitch aptamer where another stretch of U residues can be found (Figure 2-4, green triangles). Given the 5' terminus we previously identified, the total length of this RNA sequence would be 150 nucleotides, which agrees well with the results of the northern blot analyses. Upon further investigation, the poly U tract sequence appeared to be part of an intrinsic terminator hairpin, which is formed from sequences at the 3' end of the AdoCbl riboswitch aptamer (Figure 2-4, grey box). This observation strongly suggests that a riboswitch-associated intrinsic terminator site is formed under conditions of high AdoCbl or in the presence of AdoCbl, but absence of EA. This is in contrast to previous studies, which suggested the AdoCbl-sensing element was involved in transcription antitermination (113).

2.3.4 Deletion analysis of EutX

We postulated that the P3P4 ANTAR substrate located 3' of the AdoCbl riboswitch, which is not transcribed in the presence of AdoCbl, should not affect transcription of the AdoCbl riboswitch itself. To test our hypothesis, we extracted total RNA from *E. faecalis* OG1RF strains containing deletions of the riboswitch,

P3P4, or both the riboswitch and P3P4. The RNA was then resolved by denaturing polyacrylamide gel electrophoresis, and transferred to a positively charged membrane for northern blot analyses. These membranes were probed with radiolabeled DNA oligonucleotides JRG174 and JRG175. As expected when both the riboswitch and P3P4 element were deleted, neither DNA probe was able to detect a band (Figure 2-5, right). Probing the deletion of just the riboswitch, which should leave P3P4 unaltered, showed that indeed the riboswitch was undetected. However, deletion of the riboswitch prevented detection of P3P4 (Figure 2-5, left). We speculate that this may be due to instability of the RNA transcript in the absence of the 5' AdoCbl riboswitch. When we probed the deletion of P3P4, we, as expected, lost the ability to detect P3P4 by JRG174. However, this deletion also affected the detection of the AdoCbl riboswitch. Instead of seeing the EutX 150mer in the presence of AdoCbl and AdoCbl/EA, we saw the 150mer product present in minimal media alone and AdoCbl, but not in the presence of AdoCbl/EA (Figure 2-5, middle).

2.3.5 The EutX AdoCbl riboswitch can terminate transcription *in vitro*

The *in vivo* evidence demonstrated that the EutX AdoCbl riboswitch promotes formation of a 150mer RNA product when cells are cultured in the presence of AdoCbl. This suggested that the riboswitch is likely to promote premature transcription termination in response to binding of AdoCbl. To test this hypothesis, we utilized a single-round synchronized transcription termination assay (128). We designed a PCR-generated, double-stranded DNA template for transcription of the EutX- associated AdoCbl riboswitch. Grafted upstream of the EutX AdoCbl riboswitch,

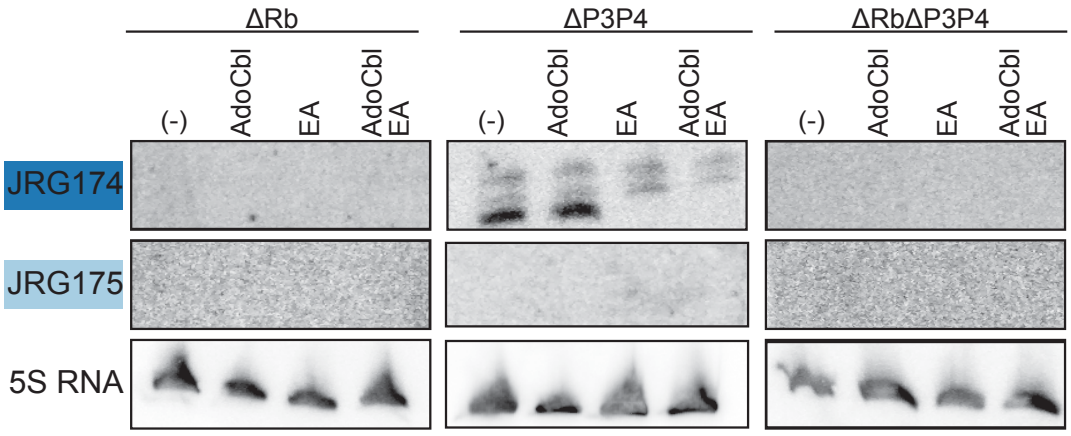

Figure 2-5. Northern blot analysis of deletion strains. Total RNA was isolated from cells grown in the indicated conditions, resolved by denaturing polyacrylamide gel electrophoresis, transferred to a membrane, and hybridized with different DNA oligonucleotide probes. Left corresponds to a strain with a deletion of the AdoCbl riboswitch; middle, deletion of the P3P4 element; right, deletion of both the AdoCbl riboswitch and P3P4 element.

from +5 to +170 of EutX sequence, was nucleotides -135 to +1 of the *B. subtilis glyQS* promoter. The DNA template also included 20 nucleotides that follow the putative EutX riboswitch intrinsic terminator. We hypothesized that the presence of this additional sequence at the 3' end would permit us to quantitatively measure transcripts resulting from "run-off" versus "premature termination", by denaturing polyacrylamide gel electrophoresis.

The DNA template was designed such that it lacked cytidine residues until position +30, relative to the transcription start site. By omitting CTP from an initiating mixture of NTPs, the transcription elongation complex was therefore halted at position 29 (Figure 2-6A, highlighted in red). To release the elongation complexes from the halt site, all four NTPs were added. Heparin was simultaneously added for prevention of reinitiation. Synchronizing transcription elongation complexes in this manner increases the overall sensitivity of the assay and prevents cooperative effects that emerge from the presence of multiple polymerase complexes on the same DNA template (129).

This assay revealed the presence of two prominent bands in the presence of AdoCbl that could be resolved by 6% denaturing PAGE (Figure 2-6B, WT). A template ending at the predicted terminator site, +150 nt as a size standard for "termination". Additionally, a template with an extra 20 nucleotides, +170 nt, added downstream of the predicted terminator site was used as a size standard for "run-off". The longer product was interpreted to result from run-off transcription on the DNA template since it ran alongside our +170 nt size standard. The shorter RNA product was interpreted to result from premature transcription termination, since it ran beside

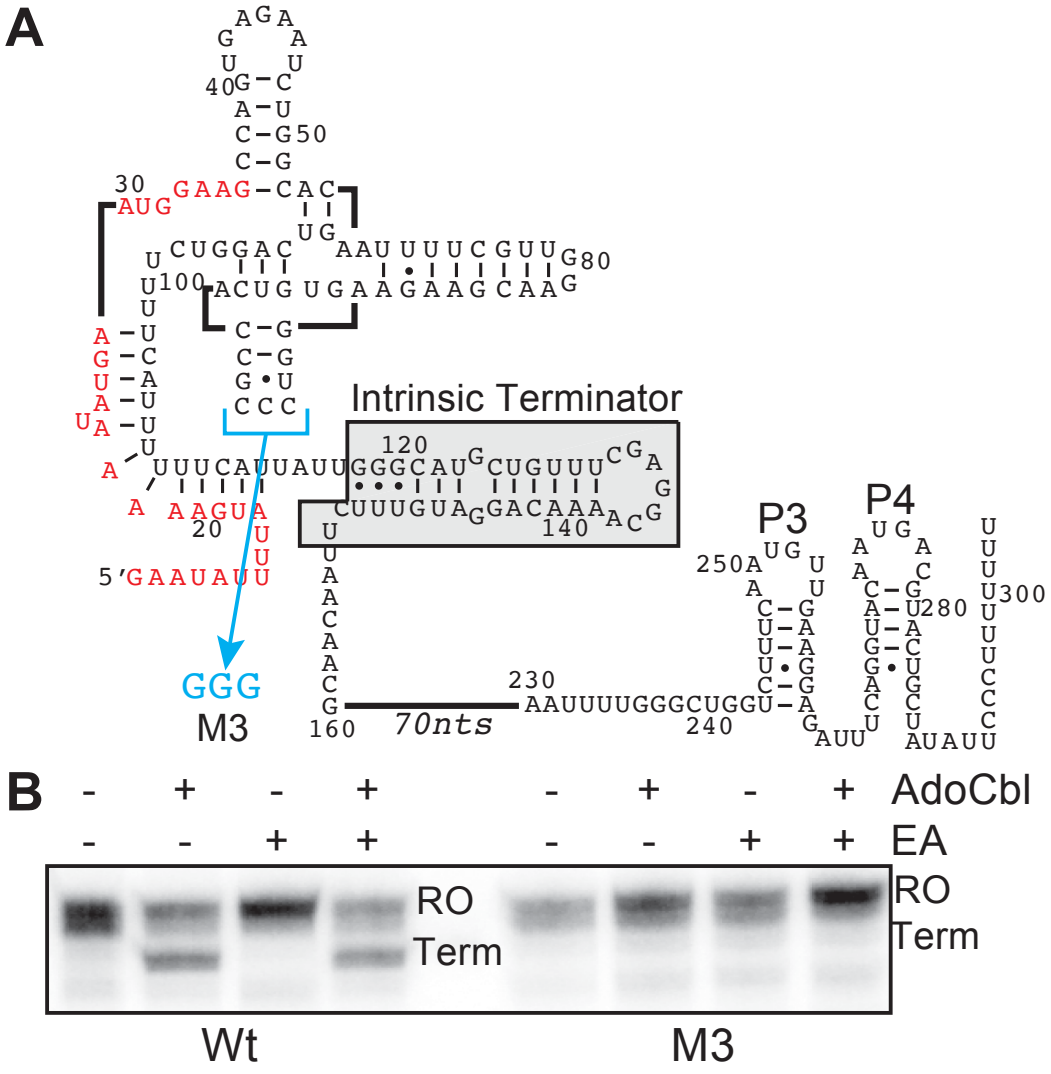

Figure 2-6. AdoCbl binding to the riboswitch causes premature termination. (A) Schematic showing the secondary structure of the AdoCbl riboswitch in EutX. A putative intrinsic terminator is highlighted in a grey-shaded box. The red nucleotides denote the halted complex for the single-round transcription. Mutation to the L5 loop, M3, shown in blue. (B) *In vitro* transcription termination assay with wild-type EutX template displays a premature termination product only in the presence of AdoCbl. *In vitro* transcription termination assay with M3 EutX template displays that termination is lost in the M3 mutant template. Wt, wild-type; RO, run-off; Term, terminated.

the +150 nt "termination" standard. Under these assay conditions and no AdoCbl and the absence or presence of EA, approximately 92.7% and 97.1%, respectively**,** of transcription corresponded to full-length transcripts, which we hypothesize requires readthrough of the riboswitch intrinsic terminator site. In support of this, addition of 250 μM AdoCbl resulted in a significant increase in the prematurely terminated product, reducing terminator readthrough to ~60%. However, while these data support our hypothesis that the AdoCbl riboswitch promotes transcription termination in an AdoCbl-dependent manner, it is important to note that the readthrough frequency is likely to be very different *in vivo*, where the cell experiences different ionic conditions, NTP concentrations and the presence of transcription elongation factors that are absent in our biochemical assay. Therefore, we conclude that the transcription termination assay provides important proof-of-principle demonstration that the EutX riboswitch promotes termination in response to AdoCbl.

To test our hypothesis on the riboswitch further, we chose to introduce a mutation that would be predicted to result in a loss of AdoCbl binding. To identify an appropriate mutation, we considered canonical AdoCbl riboswitches. The secondary structure of canonical AdoCbl riboswitches includes a small loop, which contains three highly conserved cytosines that form a pseudoknot with downstream guanosines (39). The formation of this pseudoknot is necessary for ligand binding (123). The EutX AdoCbl riboswitch is predicted to contain this same characteristic stem-loop, which has been described as L5 (Figure 2-6A, blue). We then mutated these three residues to guanines and found that the mutant sequence led to an increase in run-off transcription (from ~60% to ~95%), which was not influenced by the addition of

AdoCbl (Figure 2-6B, M3).

2.3.6 The EutX P3P4 site is required for genetic regulation of the *eut* gene cluster

EA was previously shown to activate autophosphorylation of EutW and the subsequent phosphoryl transfer between EutW and EutV, resulting in dimerization of EutV causing an overall increase in EutV RNA-binding affinity (111–114). Identification of the additional two-hairpin motif in the *eut* gene cluster, P3P4, led us to hypothesize that EutV binds EutX at this site. To further characterize the function of EutX and this putative ANTAR substrate, P3P4, the ability to bind EutV was examined by electrophoretic mobility assays (EMSAs) and differential radial capillary action of ligand assays (DRaCALA) (130). Upon addition of micromolar concentrations, purified EutV bound to a portion of EutX that contained P3P4 *in vitro* (Figure2-7A and B) much like it binds the first two-hairpin substrate in the *eut* operon, P1P2 (Figure2-7C and D). Unlabeled 30 μM competitor RNA (EutX or P1P2) could fully compete for binding to EutV in the context of radiolabeled EutX or P1P2 (Figure 2-7A and C second to last lane). However, an unrelated RNA of similar length was unable to compete for EutV, which demonstrated that binding of EutV to EutX is specific (Figure2-7A and C last lane).

However, to determine whether the EutV-EutX interaction is functionally significant *in vivo*, we sought to gain further genetic evidence for this hypothesis. We generated a *eutG-lacZ* fusion that contained the entire intergenic region between *eutT* and *eutG* (Figure 2-8A). We transformed this construct into *E. faecalis* strain SD289, where in the endogenous locus of this strain, the region comprising EutX was deleted.

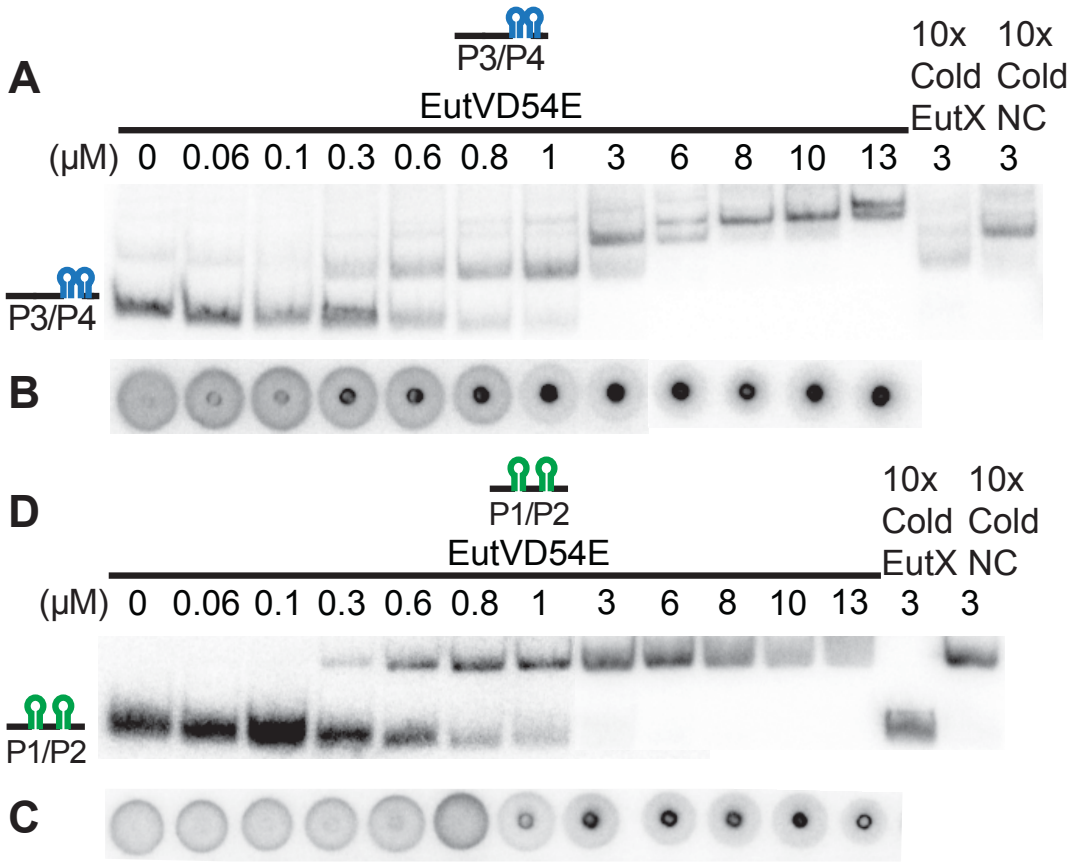

Figure 2-7. EutV specifically binds EutX. (A) EMSAs were used to analyze binding of EutV to EutX. A 117 nt fragment of EutX ('117EutX'), and EutVD54E protein were utilized (constitutively active for RNA binding). Controls included addition of 30 μM of an unlabeled competitor EutX RNA and a completely unrelated RNA. (B) Differential radial capillary action of ligand assay (DRaCALA) was used to assess EutV binding to 117EutX. (C) EMSAs were used to investigate whether EutX could compete for EutV binding to P1P2, and EutVD54E protein. Controls included addition of 30 μM of an unlabeled competitor EutX RNA and a completely unrelated RNA. (D) Differential radial capillary action of ligand assay (DRaCALA) was used to assess EutV binding to the P1P2 RNA using the same reaction conditions and control RNAs as (B).

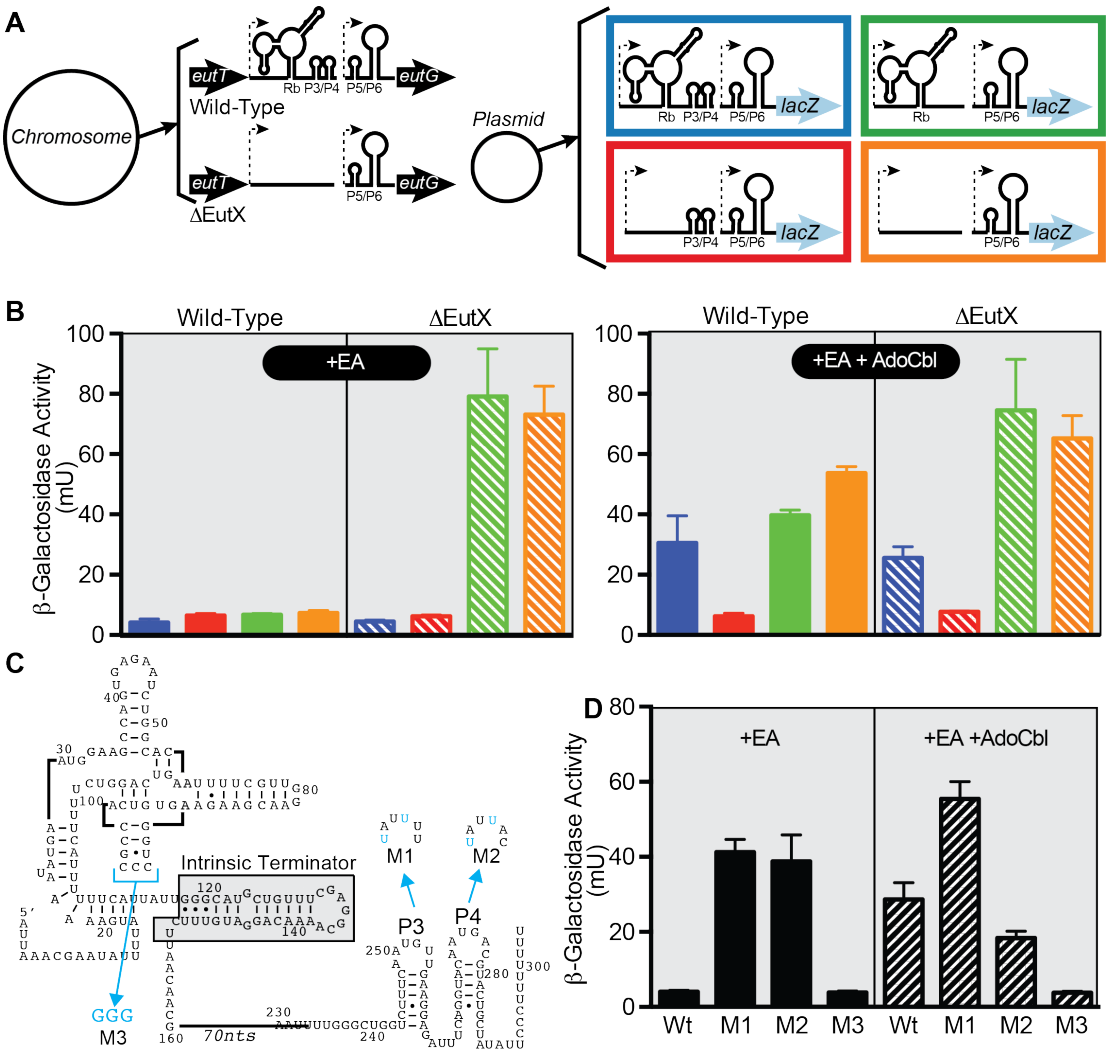

Figure 2-8. EutX acts in *trans* to sequester EutV. (**A**) A diagram of the experimental design. A plasmid containing a *eutG-lacZ* fusion was manipulated to contain a deletion of the riboswitch, the P3P4 hairpin, or both. The plasmid was introduced into *E. faecalis* strains containing either wild type EutX or a deletion. (B). β-galactosidase assays of strains shown in (A) when grown in modified minimal medium containing EA or both EA and AdoCbl. (C) Schematic of M1, M2, and M3 mutations. (D). β-galactosidase assays of strains missing EutX but carrying *eutG-lacZ* plasmids that contain either the M1 or M2 or M3 mutations. M1 mutation contains A249U and G252U site-directed changes. M2 mutation contains A273U and G276U site-directed changes. These mutations alter the conserved 1st and 4th positions within the P3 and P4 terminal loops. M3 mutation contains G58-60C. The data presented in (B) and (D) is the average of three or more independent experiments and the error bars represent the standard deviation.

Using this background was necessary for accurate analysis of mutations in the *eutG-lacZ* construct; if EutX functions to sequester active EutV~P, this activity would be predicted to affect gene regulation in *trans*. As shown in Figure 2-8B (dashed blue bars), expression of wild-type *eutG-lacZ* was induced by the presence of EA and AdoCbl when cells were grown in a modified minimal medium. These requirements for *eut* gene expression were previously reported by another group using a similar *eutG-lacZ* construct (112), and by our own studies employing *lacZ* fusions to different *eut* genes (114). Next, a construct with a deletion for the putative P3P4 ANTAR substrate was tested. Surprisingly, this construct was induced to very high levels, just by adding EA (Figure 2-8B, dashed green bars). Though the riboswitch aptamer was still present, the requirement for AdoCbl was completely relieved. Additionally, deleting the riboswitch aptamer had no further effect on expression (Figure 2-8B, dashed orange bars). These data are consistent with the hypothesis that the P3P4 two-hairpin motif of EutX sequesters active EutV~P; when P3P4 is not present, however, EutV~P is not sequestered and can activate *eut* gene expression by its antitermination activity at the other ANTAR substrates in *eut* gene cluster. We postulate that the riboswitch's function is to prevent protein sequestration when it is bound to AdoCbl. As predicted by this model, a construct in which the riboswitch aptamer of EutX was deleted, but still contained P3P4, was not induced under any condition, even when both AdoCbl and EA were added to the medium (Figure 2-8B, dashed red bars). When a wild type strain containing the endogenous locus of EutX was utilized, control by AdoCbl was regained; induction was observed only when

both AdoCbl and EA were present (Figure 2-8B, solid green and orange bars relative to striped bars). Because the regulation of the *lacZ* constructs can be altered by the presence or absence of EutX in the chromosome, these data suggest that the sRNA can regulate in *trans* and are consistent with the a sequestration model (Figure 2-9).

The deletion mutants tested in Figure 2-8B removed significant amounts of primary sequence in the *eutT-eutG* intergenic region. These large disturbances could effect *eutG* gene regulation by mechanisms other than what we propose. To ameliorate this concern and further test the sequestration model, we made targeted point mutations in the *eutG-lacZ* construct (Figure 2-8C). For example, the sequence of the six-nucleotide loops of the ANTAR dual hairpin substrates was previously shown to be important for effective binding by ANTAR protein. In particular, an adenosine (A) at the first position and a guanine (G) at the fourth position were required (114). The putative P3P4 substrate also contains A's at the first position and G's at the fourth position of the loops (Figure 2-8C). We tested one mutant, in which the A1 and G4 of the P3 loop were mutated to U's (Mut1) and a second mutant in which the P4 loop contained these same mutations (Mut2). The strains containing the Mut1 and Mut2 point mutations, respectively, were both induced by EA alone, in contrast to the wild-type *eutG-lacZ* construct (Figure2-8D, black solid bars relative to dashed bars). Since these mutations are predicted to prevent EutV from binding to the putative P3P4 sequestration substrate, the results were as predicted.

Recent structural work on AdoCbl riboswitches has confirmed the importance of the L5 loop of the ligand-bound aptamer (113). L5 is characterized by three cytosines that are predicted to base pair with three guanines in an accessory loop to

form a pocket that enables AdoCbl binding. The *eut* AdoCbl riboswitch is predicted to contain this characteristic L5 loop (Figure 2-8C, highlighted in blue) (113), thus we mutated the three cytosines to guanines to create Mut3. Like complete deletion of the riboswitch, this resulted in a *lacZ* construct that was not inducible (Figure 2-8D, M3 black and dashed bars).

2.4 Conclusion

Bacteria must be able to sense their environment and respond readily to any environmental stimuli or changes, including from within the GI-tract, where microbial inhabitants are typically required to exhibit metabolic flexibility. Since one of the nutritional options for GI-tract microbes is EA, it is reasonable that the gut commensal *E. faecalis* synthesizes the *eut* gene cluster for EA utilization. Through our studies on *eut* genes, we have now discovered multiple genetic regulatory mechanisms that affect their expression. Previously, we found that the sensor kinase, EutW, directly senses EA to promote autophosphorylation and then phosphoryl-transfer to activate the response regulator, EutV, which through its RNA-binding ANTAR domain activates readthrough of several transcription terminator sites permitting downstream *eut* transcription. However, in this study, we discovered that *E. facaelis*, like *Salmonella* species, also incorporates sensing of AdoCbl for genetic regulation of *eut* genes. In contrast to the transcription factor EutR for *Salmonella* species, *E. faecalis* employs an AdoCbl-sensing riboswitch. The AdoCbl-riboswitch in *E. faecalis* is unique in that it is part of a *trans*-acting RNA element, EutX. The two-hairpin motif present in EutX, P3P4, is responsible for the binding and sequestration of EutV. Under conditions in which EA is readily available and

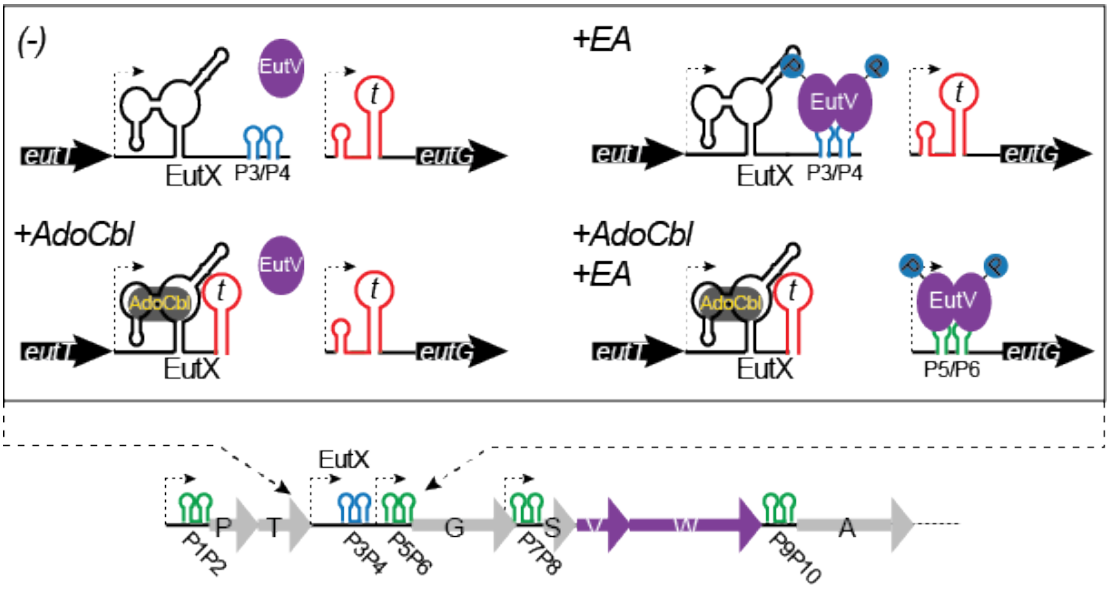

Figure 2-9. Model of the regulation of the *eut* gene cluster by EutX. Model for how EutX uses both a riboswitch and an two-hairpin binding site to regulate *eut* gene expression. See main text for details.

AdoCbl is not, EutV~P will be bound to EutX, preventing antitermination at the other sites within the *eut* gene cluster, thus shutting down *eut* gene expression (Figure 2-9, top right). However, if AdoCbl is present, the free and active EutV~P is available to antiterminate and promote *eut* gene expression (Figure 2-9, bottom right).

Sequestration of EutV by EutX makes logical sense since it would prevent the overproduction of accessory *eut* genes, which are presumably energetically taxing, and wasteful when the primary cofactor for EA catabolism is limited. However, several questions remain to be answered. It is not understood what the affinity of EutV for EutX is, or more simply the affinity for P3P4. Understanding the affinity of EutV for P3P4 combined with an understanding of the relative concentrations of EutV, EutX, and the AdoCbl riboswitch *in vivo* would greatly aid in the understanding of this regulatory circuit, and its biological relevance. The sequestration of EutV by EutX could be as simple as a much higher concentration of EutX relative to EutV in the absence of AdoCbl, thus EutX would be acting as a sponge for EutV. Alternatively, in conditions where there is limited AdoCbl, EutX could merely have higher affinity for EutV than the other ANTAR substrates present in the *eut* gene cluster. Understanding the biophysical interactions between the ANTAR domain of EutV relative to the ANTAR substrates would shed light on other ANTAR-domain containing proteins. Much has been characterized in terms of the RNA side of the binding requirements for antitermination *in vitro*, however, little is understood about the protein side. Further understanding of the direct RNA-protein interactions is needed to better understand the sequestration of EutV by EutX.

In this chapter, we present a new example of a protein-sequestering small

regulatory RNA. However, this sRNA is particularly novel in that it also contains an AdoCbl-binding riboswitch that regulates in-*cis* the protein-sequestration activity. This discovery therefore presents a new type of protein-sequestering sRNA as well as a novel mechanism for riboswitch action. In Figure 2-10 we see the current modes of riboswitch action, where they have been shown to function in transcription termination/antitermination (Figure2-10 A), translation inhibition (Figure 2-10 B), and affecting mRNA stability (Figure 2-10 C), but now we have identified a riboswitch that controls the formation of a two-hairpin motif, responsible for protein sequestration, and downstream gene expression (Figure 2-10 D). Not only is this a new mechanism for riboswitch action, it is not the only one of its kind. Shortly after identification of the *E. faecalis* EutX, a similar riboswitch-containing sRNA, Rli55, in *L. monocytogenes* was discovered suggesting this may be conserved in bacteria that utilize the EutV/W TCS (131, 132).

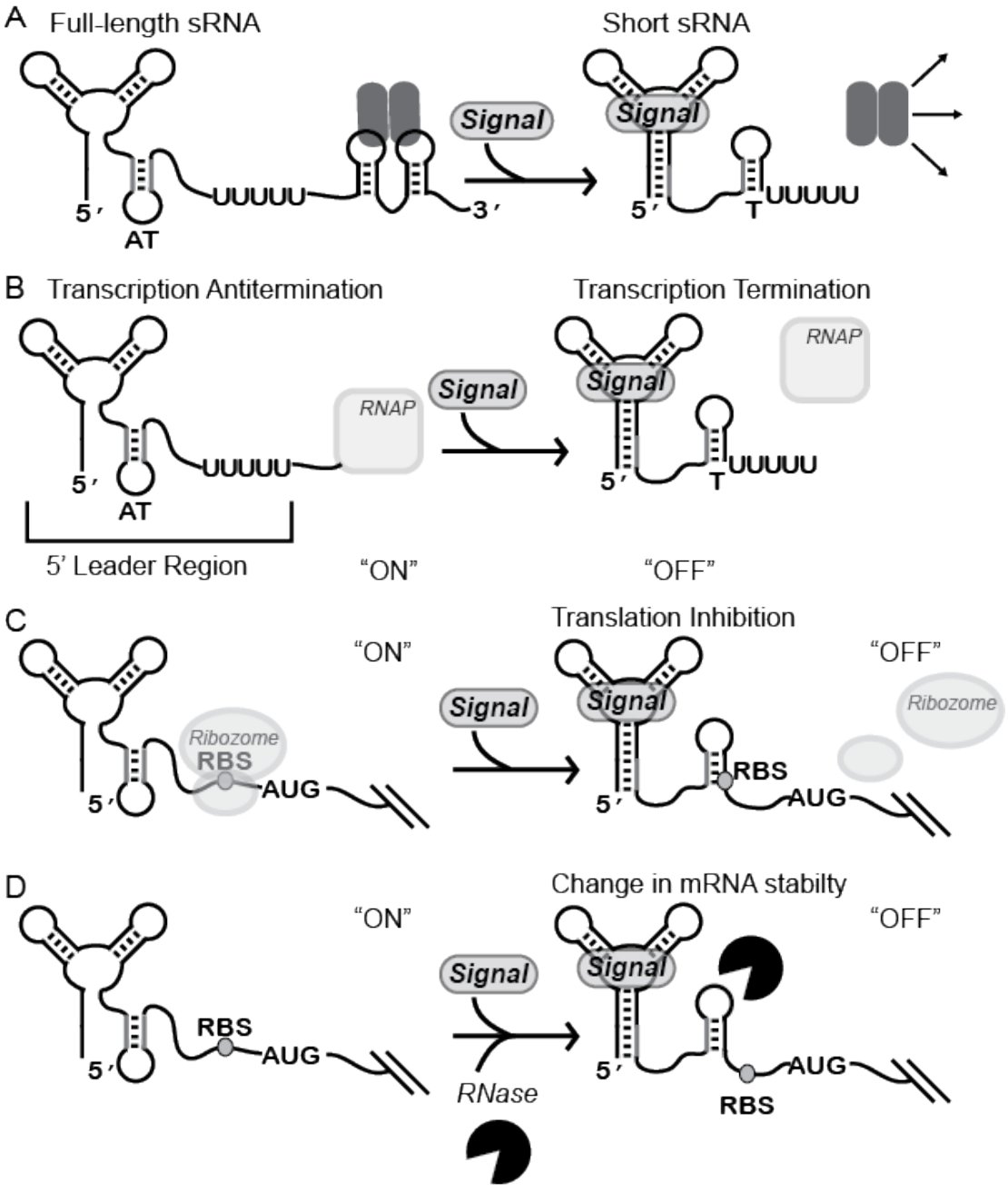

Figure 2-10. Updated summary of different modes of action for *cis*-acting RNAs. (A) Control of formation of an intrinsic transcription terminator by a metabolite-binding riboswitch, 'M' denotes the metabolite ligand. (B) Control of translation initiation efficiency by a metabolite-sensing riboswitch. (C) Control of mRNA stability by conformational change of a riboswitch upon binding its ligand. (D) Control of a protein sequestration site by conformational change of a riboswitch upon binding its ligand

Chapter 3: Antitermination determinants of the *Enterococcus faecalis* ethanolamine-responsive two component regulatory system

3.1 Author contribution

Data gathered by Gebbie, M. are in Figs 3-1 through 3-16, while data gathered by Weiss, C.A., are in Fig. 3-3B.

3.2 Introduction

The ability to sense and respond to sudden changes within an environmental niche is essential for all microorganisms. Correspondingly, microorganisms have evolved many genetic regulatory strategies for coupling specific recognition of cellular changes to control of gene expression. One of the most universal bacterial strategies for coordinating signals with cellular adaptations is through the use of two-component regulatory systems (TCS) (133). In the typical TCS, a sensor kinase protein perceives a chemical or physical signaling cue and couples this information to autophosphorylation of a conserved histidine (133). Autophosphorylation then promotes transfer of the phosphoryl group to an aspartate located within a highly conserved receiver domain of the cognate response regulator protein. Phosphoryl transfer to the response regulator protein triggers changes to an output domain, which can affect protein:protein interactions, enzymatic activity, or gene expression (134). A majority of these signal transduction pathways rely on DNA-binding output domains to regulate gene expression at the level of transcription initiation (134).

However, an increasing number of RNA-mediated mechanisms have been identified that can affect gene expression at the post-initiation level (2, 3). In recent years our laboratory, as well as other laboratories, have found that a subset of these post-initiation regulatory mechanisms utilize TCS for coupling signal detection to post-initiation regulatory control (131, 132).

To exert post-initiation regulatory control, some TCS incorporate an RNA-binding output domain into the response regulatory protein. A characteristic domain called the AmiR and NasR Transcriptional Antiterminator Regulator domain (ANTAR) is used in these instances (106, 108, 113, 114, 135, 136). Approximately 1% of TCS appear to utilize RNA-binding ANTAR domains (134, 137); however, ANTAR domains can also be found widespread in other protein architectures (106). Sequence homology-based searches have predicted more than 1100 ANTAR domain-containing proteins, widely distributed across at least 644 bacterial species (114). ANTAR-domains are usually found within multi-domain proteins. The largest class of ANTAR-containing proteins, comprising nearly 50%, consists of ANTAR-based response regulators; however, several other arrangements are common. One important class of ANTAR proteins seem to possess an N-terminal pseudo-receiver domain capable of protein-protein interactions (137). This class of ANTAR protein, therefore, may regulate gene expression via interactions with a modulator protein that itself may possess signal-sensing function. For example, the *Pseuodomonas aeruginosa* ANTAR protein, AmiR, dimerizes upon binding two molecules of its negative regulator AmiC (108, 137). Under inducing conditions, AmiC binds a small amide compound and does not associate with AmiR, thereby allowing AmiR to

associate with the 5' leader of its target gene. Binding of AmiR to the nascent transcript prevents formation of an intrinsic terminator site and instead permits transcription elongation of downstream portions of the transcript. The ANTAR domain also occurs in conjunction with a diverse set of signal-sensing domains. For example, in *Klebsiella oxytoca,* NasR contains a nitrate and nitrite-sensing (NIT) domain fused to an ANTAR domain and regulates the *nasFEDCBA* operon (107, 115). In the presence of nitrate, NasR is activated to bind to the 5' leader region of the nascent *nasF* transcript. Similar to AmiR, association of NasR inhibits formation of an intrinsic terminator within the 5' leader region, thereby allowing synthesis of the downstream *nas* operon. However, while these prior data demonstrated that both AmiR and NasR are RNA-binding proteins, that prevent formation of an intrinsic terminator, the determinants for RNA-binding remained unidentified.

E. faecalis EutV, much like AmiR and NasR, contains an ANTAR domain at its C-terminus that associates with RNA to prevent formation of intrinsic terminators, allowing for downstream expression of the *eut* gene cluster (113, 114). EutV is also conserved in other Firmicutes that contain *eut* gene clusters, and is likely to participate in common regulatory mechanisms (72, 111, 113). For *E. faecalis*, the corresponding sensor kinase, EutW, undergoes autophosphorylation upon sensing ethanolamine, which correspondingly triggers phosphorylation and dimerization of EutV (111, 113, 114). In this organism, dimerization promotes binding of EutV at four separate locations within the *eut* gene cluster, preventing formation of intrinsic terminators in all instances. More specifically, prior results from our laboratory showed that these intrinsic terminators each overlap with a two-hairpin motif

comprised of two helices separated by a short oligonucleotide linker. Each hairpin is capped by hexanucleotide terminal loops that resemble one another (113). These EutV binding sites are located 5' of *eutP*, *eutG*, *eutS*, and *eutA*, and have been named P1P2, P5P6, P7P8, and P9P10, respectively (113, 132). It was previously speculated that the *E. faecalis* EutV binding site may share conserved features with terminator-disrupting, putative ANTAR-binding sites in other organisms (114). Interestingly, the same features could be identified within the leader regions that are targeted by AmiR and NasR, suggesting a common ANTAR:RNA interaction has been evolutionarily maintained.

From these aggregate findings, we assume ANTAR domains are likely to bind to a common RNA substrate. However, putative RNA substrates could not be identified for all organisms that have ANTAR proteins, thus suggesting that there will be specificity determinants that occur in the nucleotides that are less conserved at the primary sequence level. Furthermore, it is likely that there will be specificity determinants on the ANTAR-containing proteins that accommodate modified substrates. Therefore, the detailed characterization of ANTAR:RNA interactions is important, as it will provide key information that can be used to predict ANTAR-based regulatory networks across diverse organisms. In this chapter, we aimed to understand some of the key determinants for ANTAR:RNA interactions.

3.3 Results and Discussion

3.3.1 *E. faecalis* EutV/W TCS promotes antitermination in a heterologous host

Previously, we predicted that those proteins that contain an ANTAR domain, including *P. aeurginosa* AmiR, *K. oxytoca* NasR, and *E. faecalis* EutV, recognize a

two-hairpin motif whose distal hairpin directly overlaps a mutually exclusive intrinsic terminator (114). However, the strict determinants for RNA recognition by EutV have yet to be determined. Moreover, EutV antitermination has not been examined in other bacterial hosts, thereby leaving open the possibility that cellular accessory factors could participate in antitermination. Unfortunately, *in vitro* analysis of EutV has been complicated by the fact that the purified protein is highly insoluble and exhibits an increasing fraction of inactive protein upon purification. Therefore, to characterize the requirements of EutV antitermination further we attempted to develop an *in vivo* approach. Specifically, we decided to recapitulate the EutV/W regulatory circuit within the heterologous host *Bacillus subtilis* along with a genetic reporter for quantification of antitermination activity. *B. subtilis* lacks a *eut* operon, thus eliminating any existing complications from accessory Eut factors. *B. subtilis* genetics are considerably more tractable, thus allowing us to test hypotheses for the protein: RNA complex from within this organism.

To begin to recapitulate the EutV/W genetic circuitry in a heterologous host, the *E. faecalis eutV* and *eutW* genes were integrated as a two-gene operon into the nonessential *B. subtilis thrC* locus. Furthermore, transcription of the *eutVW* operon was placed under control of P_{liaI}, which is a bacitracin-inducible promoter (Figure 3.1) (138–140). The P1P2 two-hairpin motif was then placed upstream of a *yfp* reporter gene, but downstream of a constitutive promoter (P_{const}) (Figure 3.1). This reporter construct was integrated into another nonessential locus, *amyE*. Since EutV is predicted to promote readthrough of P1P2 in response to ethanolamine in *E. faecalis*, YFP activity of the P_{const}-P1P2-*yfp* reporter was also predicted to be subject

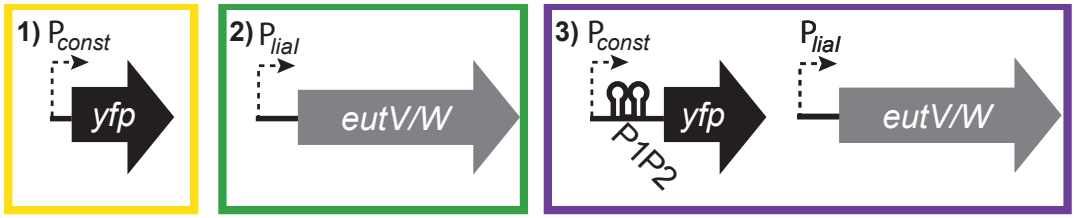

Figure 3-1. Schematic for expression of an *E. faecalis* EutV antitermination circuit in a heterologous host. 1) Control strain containing constitutively expressed *yfp* for "high" *yfp* control. 2) Control strain containing bacitracin inducible reporter: P_{lialI} eutV/W, for "low" yfp control. 3) Test strain containing the constitutively active P1P2T-*yfp* reporter and a bacitracin inducible (P_{lia+}) *eutVW*.

to ethanolamine dependency (Figure 3-1). However, while ethanolamine is catabolized by *E. faecalis*, *B. subtilis* is predicted to be unable to metabolize ethanolamine. This is likely to explain the high tolerance of ethanolamine exhibited by *E. faecalis* in liquid culture. Since ethanolamine would need to be added to *B. subtilis* in order to study the EutV/W TCS in this organism, we examined the impact of ethanolamine on its growth (Figure 3-2). *B. subtilis* cultures were grown in the presence of varying ethanolamine, from 10 μM to 33 mM. Indeed, in 33 mM ethanolamine, the concentration we commonly use for induction of the *E. faecalis eut* gene cluster, growth of *B. subtilis* was dramatically inhibited. We hypothesize that the decreased tolerance for ethanolamine in *B. subtilis* is due to the absence of an ethanolamine ammonia lyase enzyme, EutBC. Given these results, we decided on a concentration of 5 mM, which resulted in no apparent change in growth yield but is likely to still be sufficiently high enough for activation of EutV/W, for study of ethanolamine regulation in *B. subtilis*.

While 5 mM ethanolamine was not inhibitory to growth of *B. subtilis,* we were not certain whether ethanolamine could traverse the *B. subtilis* cell wall, and thus be used to promote antitermination in a heterologous host. However, it appears that *B. subtilis* might encode for a putative ethanolamine transporter, *yxeR* (141, 142), suggesting EA should be able to enter the cytosol, although this gene has not yet been characterized. We then cultured *B. subtilis* cells in the presence and absence of 5 mM ethanolamine and the presence and absence of bacitracin, for induction of *eutVW* expression, and monitored *yfp* expression by fluorescence microscopy (Figure 3-3).

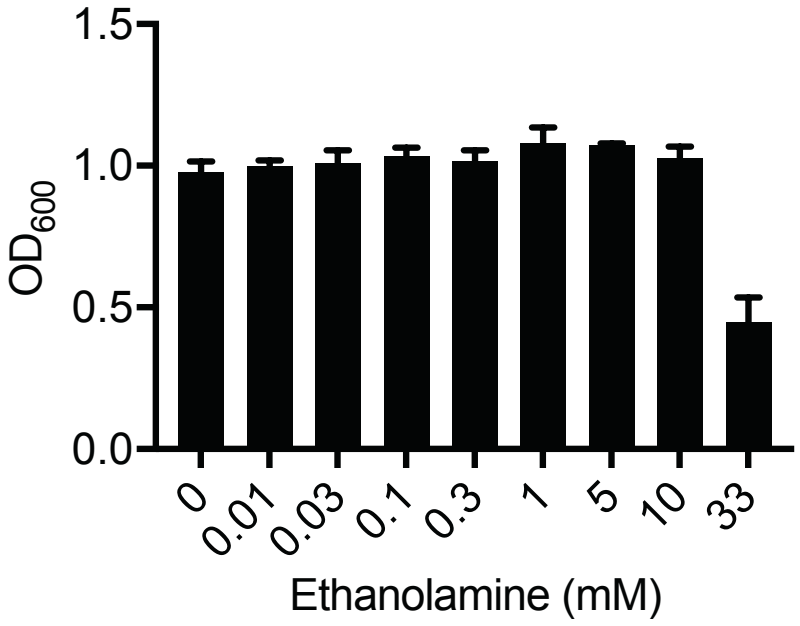

Figure 3-2. High [EA] inhibits growth of *B. subtilis*. *B. subitilis* grown in the presence of increasing EA, and growth monitored at OD_{600} ~1 relative to the no EA culture.

This test revealed that both EutV/W and ethanolamine were required for YFP expression. The presence of EutV/W alone is insufficient to promote readthrough of the P1P2 terminator site, resulting in low YFP levels. Therefore, *E. faecalis* EutV/W antitermination could be reduced to only its simplest constituents and fully recapitulated in a heterologous host.

3.3.2 EutX antagonizes EutV antitermination

Previously we hypothesized that the EutX sRNA functions to sequester EutV, preventing antitermination, thus down regulating gene expression of the *eut* gene cluster (132). To further test the sequestration model of EutX function, we chose to introduce it in the *B. subtilis* host alongside the other components of EutV/W genetic circuitry. We speculated if EutX truly acts in *trans* to sequester EutV in *E. faecalis*, without aid of accessory factors, it would do so in *B. subtilis*. To test this hypothesis, we subcloned the full-length EutX into an extra-chromosomal multicopy vector, pDG148 (143–145). However, since no terminator was found downstream of the P4 stem-loop in EutX, we chose to add a "semi-synthetic terminator" to ensure the full EutX transcript was produced. We employed the polyU tract, identified by 3' RACE (132), to assist in creation of a strong terminator sequence "CCCAACGUUGGGUUUUUUU". The gene encoding for synthesis of EutX with "semi-synthetic terminator" was then transformed into the strain already containing the P_{const} P1P2-*yfp* and the P_{liaI} *eutVW*. EutX expression was induced using 1 mM IPTG. As a control for IPTG induction, we also transformed an empty vector control, pDG148, into the aforementioned strain.

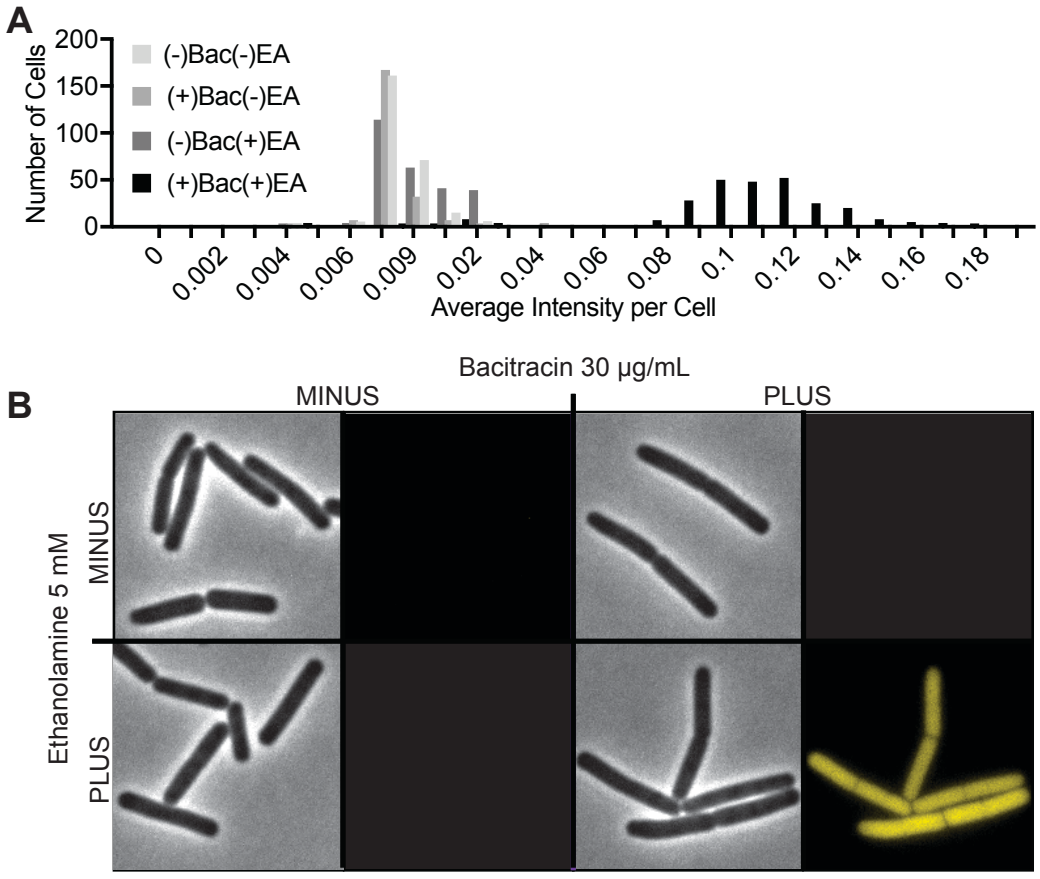

Figure 3-3. *E. faecalis* EutVW TCS promotes antitermination in a heterologous host. (A) Quantification of a strain containing *amyE::*P_{const}P1P2T-*yfp* and *thrC::*P_{liaI} EutVW in the presence and absence of 5 mM EA and/or 30 µg/mL bacitracin. Data quantified using Oufti CellDetection (146). (B) Representative fluorescence microscopy images of above conditions. Images collected by Cordelia A. Weiss.

Strains were cultured and prepared as the previously described with a few exceptions. The pDG148 vectors (empty or containing EutX) were induced at the start of outgrowth in the presence or absence of 1 mM IPTG, to ensure that sufficient EutX was present at the time of *eutVW* induction. At mid-exponential growth phase, *eutVW* was induced with 30 μg/mL bacitracin, and 5 mM EA. After 40 minutes of incubation, cells were centrifuged and prepared for analytical Fluorescently Activated Cell Sorting (FACS) analysis. 30,000 events were counted per strain/condition, and then analyzed using Cytobank (147). While the presence of IPTG unexpectedly appeared to modestly enhance antitermination activity in the empty vector control, induction of EutX exhibited significant reduction of YFP (Figure 3-4). This suggested EutX antagonizes EutV antitermination and requires no accessory factors for its effects on EutV.

Total RNA was then extracted from the appropriate strains and was resolved by denaturing polyacrylamide gel electrophoresis, and transferred to a positively charged membrane for northern blot analyses. We probed the membrane with a radiolabeled antisense RNA to EutX. As a positive control, total RNA was also extracted and analyzed from OG1RF grown in the presence of ethanolamine. The latter RNA material provided a reference for EutX expression on the northern blots (Figure 3-5, first lane). An empty vector control did not show presence of EutX, and nor did the undinduced EutX-containing strain (Figure 3-5, middle two lanes), suggesting that EutX is expressed as expected. We hypothesize that EutX is indeed responsible for the loss of antitermination activity. Under IPTG-inducing conditions a band corresponding to EutX was observed (Figure

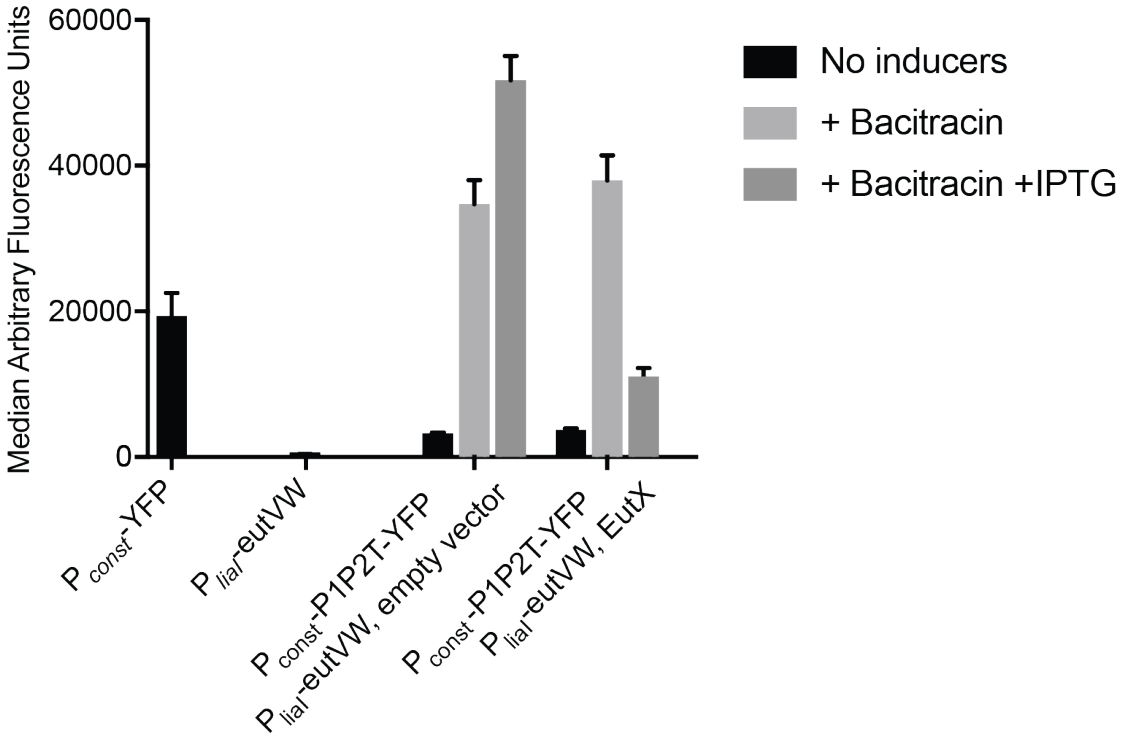

Figure 3-4. EutX antagonizes EutV antitermination in a heterologous host. FACS data showing effect EutX has on antitermination in the presence or absence of 30 μg/mL bacitracin, presence of 5 mM EA, and presence or absence of 1 mM IPTG. P_{lial} *eutVW* expression is under bacitracin control, and EutX is under IPTG control.

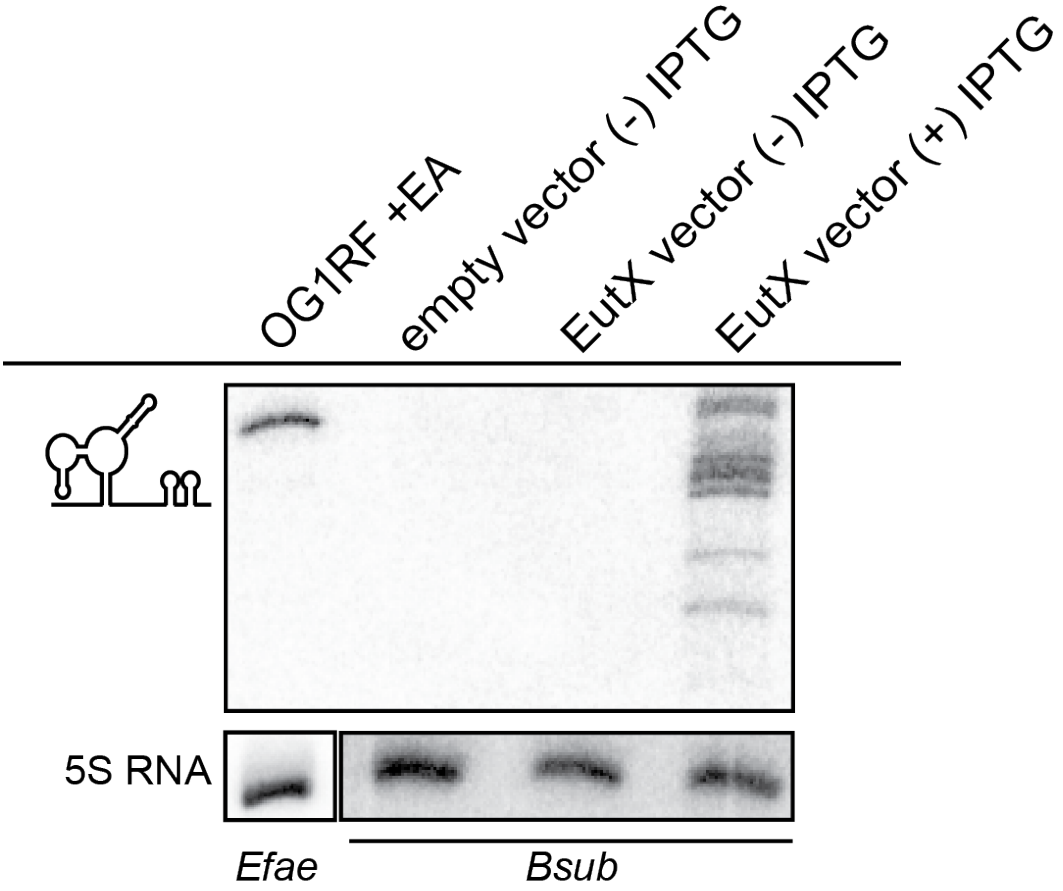

Figure 3-5. Northern blot analysis of EutX expression in *B. subtilis*. Total RNA was isolated from cells grown in the indicated medium conditions, resolved by denaturing polyacrylamide gel electrophoresis, transferred to a membrane, and hybridized with different antisense RNA probes. *Efae, E. faecalis* RNA; *Bsub, B. subtilis* RNA

3-5, last lane); however, instead of seeing a single full-length product, several smaller RNA species could also be identified. We speculate these RNA transcripts result from degradation of the EutX sequence expressed in *B. subtilis*, although it is not immediately clear why the sequence is less stable in *B. subtilis* than that of *E. faecalis*. It is possible that this increased EutX degradation may be due to an additional 50 nt of non-native sequence located just upstream of the previously identified 5' end of EutX, which might contain instability determinants (132). This extra sequence results from subcloning of EutX into the plasmid vector for *in vivo* expression. It is possible that the degradation of EutX in *B. subtilis* may also explain why its production did not result in complete loss of P1P2 antitermination in *B. subtilis*.

Alternatively, it is also possible that overexpression of EutV throws off the balance of EutV:EutX interactions, thus preventing EutX from completely antagonizing EutV antitermination. Under the inducing conditions we chose, it is possible that the abundance of EutV is too high, such that EutX cannot sequester all of the EutV proteins. As a preliminary test of this hypothesis we sought to determine the minimum bacitracin concentration that still promotes EutV antitermination, and that could still affect EutX's ability to antagonize EutV. The level of fluorescence, and therefore antitermination activity, was analyzed via analytical FACS upon incubation with varying bacitracin. We anticipated that less bacitracin, but constant IPTG (for EutX production) would promote further abrogation of antitermination; however, instead we saw a proportional change in the amount of antitermination relative to the amount of bacitracin added (Figure 3-6). This data suggests that there

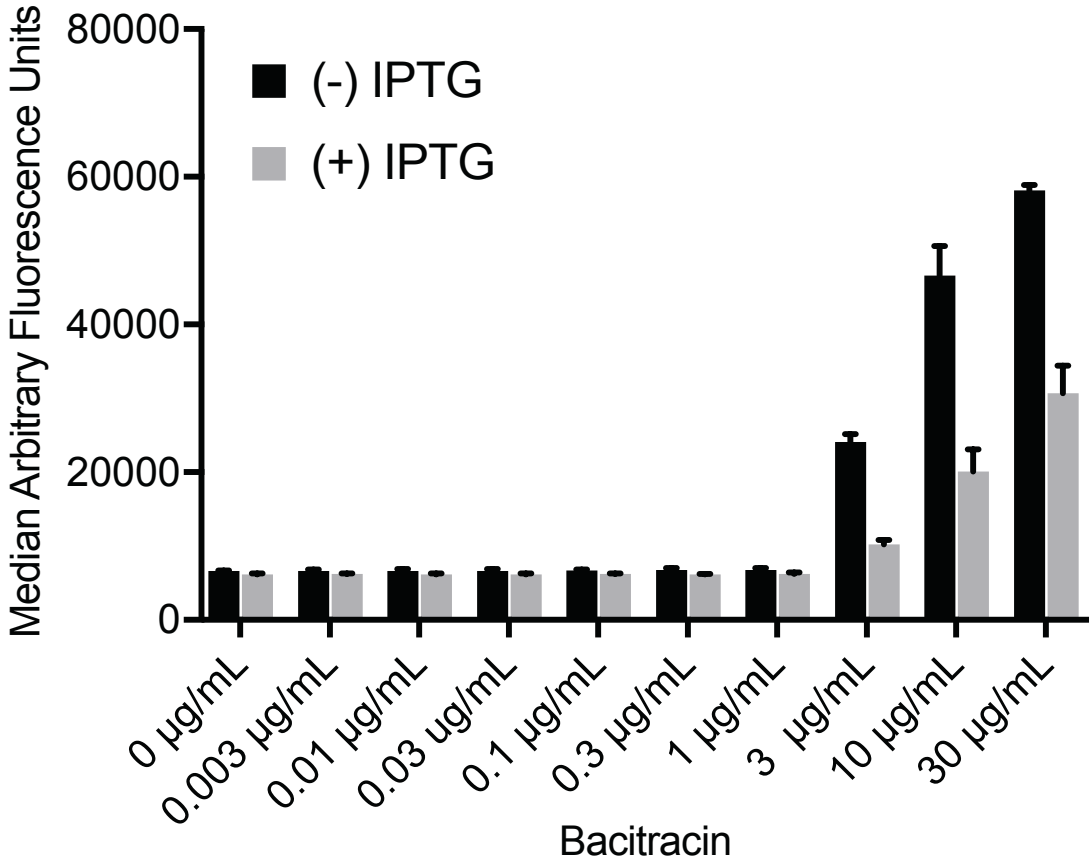

Figure 3-6. Heterologous expression of EutX antagonizes EutV. FACS data showing affect EutX on antitermination when titrating in bacitracin, presence of 5 mM EA, and presence or absence of 1 mM IPTG.

is at least a low level of antitermination always occurring in our heterologous system, and that EutX is unable to fully antagonize EutV antitermination activity, possibly due to the degradation of EutX in the heterologous host.

3.3.3 Charged amino acids in ANTAR domain that may be important for antitermination

In our previous work we identified that the EutV ANTAR domain is sufficient for binding P1P2; however, inclusion of the adjacent α-helical region substantially increased binding affinity (Figure 3-7, A) (114). Additionally, we established the minimum binding requirements for the RNA substrate: a two-hairpin motif whose 3' loop overlaps a terminator structure, and that includes two hexanucleotide loops where the A in position 1 and G in position 4 appear to be required for antitermination function. Conversely, little is known about which portions of ANTAR domains are required for binding RNA substrates. To begin investigating which amino acids were likely to be important for antitermination activity, we aligned the protein sequences of ANTAR-domains from a diverse group of distinct organisms that are predicted to be a part of *eut* gene clusters (79). This comparative sequence alignment revealed 33 conserved amino acids in the ANTAR-domain and an adjacent extended α-helical segment (Figure 3-7). To investigate where in the EutV protein these conserved residues were likely to be positioned, a structural model of the *E. faecalis* EutV protein was generated using the I-TASSER suite (Figure 3-7) (148, 149). Protein threading of *E. faecalis* EutV structural model was based on *Mycobacterium tuberculosis* Rv1626. Using this three-dimensional structural model, it appeared that of the 33 conserved amino acids, 18 were hydrophobic and were

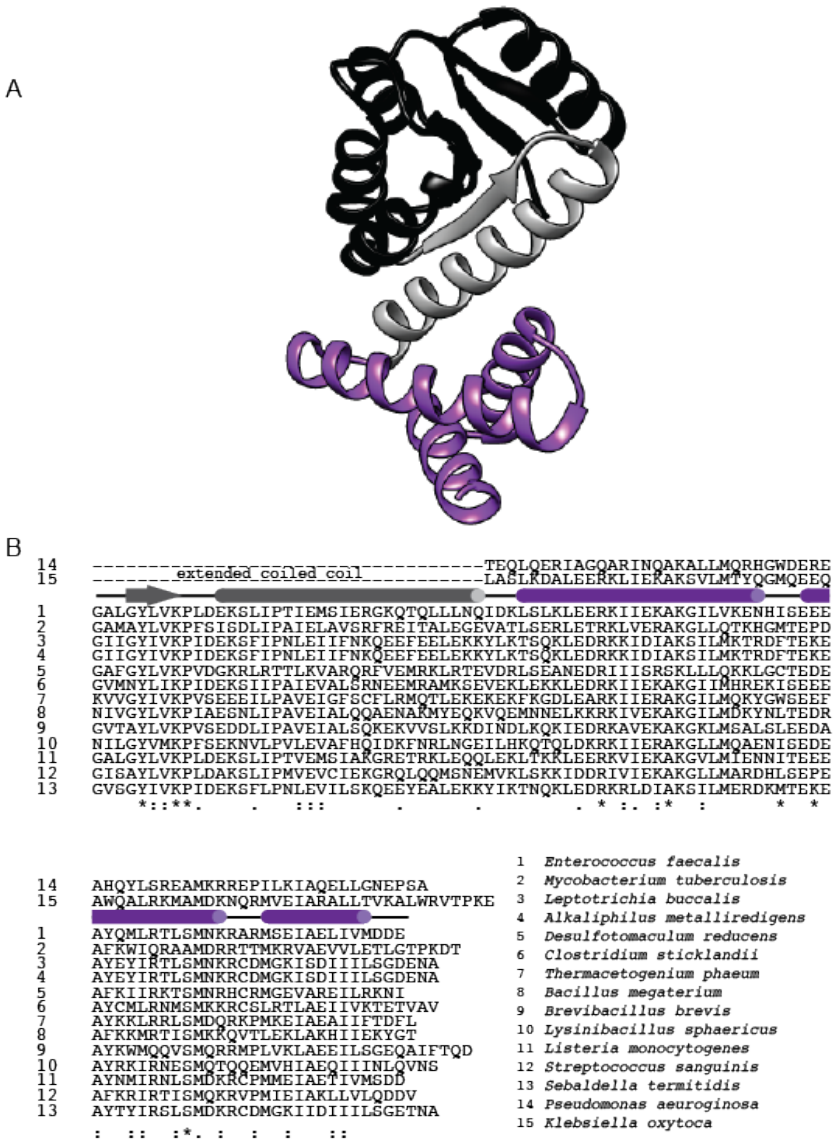

Figure 3-7. Alignment of putative EutV homologues. (A) Structural model of *E. faecalis* EutV. Modeled by threading *M. tuberculosis* Rv1626 via I-Tasser and Chimera (148, 150). Reciever domain in black, α-helix in grey, and ANTAR domain in purple. (B) Aligned *eut*-associated **response regulator ANTAR domain proteins ~1-14. *P. aeruginosa* and *K. oxytoca*, are not** aligned, but their ANTAR domain is shown above the secondary structure. Identical residues are denoted by an "*", highly conserved by a ":", and semi-conserved by a ".".

likely to be involved in intramolecular interactions within the protein core, such as the structural interactions formed between alpha helices. (Figure 3-8A, grey boxes). The remaining 15 were then chosen for mutagenesis. Specifically, these putatively conserved residues were mutagenized to alanines via site-directed mutagenesis (Figure 3-8A, blue boxes).

YFP output was measured for the alanine mutants relative to a strain that contained wild type EutV, and then quantified by FACS analysis (147). Of these mutants, four of the 15 had no effect on antitermination (Q124A, Q131A, E140A, and S171A), two had moderate effect on antitermination (N173A, R175A), and nine appeared to completely abolish antitermination activity (Y101A, K104A, M117A, R142A, K147A, E160A, Y164A, M172A, and M178A) (Figure 3-8B). To investigate whether intracellular abundance of the proteins was altered upon alanine mutagenesis, protein extracts were analyzed by Western blotting, using antisera to *E. faecalis* EutV. This revealed that several of the amino acids that exhibited a reduction in antitermination activity also exhibited reduced protein levels relative to WT (~40% or less; Y101A, K104A, E160A, Y164A, M172A, N173A, R175A, and M178A, data not shown). This suggested that the decreased antitermination activity of these particular mutants might not be due to the amino acid mutation, but rather to instability of the proteins *in vivo*. Of the three remaining mutants that exhibited full loss of antitermination activity, further inspection of the M117A mutation suggested the 117 residue could be involved in intramolecular interactions, thus leaving only R142A and K147A as representing the most critically important residues for antitermination activity.

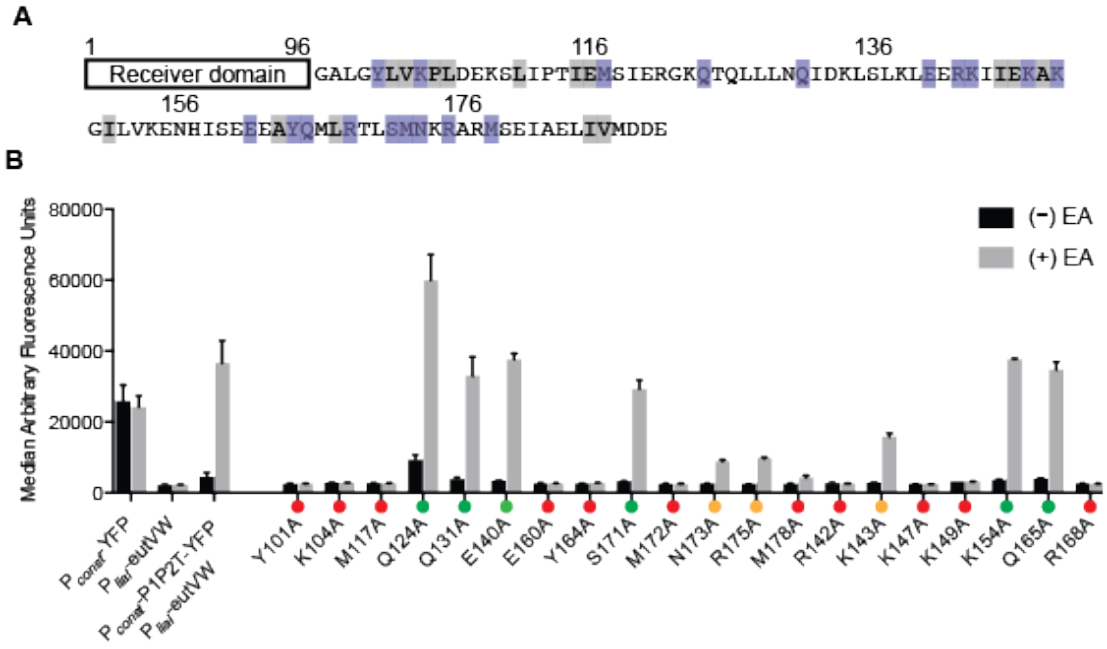

Figure 3-8. ***In vivo*** **antitermination activity of mutant EutV proteins.** (A) Schematic showing the ANTAR domain with conserved residues highlighted. Grey boxes denote hydrophobic amino acids predicted to be involved in intramolecular contacts, while purple boxes denote non-hydrophobic residues mutagenized to alanines. (B) FACS data showing affect of mutation on antitermination. Red dots denote loss of antitermination activity, yellow dots denote moderate loss of antitermination activity, and green dots denote no loss in antitermination activity. Reported median value, error bars denote SD of at least 3 replicate experiments.

Since structural models for both AmiR and NasR ANTAR domains have been published (135, 137, 151), and since our laboratory has predicted they are likely to recognize a similar RNA substrate to *E. faecalis* EutV, we inspected these protein structures for insight into residues that might be generally involved in RNA recognition (Figure 3-9). Similarly, a high-resolution X-ray crystallographic structure has been resolved for an ANTAR-containing protein (Rv1626) from *Mycobacterium tuberculosis*.

In the absence of amides, AmiC monomers bind separately to the AmiR dimer, causing a conformational change and therefore reducing AmiR RNA-binding activity (137). Similarly, in the absence of nitrates or nitrites, the ANTAR domain of NasR is unable to bind its RNA substrate (107, 109, 115, 135). Interestingly, the X-ray crystallographic structures revealed that the ANTAR portions of the monomers of NasR and AmiR are superimposable, despite dramatic differences in the dimerization interfaces of the overall proteins. NasR utilizes the C-terminus of the third helix in the ANTAR domain to assist with dimerization. In contrast, for the unbound state of AmiR, the first helix of the ANTAR domain and the extended α-helical region appear to participate in creating dimerization interfaces (137). However, only the monomeric structure of Rv1626 was resolved by crystallographic analysis, although its ANTAR portion superimposes well with structures of NasR and AmiR monomers (151).

Since both AmiR and NasR crystallographic structures are in the unbound/inactive state, and their dimeric configuration is completely different from one another, assumptions regarding general EutV structures must be viewed cautiously. We assume that upon activating the NIT domain of NasR, and with

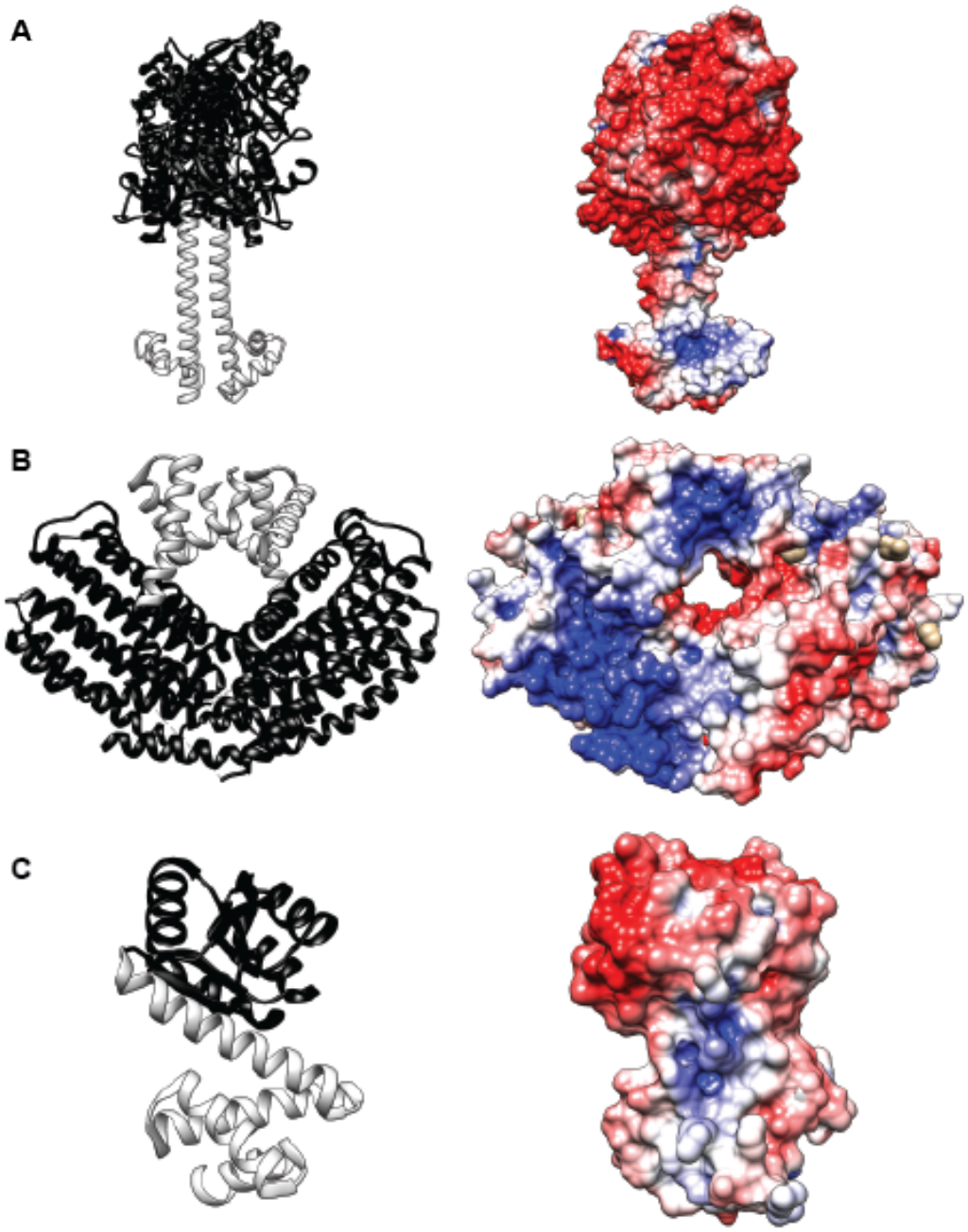

Figure 3-9. Electrostatic surface potential maps of AmiR, NasR, and Rv1626 X-ray crystal structures. Left are the ribbon structures, and right are the calculated electrostatic surface potential maps. Calculations were obtained using the Chimera PDB2PQR APBS (150). The ANTAR domain is colored in grey. (A) Crystal structure of AmiR, PDB: 1qo0, a putative positive patch referenced in the text is located in the bottom right in the ANTAR domain. (B) Crystal structure of NasR, PDB: 4akk, the putative positive patch referenced in the text is located at the top of the "handle". (C) Crystal structure of Rv1626, PDB:1s8n, the putative positive patch referenced in the text is located in the center of the image.

presence of amides to prevent AmiC silencing of AmiR, that conformational changes will occur that facilitate the ANTAR domains recognition of RNA. Furthermore, sequence comparison of the ANTAR domains from AmiR, NasR, Rv1626, and EutV indicate that the majority of residues that are strictly conserved are those that are part of the hydrophobic cavity of the domain.

The three crystal structures, modeled against Rv1626 EutV, suggested there may be a putative salt bridge at Q165A/R168A in EutV, and thus were mutagenized as follows: Q165A, R168A, Q165A/R168A, and Q165R/R168Q (Figure 3-8 and 3-10). Interestingly, Q165A showed no change in antitermination activity; however, R168A completely abolished antitermination. Not surprisingly Q165A/R168A also showed a drop in antitermination activity, but this can be explained by the mutation at R168 and not Q165 (Figure 3-8). Finally, Q165R/R168Q restored antitermination activity. These mutations argue that the putative salt bridge is unlikely to exist. Instead, the Q165R/R168Q mutation suggests that the general location and charge of R168 is likely to be important for antitermination activity.

3.3.4 Putative positively charged patch on surface of the ANTAR domain

Previously, calculations of electrostatic surface potential maps revealed that there may be a putative positive patch on the surface of the ANTAR domain of NasR (135). We found a similar patch in AmiR, and Rv1626 (*M. tuberculosis*) ANTAR domains (Figure 3-9). A similar positive patch is found in *E. faecalis* EutV. Interestingly, upon modeling the other *eut*-associated ANTAR-domain containing response regulators, we see the positive patch appears to be reasonably conserved (Figure 3-11).

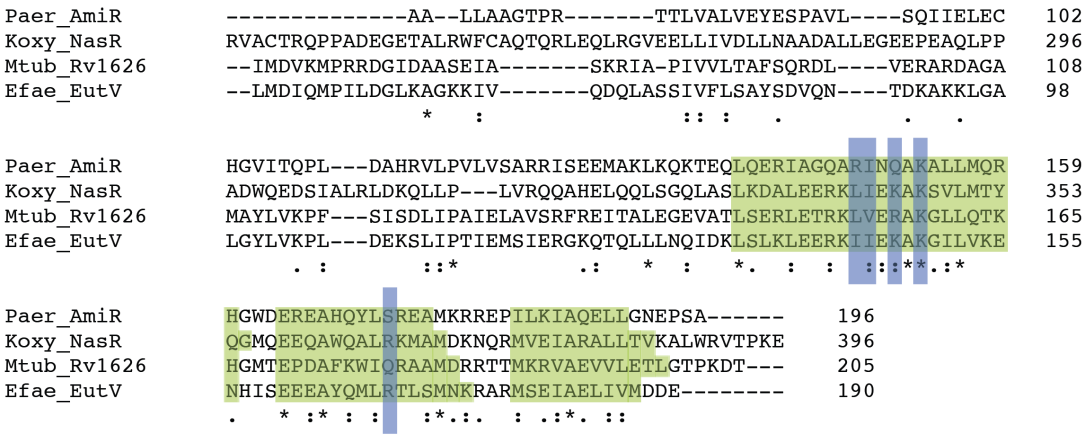

Figure 3-10. Putative positive patch in AmiR, NasR, Rv1626, and EutV. ANTAR helices are highlighted in green. Blue boxes denote the residues that are predicted to participate in formation of a putative positive patch.

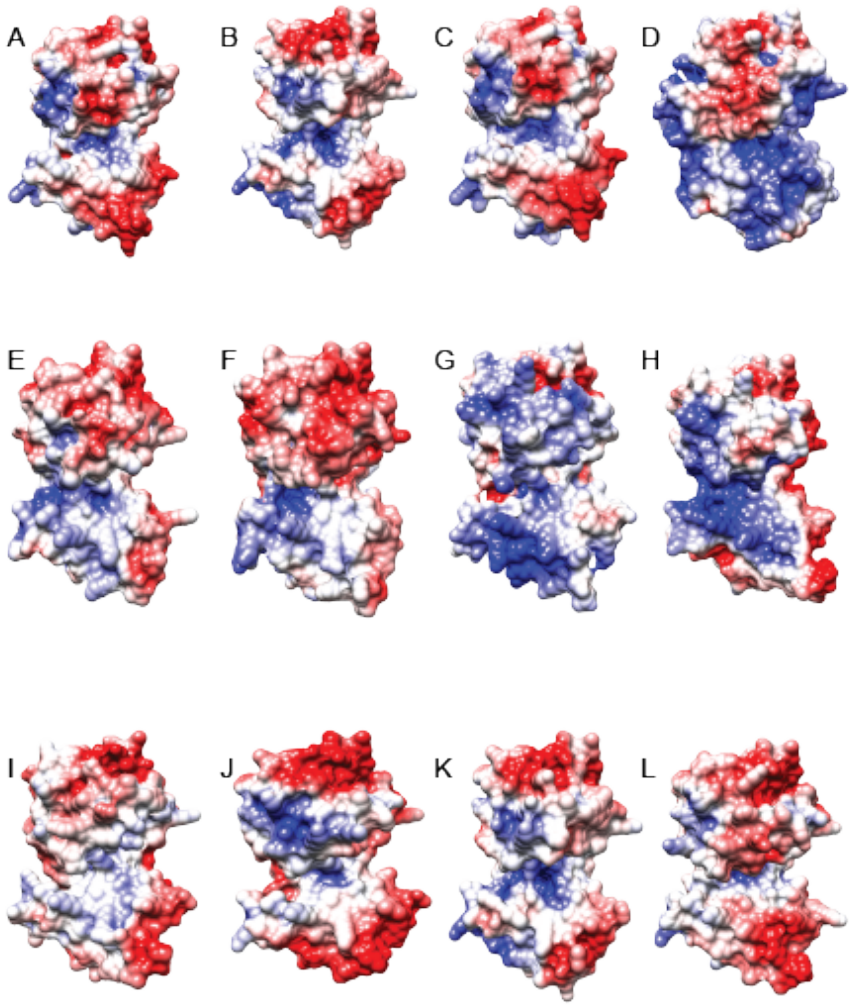

Figure 3-11. Electrostatic surface potential maps of putative EutV homologues. (A) EutV from *E. faecalis*.;(B) *L. buccalis*; (C) *A. metalliredigens*; (D) *D. reducens*; (E) *C. sticklandii*; (F) *T. phaeum*; (F) *B. megaterium*; (H) *B. brevis*; (I) *L. sphaericus*; (J) *L. monocytogenes*; (K) *S. sanguinis*; (L) *S. termitidis*. All proteins were modeled using I-Tasser, (148) and electrostatic surface potentials were calculated using the Chimera PDB2PQR APBS (150).

To investigate whether this positive patch was important for antitermination activity we chose to mutagenize the positively charged residues in this region to alanines: K143, K149 (R142 and R147 were previously caught in the screen for conservation), and an additional residue R154 was chosen for this screen due to its proximity to the positive patch. Of these alanine mutants K143A showed moderate loss in antitermination activity, K149A and R168A showed complete loss in antitermination activity, while K154A showed no change in antitermination activity (Figure 3-8). Since R142, K143, K147, K149, and R168 alanine mutants showed diminished antitermination activity we chose to make more subtle mutations to test whether the charge was important for antitermination (lysine to arginine and vice versa). If the charge is important for antitermination, we reasoned that replacing a similarly sized and charged amino acid with another would restore antitermination activity. Interestingly, K143R, K147R, K149R, and R168K were all capable of restoring antitermination activity (Figure 3-12). However, R142K did not restore antitermination activity, suggesting that the charge at R142 is not the only requirement for antitermination activity (Figure 3-12). This suggested that R142 is required for antitermination activity, and that the 142 amino acid may be important for RNA:protein interaction to promote antitermination activity, perhaps through direct RNA contacts or by affecting protein dimerization.

3.3.5 *nasF* ANTAR substrate and NasR cannot function in a heterologous host

Ramesh et al. showed that both *P. aeurginosa* AmiR, and *K. oxytoca*

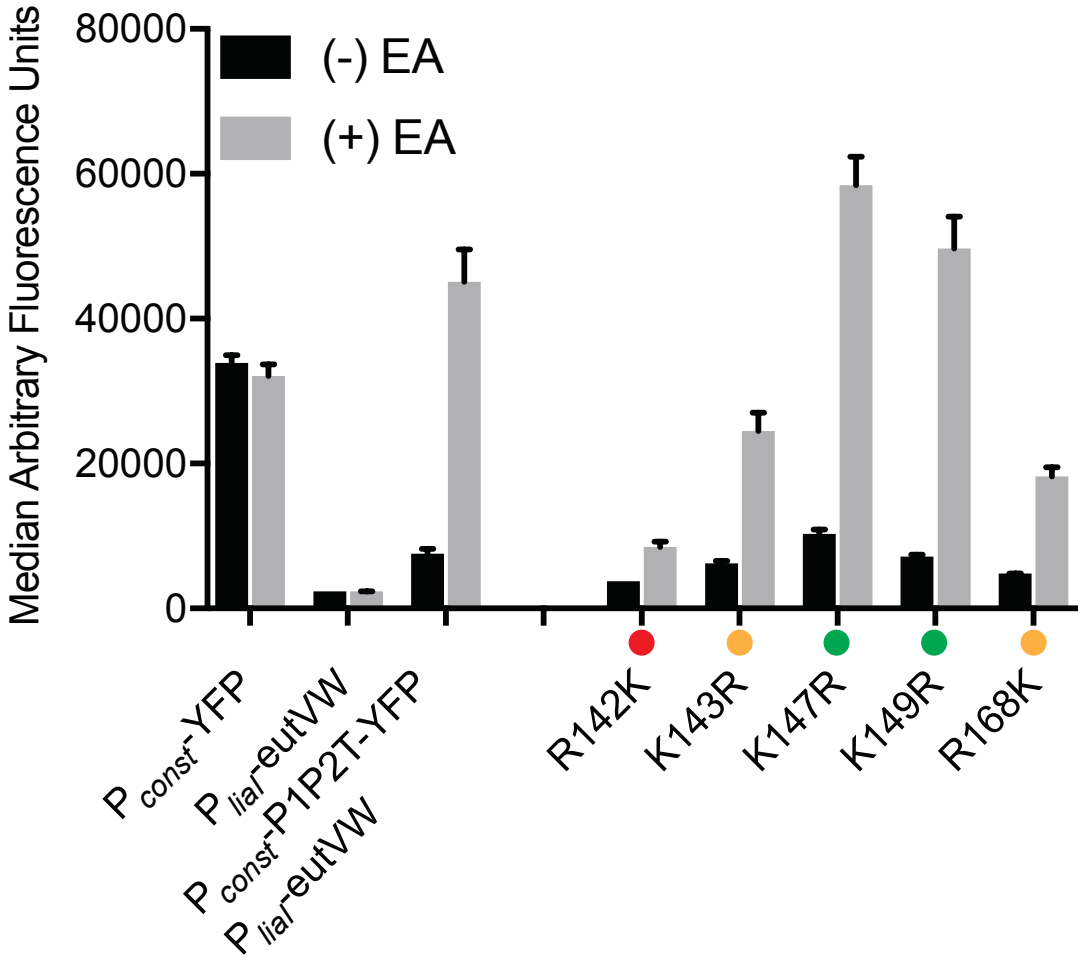

Figure 3-12. Positive patch is required for antitermination activity. FACS data showing affect of mutation on antitermination in the presence of 30 μg/mL bacitracin and presence or absence of 5 mM EA. Reported median value, error bars denote SD of at least 3 replicate experiments. Red dots denote loss of antitermination activity, yellow dots denote moderate loss of antitermination activity, and green dots denote no loss in antitermination activity. Reported median value, error bars denote SD of at least 3 replicate experiments.

NasR have similar two-hairpin motifs found in their respective genomes (114). We hypothesized that, given the high similarity between two-hairpin motifs in different organisms, *E. faecalis* EutV could promote antitermination at another organisms' ANTAR substrate. To test our hypothesis, we subcloned the *nasF* two-hairpin motif from *K. oxytoca* into the *yfp* reporter construct and tested whether EutV could promote antitermination at the *nasF* leader region.

As before, the strain containing *nasF*/EutV (*amyE::*P_{const} P1P2$_{nasF}$-*yfp*, *thrC::*P_{liaI} *eutVW*) were grown in rich media to mid-exponential and EutV was induced with 30 μg/mL bacitracin and 5 mM EA (Figure 3-13A, blue). After 40 minutes of incubation, cells were harvested and analyzed using analytical FACS (147). Analysis of our FACS data showed that EutV was unable to promote antitermination of *nasF* (Figure 3-13B). We hypothesize that this may be due to degradation of *nasF*, since there are likely to be RNases present in *B. subtilis* that are not present in *K. oxytoca*. Alternatively, the terminator in the *nasF* construct may be too strong, and thus preventing the two-hairpin motif from forming. However, we cannot ignore the possibility that there may be other accessory factors required for antitermination in *K. oxytoca* that are not required in *E. faecalis*.

To further test our hypothesis that ANTAR proteins and their substrates can be interchanged between organisms, we generated a bacitracin inducible NasR, to be paired with *E. faecalis* P_{const} P1P2-*yfp* reporter (Figure 3-13A, orange). This strain was cultured and prepared as the previous FACS experiment with one exception: since NasR is predicted to promote read-through of its RNA substrate in a nitrate-

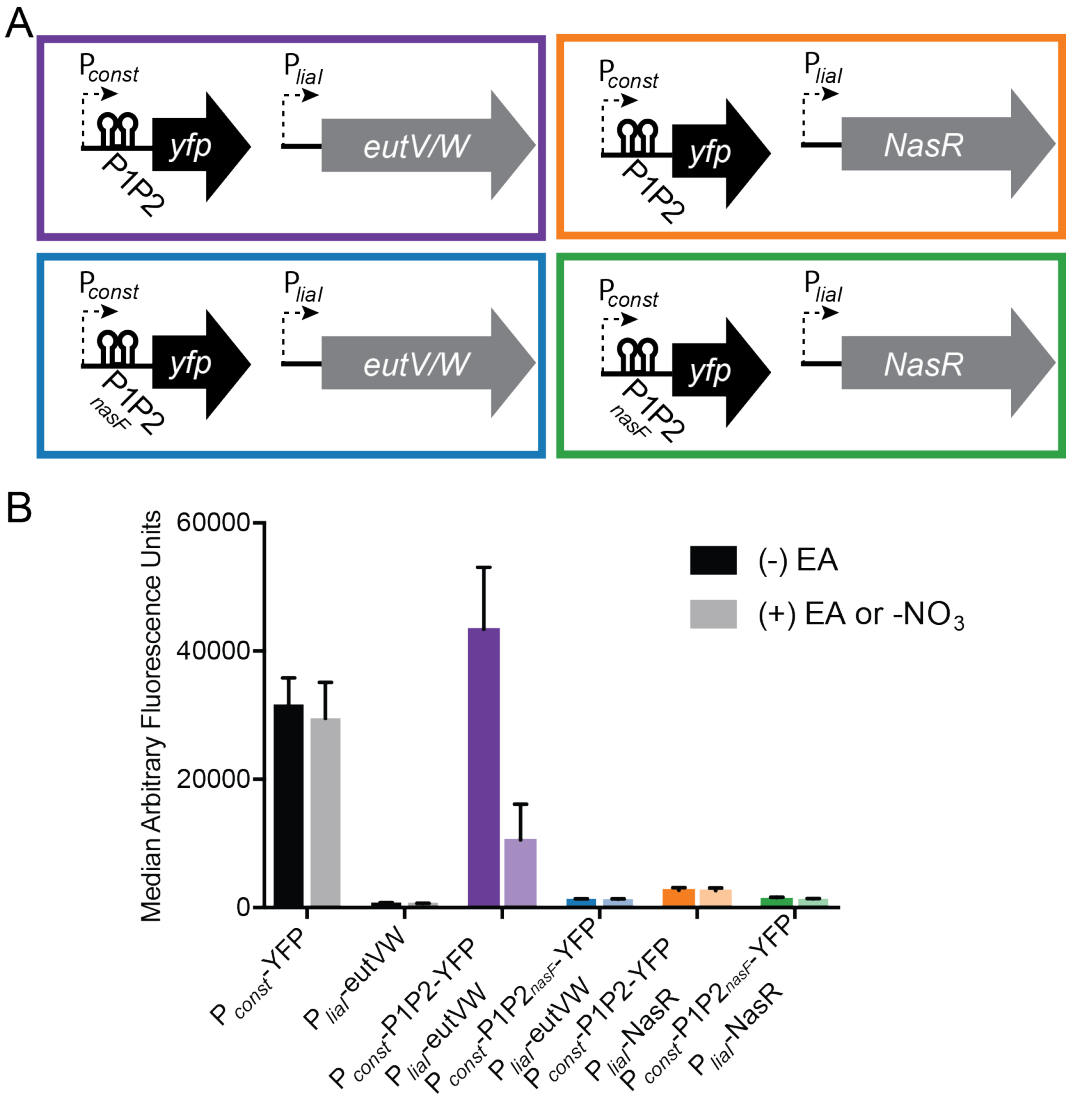

Figure 3-13 . *nasF* RNA and NasR protein do not support antitermination in a heterologous host. (A) Schematic of constructs. Purple is the original strain *amyE::*P_{const}P1P2-*yfp, thrC::*P_{liaI} EutVW. Blue: *amyE::*P_{const}P1P2$_{nasF}$-*yfp, thrC::*P_{liaI} EutVW. Orange: *amyE::*P_{const}P1P2-*yfp, thrC::*P_{liaI} NasR. Green: *amyE::*P_{const}P1P2$_{nasF}$-*yfp, thrC::*P_{liaI} NasR. (B) FACS data showing affect NasF and/or NasR have on antitermination in the presence of 30 µg/mL bacitracin and presence or absence of 5 mM EA or 10 mM $NaNO_3$. $NaNO_3$ was used for strains containing NasR only.

responsive manner, we induced its activity with 10 mM nitrate as previously reported by the Stewart lab (115, 116, 152). The two-hairpin motif from *E. faecalis* coupled with NasR was still unable to support antitermination (Figure 3-13B). We postulated that this might be due to an inability for nitrates/nitrates to cross the membrane into the cytosol to levels that are sufficient to promote antitermination activity. *B. subtilis* has only one nitrate transporter, *nasA*, which has only been inferred computationally, and not experimentally verified, or characterized (141, 142). This could explain why NasR might be unable to promote antitermination of the *E. faecalis* ANTAR substrate. Alternatively, *nasF* RNA and NasR may not function within a heterologous host, or perhaps required additional accessory factors. To discern whether it was incompatibility of the RNA:protein or issues with the heterologous host, we constructed a *nasF* two-hairpin reporter (*amyE::*P_{const} NasF- *yfp*) and an inducible NasR (*thrC::*P_{liaI} NasR) strain (Figure 3-13A, green). When we tested this strain for antitermination activity we saw that NasR was unable to promote antitermination of *nasF* leader region in a heterologous host (Figure 3-13B).

3.3.6 *In vitro* analysis of the ANTAR domain

Previous attempts addressing *in vitro* binding of the ANTAR domain to its RNA substrate have been complicated by EutVs' overall poor solubility and the overall high proportion of inactive protein after purification. In contrast, the Stewart lab already purified NasR and began to address binding affinity to its RNA substrate (109) in response to nitrate/nitrite binding. Therefore, we purified a His_{10}-MBP tagged NasR to test whether it bound the *nasF* two-hairpin motif. Binding was tested using differential radial capillary action of ligand assays (DRaCALA) by titrating

increasing amounts of the His_{10}-MBP-NasR dimer to trace levels of radiolabeled *nasF* two-hairpin motif (130). Upon addition of micromolar RNA concentrations, purified NasR bound to the two-hairpin motif in *nasF* RNA (Figure 3-14, A). Competition with 20 μM cold *nasF* RNA successfully competed away the interaction, while 5 μM and 20 μM of an unrelated RNA, mini-ykkC, was unable to compete away the interaction, indicating that binding of NasR is likely to be specific to the two-hairpin motif in *nasF*. As an additional confirmation of specificity, antisense-EutX (As-EutX) RNA was tested for binding activity, which revealed no binding activity for this RNA. In contrast, the apparent K_d value of the His_{10}-MBP-NasR-$P1P2_{nasF}$ leader RNA complex was estimated to be about 1.6 ± 0.2 μM (Figure 3-14, B), which is very similar to the previously reported K_d from the Stewart lab (109). At higher concentrations of protein, binding activity began to be reduced, which we speculate is due to the increased salts from suspensions of purified NasR. Based on the observation that maximal binding activity was observed between 1.25 and 2.5 μM *nasF* RNA to 5 μM NasR dimer, we speculate that the fraction of active NasR was between 40 and 50% of total purified protein.

To ascertain whether NasR exhibits properties similar to EutV, we investigated the functional relevance of a NasR R340 mutation, which corresponds to the EutV R142 mutation (Figure 3-10). To that end, we purified His_{10}-MBP-NasR R340A and R340K to test *nasF* RNA binding. Addition of increasing micromolar concentrations of the proteins to *nasF* RNA demonstrated that both mutations abolished RNA binding activity. Therefore, the amino acid identity at this position is important for both EutV and NasR proteins, consistent with a putative role in RNA

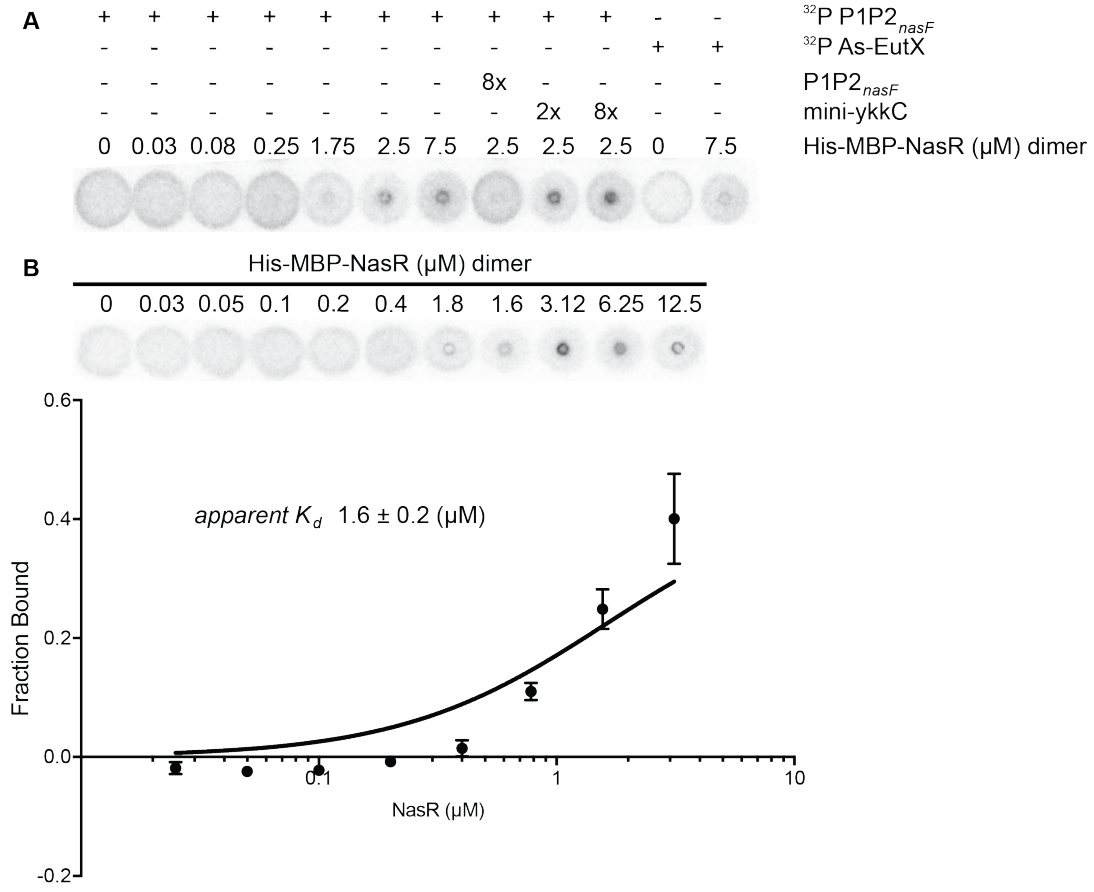

Figure 3-14. *nasF* RNA binds specifically to NasR. (A) Differential radial capillary action of ligand assay (DRaCALA) was used to assess NasR dimer binding to the P1P2$_{nasF}$ RNA. Unrelated RNA, mini-ykkC, was used as a negative control for binding specificity. Cold P1P2$_{nasF}$ RNA was used as a positive control for competition. (B) DRaCALA was used to assess the apparent K_d of the NasR-P1P2$_{nasF}$ RNA complex. 10 mM $NaNO_3$ was added to each reaction to activate NasR binding to P1P2$_{nasF}$ RNA.

binding interactions. (Figure 3-15 A and B, respectively). However, further studies are required to ensure that these mutations do not simply disrupt the structure of NasR.

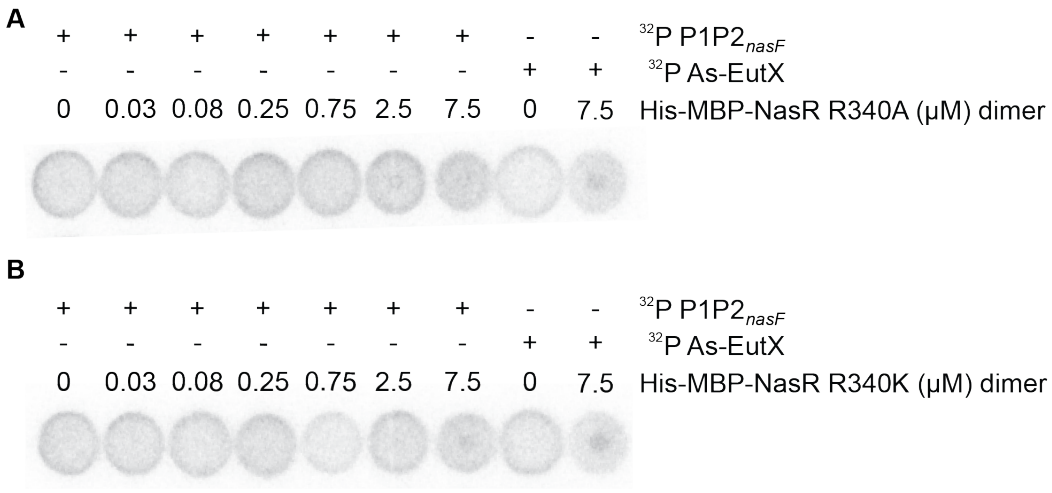

Figure 3-15. Mutation of NasR disrupts binding to *nasF* RNA. (A) DRaCALA was used to assess NasR R340A dimer binding to the *nasF* leader two-hairpin motif. (B) DRaCALA was used to assess NasR R340K dimer binding to the *nasF* leader two-hairpin motif. 10 mM $NaNO_3$ was added to each reaction to activate NasR binding to *nasF* RNA.

3.4 Conclusion

Our approach to recapitulating this circuit has allowed us to test whether EutX was able to sequester EutV without any accessory factors. Indeed, EutX was able to sequester EutV; however, it appears that in a heterologous host it may be subject to posttranscriptional problems, such as elevated RNA degradation.

One residue, R142, was found to be well conserved in NasR, Rv1626, EutV and other Eut-associated ANTAR domain containing proteins. When R142 was mutated to an alanine, antitermination activity was eliminated, suggesting the residue provides an important role for antitermination function. Upon mutating R142 to a lysine, which should maintain the charge and relative size of the amino acide side chain, antitermination activity was not restored. This suggested that the charge alone was not sufficient for antitermination activity. For NasR the corresponding residue is R340, and mutating this residue to an alanine or a lysine, much like the *in vivo* EutV studies, abolished binding, suggesting that this amino acid is required for antitermination activity, possibly through RNA-binding interactions. However, further studies on the NasR protein are required to ensure that these mutations do not simply disrupt its tertiary structure.

The catabolism of ethanolamine by *E. faecalis* has been proven to be quite complex with the use of a TCS (Fox et al. 2009; Ramesh et al. 2012) to sense ethanolamine, promote antitermination and therefore gene expression of the *eut* gene cluster. In addition, a sRNA, EutX, senses AdoCbl and functions to sequester EutV when AdoCbl levels are low (132). Experiments in this chapter demonstrate that the TCS, EutV/W, and an ANTAR substrate can recapitulate antitermination in a

heterologous host, *B. subtilis*. The tools we created will allow us to continue studying the ANTAR protein determinants required for antitermination in more detail.

Chapter 4: *Enterococcus faecalis* can utilize ethanolamine under anaerobic and aerobic conditions

4.1 Author contribution

Data gathered by Gebbie, M. are in Figs 4-1 through 4-14, data gathered by Weiss, C.A., are in Fig 4-4, data gathered by Dr. Kaval, K., are in Figs. 4-5, 4-8, 4-10, and 4-12, data gathered by Dr. Debroy, S., are in Figs. 4-13.

4.2 Introduction

Organisms that inhabit the gastrointestinal (GI) tract must exhibit metabolic flexibility, such as the ability to catabolize ethanolamine (EA). EA is a highly accessible source of nitrogen and carbon as it is a breakdown product of phosphatidylethanolamine, an abundant phospholipid in both mammalian and bacterial cell membranes (76, 153). It is believed that the host diet and the bacterial and epithelial cells in the intestines provide rich sources of EA in nature (76, 77, 153). A number of bacteria that inhabit the GI-tract and other environments, including, but not limited to species of *Salmonella*, *Enterococcus*, *Klebsiella*, *Mycobacterium*, *Clostridium*, *Escherichia*, and *Pseudomonas*, can utilize EA as a sole source of carbon and/or nitrogen (Del Papa and Perego 2008; Scarlett and Turner 1976). EA is catabolized into ammonia and acetaldehyde by an ethanolamine ammonia lyase, EutBC, where the ammonia can serve as a cellular supply of reduced nitrogen, and the acetaldehyde can be converted into the metabolically useful compound acetyl-CoA (155). The ethanolamine ammonia lyase requires adenosylcobalamin (AdoCbl) as an important cofactor for this initial breakdown of EA.

Previous studies on *E. faecalis* by the Perego lab investigated whether the presence of oxygen affects utilization of EA, by growing WT *E. faecalis* in a modified minimal media and measuring growth of the bacterial culture (111). They found that *E. faecalis* could grow both in aerobic and anaerobic conditions when grown in the modified minimal media supplemented with 100 mM glucose. As expected, when no carbon source was supplemented, no growth was observed in either condition. EA alone did not support *E. faecalis* growth in the modified minimal media in either aerobic or anaerobic condition. However, the Perego lab found that growth of *E. faecalis* cultures could be enhanced by addition of EA when the media was also supplemented with AdoCbl, but only in an anaerobic environment. The restriction of growth in the presence of EA to only anaerobic conditions suggested oxygen might contribute to genetic regulation of the EA catabolism gene cluster. Based on this observation, we sought to investigate whether anaerobiosis plays a role in *eut* gene regulation and EA utilization in *E. faecalis*.

4.3 Results and Discussion

4.3.1 *E. faecalis* utilizes EA aerobically

There have been multiple reports on EA utilization in *E. faecalis* and its use as a carbon or nitrogen source (111, 156, 157); however, previous studies were limited by the microbe's complex growth requirements (111). These previous studies used a modified minimal media comprised of M9 salts, HEPES, yeast extract, enzyme cofactors, and essential amino acids. The media was then supplemented with AdoCbl and EA under anaerobic conditions (111) to support growth. However, the exact composition of this media was complicated by the addition of yeast extract, an

undefined source of carbohydrates and amino acids. We sought to better characterize *E. faecalis* growth and utilization of EA, so we designed and purchased a rigorously defined chemical medium mixture (Table 1), similar to a defined medium that has been used for growth of other Gram-positive cocci (158). When inoculated in CDM supplemented with 25 μM AdoCbl in the absence of any carbon source, WT *E. faecalis* (OG1RF) cultures failed to grow (data not shown). However, when the organism was inoculated in CDM supplemented with both 0.2% glucose and 25 μM AdoCbl, cultures grew robustly (Figure 4-1, black line). Therefore, we conclude that the CDM is fully sufficient to support growth of *E. faecalis* upon addition of a carbohydrate source and with addition of AdoCbl.

Table 1. Composition of chemically defined media (CDM)

Component	*Grams/Liter*
Adenine Sulfate, $2H_2O$	0.02
DL-Alanine	0.1
L-Arginine, FB	0.1
L-Asparagine, Anhydrous	0.1
L-Aspartic Acid	0.1
Biotin	0.0002
Calcium Chloride, Anhydrous	0.0005
D-Calcium Pantothenate	0.002
L-Cystine 3HCl	0.0652
Folic Acid	0.0008
L-Glutamic Acid	0.1
L-Glutamine	0.2
Glycine	0.1
Guanine, HCl, H2O	0.02
L-Histidine, FB	0.1
Hydroxy L-Proline	0.1
L-Isoleucine	0.1
L-Leucine	0.1
L-Lysine, HCl	0.1249
Magnesium Sulfate, Anhydrous	0.3419

L-Methionine	0.1
Niacinamide	0.0001
β-NAD, $3H_2O$	0.0025
Para-Aminobenzoic Acid	0.0002
L-Phenylalanine	0.1
Potassium Phosphate, Dibasic, Anhydrous	0.2
Potassium Phosphate, Monobasic, Anhydrous	1
L-Proline	0.1
Pyridoxal, HCl	0.001
Pyridoxamine, 2HCl	0.001
Riboflavin	0.002
L-Serine	0.1
Sodium Acetate, Anhydrous	2.7126
Sodium Phosphate, Monobasic, H_2O, ACS	7.35
Sodium Phosphate, Dibasic, Anhydrous	3.195
Thiamine, HCl	0.001
L-Threonine	0.2
L-Tryptophan	0.1
L-Tyrosine, 2Na, $2H_2O$	0.1442
Uracil	0.02
L-Valine	0.1
Add to Prepared Sterilized Medium:	
Carbohydrate	0.2% Glucose or Ribose
Ferric Nitrate, $9H_2O$ 0.7 mg/mL (good for one month)	2.848 mL/L (4.8 μM)
Ferrous Sulfate, $7H_2O$ 1 mg/mL (good for one month)	10 mL/L (36 μM)
Manganese Sulfate, H_2O 1.7 mg/mL	4 mL/L (40.2 μM)
$NaHCO_3$ (125 mg/mL) **make fresh**	40 mL/L
L-Cysteine (35.4 mg/mL) **make fresh**	40 mL/L

(158)

While addition of glucose supported full growth of *E. faecalis*, addition of 25 μM AdoCbl and 10 mM EA only resulted in minimal growth yields for cells cultured in anaerobic conditions (Figure 4-1, light blue). In an attempt to enhance growth in the presence of AdoCbl and EA we added ribose, which we considered to be a moderate, non-PTS (phosphotransferase system) sugar (Figure 4-1,dark blue line).

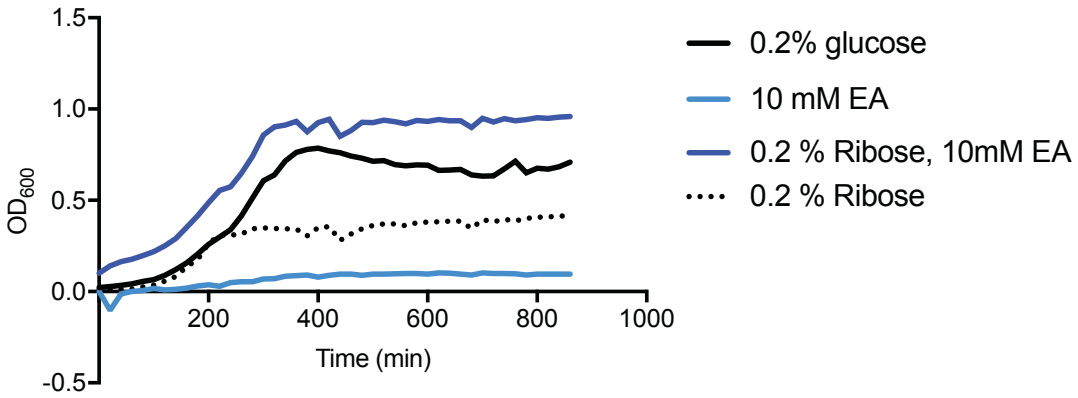

Figure 4-1. Growth of *E. faecalis* in CDM. WT *E. faecalis* exhibits enhanced growth in the presence of glucose (black), but not in EA (blue) alone or ribose alone (black dotted). 0.2% ribose and 10 mM EA supports anaerobic growth of WT *E. faecalis*. OG1RF was grown in CDM supplemented with 25 μM AdoCbl and 0.2% glucose or 10 mM EA or 0.2% glucose and 10 mM EA.

Under these conditions, growth increased to levels comparable to the glucose culture (Figure 4-1, light blue line). Based on its defined composition and its ability to support *E. faecalis* growth, all experimentation on *E. faecalis* in our laboratory was shifted to rely upon CDM supplemented with 0.2% ribose, 25 μM AdoCbl, and, when desired, 33 mM EA.

To test whether *E. faecalis* could utilize EA aerobically, we inoculated CDM supplemented with 0.2% ribose and 25 μM AdoCbl in the presence or absence of 33 mM EA with overnight cultures of a WT *E. faecalis*, EFKK6 (an OG1RF strain that contains an *eutM:gfp* integrated at a nonenssential locus; see Table 8). The bacterial cells were first prewashed in CDM to eliminate contaminating nitrogen/carbon sources, and inoculated to a starting OD_{600} of 0.1. The cultures were incubated at 37°C, while shaking to promote oxygenation, and growth was monitored via klett colorimeter every 30 minutes until stationary phase was reached. Under these conditions, *E. faecalis* cultures grew poorly in the absence of EA (Figure 4-2 B, empty circles). However, growth was moderately enhanced with addition of EA, as evidenced by an overall increased OD_{600} for stationary phase cells (Figure 4-2 B, filled circles). Interestingly, aerobic growth of this strain in the presence of EA lead to slower growth rate, 141 ± 3 min per doubling, while the absence of EA was faster at 69 ± 9 min (Table 2). We hypothesize that the slower growth rate in the presence of EA in an aerobic environment may be influenced by the taxing nature of transcribing, and translating the entire ~17 kb *eut* gene cluster required for EA utilization. However, while the growth rate was slower, the total number of generations during logarithmic phase appeared to reach 2.6, while the absence of EA reached only 1.6

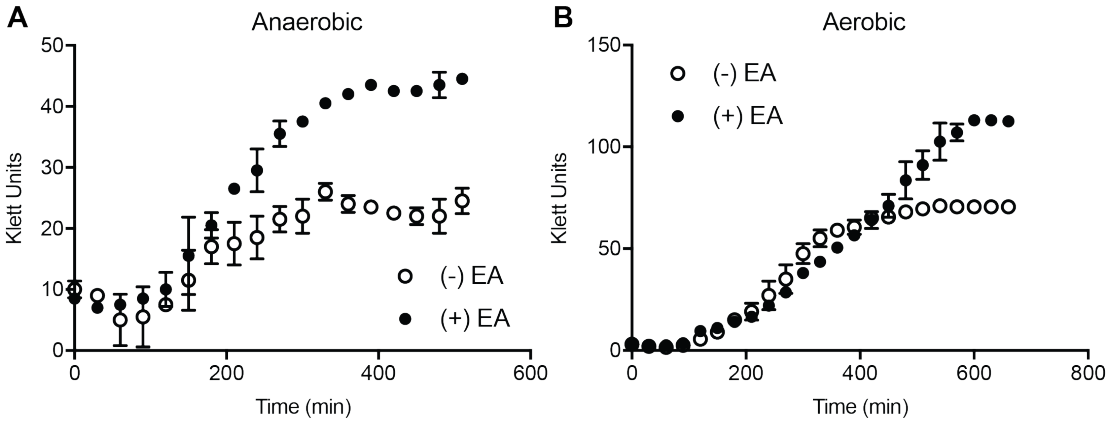

Figure 4-2. *E. faecalis* utilizes ethanolamine aerobically in CDM. (A) Growth curves of EFKK6 in anaerobic cultures, standing. (B) Growth curves of EFKK6 grown in shaking cultures. (A-B) Both cultures were grown in chemically define media (158) supplemented with 0.2% ribose and 25 μM AdoCbl in the presence or absence of 33 mM EA.

generations.

Table 2. Analysis of EFKK6, OG1RF, and ΔEutX growth in CDM

	Anaerobic			Anaerobic			Anaerobic	
EFKK6	G	n	OG1RF	G	n	*ΔeutX*	G	n
(-) EA	117 ± 4	1.8		160 ± 47	1.6		129 ± 5	1.9
(+) EA	71 ± 13	2.2		134 ± 1	2.5		1301± 17	3
	Aerobic			Aerobic			Aerobic	
EFKK6	G	n	OG1RF	G	n	*ΔeutX*	G	n
(-) EA	69 ± 9	2.6		86 ± 8	3.3		94 ± 4	2.6
(+) EA	141 ± 3	3.9		115 ± 4	3.9		79 ± 13	3.9

Data shown are mean ± SD of two independent samples.
G is generation time in min/doubling
n is number of generations

4.3.2 Complex media components inhibit aerobic EA utilization

Previous studies on EA utilization in *E. faecalis* employed a modified minimal media (111, 112) that contained 2 g/L yeast extract. Their results suggested that *E. faecalis* could only utilize EA under anaerobic conditions; however, given the added yeast extract, we could not be certain that EA was the only carbon/nitrogen source being used. We postulated that the presence of yeast extract might somehow be inhibitory to EA utilization in an aerobic environment. To test this, *E. faecalis* EFKK6 was cultured in CDM supplemented with 0.2% ribose and 25 uM AdoCbl, in the presence or absence of EA and 2 g/L yeast extract. As before, we observed enhanced growth in the presence of EA when yeast extract was omitted (Figure 4-3A, solid vs dotted line), relative to cultures that lacked EA. However, when this strain was grown in medium supplemented with yeast extract it did not exhibit growth enhancement in the presence of EA (Figure 4-3B, filled vs empty cirlces). On the contrary, while both the presence and absence of EA lead to the same number of generations, ~2.4 (Table 3), the growth rates were considerably different. In the absence of EA the culture exhibited a doubling time of approximately 114 ± 13

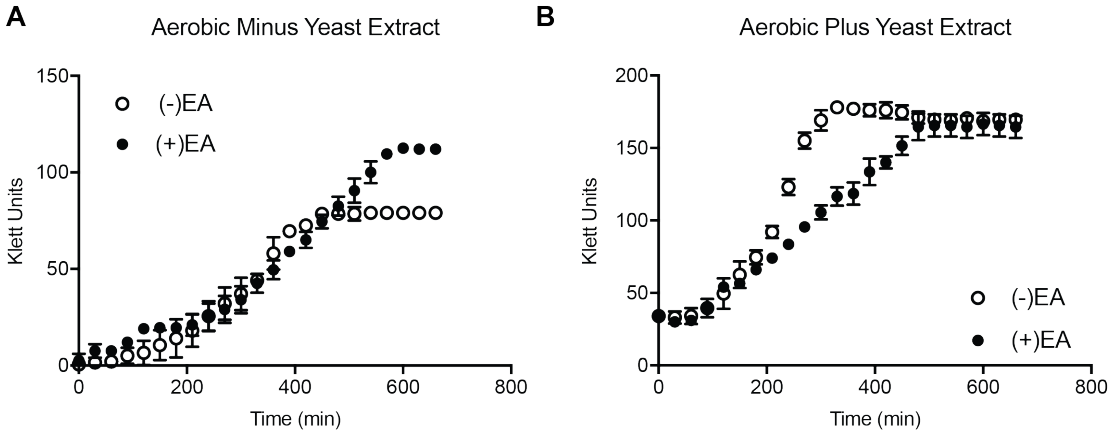

Figure 4-3. Yeast extract inhibits the ability of *E. faecalis* to utilize EA. (A) Growth curves of WT *E. faecalis* grown in shaking cultures. (B) Growth curves of WT *E. faecalis* in anaerobic cultures, standing. (A-B) Both cultures were grown in chemically define media (158) supplemented with 2 g/L yeast extract (111), 0.2% ribose and 25 μM AdoCbl in the presence or absence of 33 mM EA.

minutes, while in the presence of EA this increased to 175 ± 6 minutes (Table 3). This suggests that the presence of yeast extract inhibits utilization of EA under aerobic conditions, potentially contributing to the previously reported requirement for anaerobic conditions.

Table 3. Analysis of EFKK6 growth in CDM supplement with yeast extract

	Aerobic	
	G	n
(-) EA	114 ± 13	2.4
(+) EA	175 ± 6	2.4

Data shown are mean ± SD of two independent samples.
G is generation time in min/doubling
n is number of generations

4.3.3 Visualization of Eut microcompartments

The enhanced growth observed for *E. faecalis* in the presence of EA under both aerobic and anaerobic conditions suggested that EA is being catabolized. Many organisms that catabolize EA do so within the fully assembled microcompartment shell (159). Therefore, we speculated that catabolism of EA in *E. faecalis* may require the formation of Eut microcompartments under both aerobic and anaerobic growth conditions. To investigate this hypothesis, we attempted to apply two approaches: (1) live-cell imaging of microcompartments using fluorescence microscopy and (2) transmission electron microscopy.

We reasoned that a fluorescently-tagged microcompartment shell protein might result in fluorophores tightly associated with the microcompartment organelle. There is precedence for this idea; a prior proof-of-principle study on the imaging of carboxysomes relied on expression of a microcompartment shell protein that was

translationally fused to GFP (160). In that study, multiple GFP-intensive spots were observed inside cells, which the authors interpreted as corresponding to locations of carboxysomes. Therefore, for imaging of *E. faecalis* Eut microcompartments, we attempted to develop a similar approach. Such an approach would, in the long-term, enable the investigation of many different aspects of the biology of bacterial microcompartments, from genetic regulation, to biogenesis, to localization, and perhaps even to disassembly. To that end, a structural shell protein for Eut microcompartments, EutM, was fused in-frame to GFP, and integrated into the endogenous *eut* locus. After culturing this strain in the presence of EA under aerobic conditions, an aliquot of the cells at mid-exponential was removed and placed on an agarose pad (1.5% agarose in 1X PBS), and subjected to analysis by fluorescence microscopy. Interestingly, instead of seeing multiple foci, as predicted for microcompartment formation (132), we saw a single, exceptionally bright focal point for GFP fluorescence (Figure 4-4), which for diplococcal cells always appeared to be positioned at the cellular pole opposite to the cellular division septum. To investigate these cells further, samples of cells cultured under the same growth conditions were subjected to transmission electron microscopy (TEM). The corresponding micrographs revealed widespread cytological features that are consistent with septal defects (Figure 4-5, left vs right). No apparent microcompartments were observed for this strain. Also, for the EutM-GFP-containing strain there are aberrant cellular features at the poles opposite of the division septum that look qualitatively similar to large sites of protein aggregation. These results combined suggested that this particular strain exhibits morphological defects, most likely due to aggregation of the

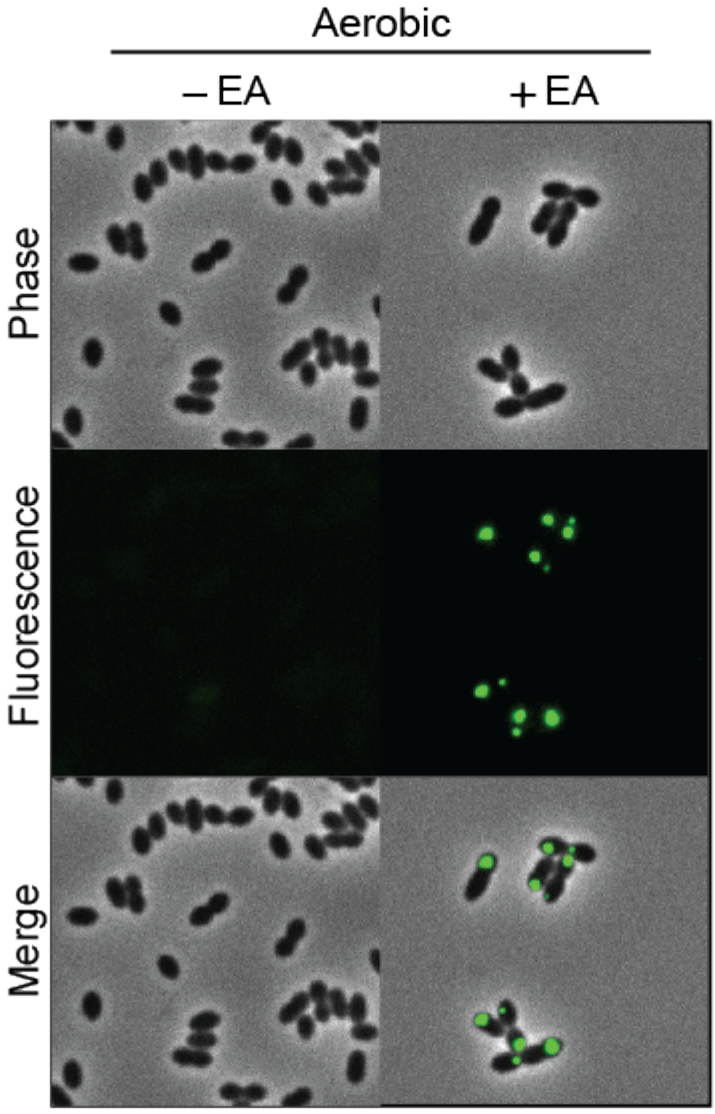

Figure 4-4. Endogenously integrated *eutM:gfp* hinders Eut microcompartment formation. This strain was grown in CDM supplemented with 0.2% ribose and 25 μM AdoCbl in the presence or absence of 33 mM EA. Cells were collected at early stationary and washed in 1x PBS, and placed on a 1.5% agarose pad for imaging. In the presence of ethanolamine only, green foci are found localized to the poles of cells. Images collected by Cordelia A. Weiss.

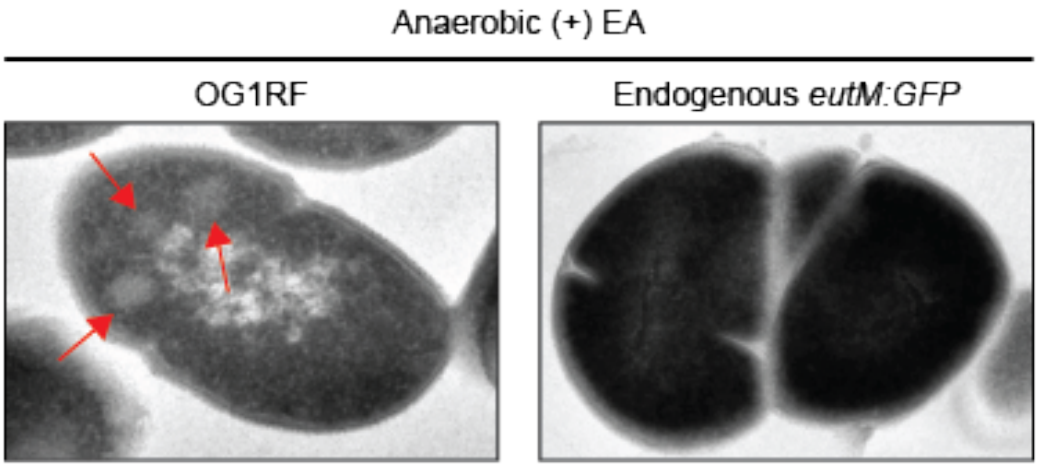

Figure 4-5. TEM images of the endogenously integrated *eutM:gfp* shows septal defects. This strain was grown in CDM supplemented 0.2% ribose and 25 μM AdoCbl in the presence of 33 mM EA. Cells were collected at early stationary and washed in 1x PBS, and resuspended in 3% glutaraldehyde. TEM images collected by Dr. Karan Kaval.

GFP-tagged EutM protein.

One possible reason why the EutM-GFP strain failed to form functional microcompartments is that the sole cellular source for EutM was the GFP fusion protein. It is possible that this fusion protein is too disruptive to assembly of microcompartments. However, it is also possible that a second copy of EutM, fused to GFP and positioned ectopically on the genome, might become incorporated within microcompartments at a lower abundance than the EutM protein expressed from the *eut* locus. In that instance, we reasoned, the GFP-containing protein could be simply incorporated at a low rate into microcompartments. To test this hypothesis we generated a copy of *eutM-gfp* that was integrated into a distally located ectopic locus. This *eutM-gfp* gene was maintained under genetic regulatory control by the endogenous *eutM* promoter region, including the EutV-responsive antitermination site just downstream of the *eutM* promoter. Therefore, EutM-GFP should be expressed from an ectopic locus under the same conditions that the *eut* gene cluster is activated. Analysis of this strain by fluorescence microscopy revealed a moderate increase in overall GFP in the absence of EA (Figure 4-6, minus EA), suggesting that the *eutM-gfp* gene was indeed responsive to the EutV regulatory protein. Moreover, multiple discrete fluorescent foci were observed within cells that had been cultured in the presence of EA, for both exponential (data not shown) and stationary growth phase, and for cells that had been cultured under aerobic and anaerobic conditions (Figure 4-6, plus EA). We did not see any discernable difference in total fluorescence between exponential or stationary phase cultures. This result is initially exciting as it suggests that Eut microcompartments can indeed be imaged via fluorescence

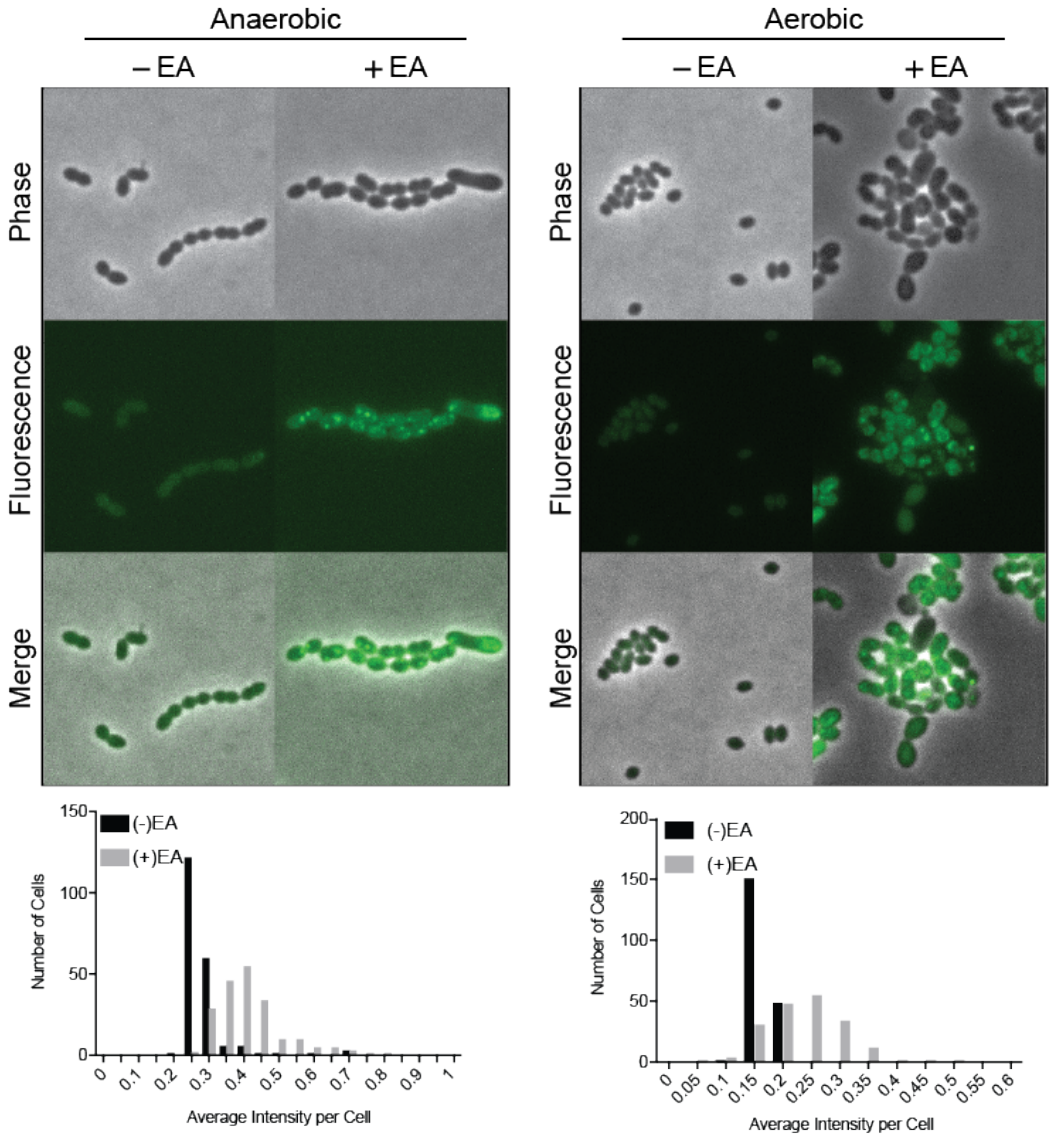

Figure 4-6. Fluorescent microscopy of EFKK6 during growth. Green foci, indicative of Eut microcompartments are found in the presence of EA. The strain containing the ectopically integrated *eutM:gfp* was grown in CDM supplemented 0.2% ribose and 25 μM AdoCbl in the presence of absence of 33 mM EA. Cells were collected at early stationary and washed in 1x PBS, and placed on a 1.5% agarose pad for imaging on the Zeiss Observer. Z1. Approximately 200 cells were quantified using Oufti software (146). Images collected by Dr. Wade Winkler.

microscopy, assuming that the fluorescent foci correspond to fully formed microcompartments. Therefore, we interpret our data as showing that *E. faecalis* is likely to form Eut microcompartments under both aerobic and anaerobic conditions in order to catabolize EA, and that anaerobiosis is not likely to be required for EA catabolism.

Fluorescence microscopy has distinct limitations, specifically that the threshold of resolution is 150 nm, which is coincidentally the average size of bacterial microcompartments. Thus, we must be cautious about interpreting the discrete foci that we observed as strictly corresponding to the formation of Eut microcompartments. Instead, a more cautious interpretation until higher resolution microscopy approaches are utilized is that an increase in GFP corresponds to confirmation that the *eut* gene cluster has been upregulated, and that the foci may correspond to microcompartments. Nonetheless, observation of EA-dependent, fluorescent foci under aerobic conditions are consistent with production of microcompartments.

To further test whether Eut microcompartments are actually being formed we decided to employ a higher resolution microscopy, TEM. Cells were cultured under the conditions described above and treated as appropriate for microtome sectioning and TEM analyses. Inspection of the corresponding micrographs revealed the presence of putative microcompartments when bacterial cells were cultured in the presence of EA, regardless of the presence or absence (data not shown) of oxygen (Figure 4-7). Furthermore, no microcompartments are detected when cells were cultured in the absence of EA. Quantification of the microcompartments in the TEM

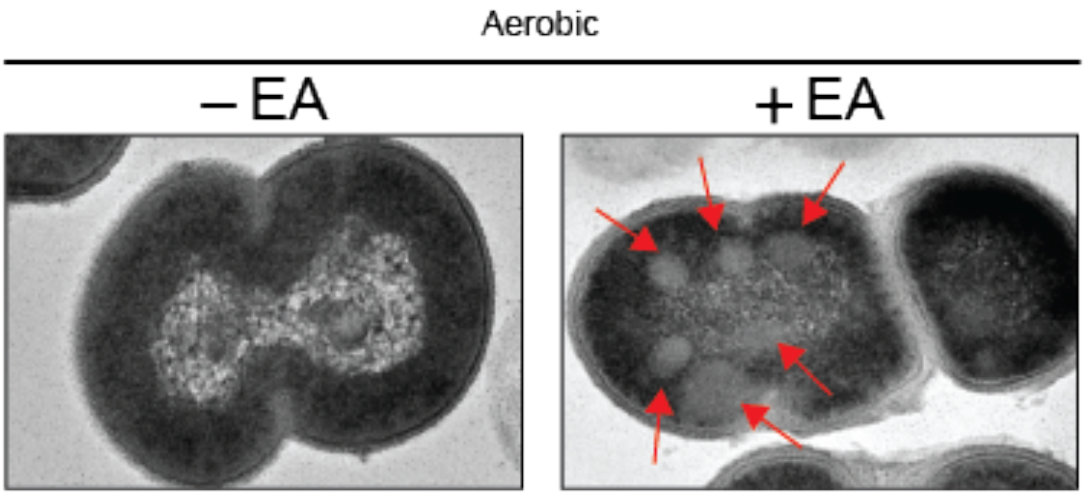

Figure 4-7. TEM images of EFKK6. EFKK6 was grown in CDM supplemented 0.2% ribose and 25 μM AdoCbl in the presence of absence of 33 mM EA. Cells were collected at early stationary and washed in 1x PBS, and resuspended in 3% glutaraldehyde. TEM images collected by Dr. Karan Kaval.

micrographs indicated that EFKK6 produces approximately 1.7 ± 0.1 microcompartments/cell (Table 4). However, TEM micrographs of cells cultured in the presence of EA demonstrated contorted cells that appear to be stressed, and that appear to have expressed biofilm extracellular polysaccharides that surround them. This suggests that while the ectopically-integrated *eutM-gfp* is likely to support Eut microcompartment formation, expression of the *eutM-gfp* appears to have adverse affects on cellular morphology, including possible septal defects.

Our data on the strain containing an ectopically integrated *eutM-gfp* (EFKK6), suggested that it might exhibit cytological defects; possibly as a result of the *eutM-gfp* fusion. Therefore, we sought to compare EFKK6 with *E. faecalis*, which is the same but simply lacks the *eutM-gfp* fusion. We began by examining growth of OG1RF in CDM supplemented with 0.2% ribose and 25 μM AdoCbl in the presence or absence of EA. This revealed that while EFKK6, the strain containing the ectopically integrated *eutM-gfp*, grew at a rate of 117 ± 4 (Table 2) minutes/doubling anaerobically in the absence of EA, OG1RF grew at a rate of 160 ± 47minutes/doubling (Figure 4-8A, filled circles, and Table 2), taking roughly 15-30% more time per generation. This difference in anaerobic growth rate was also apparent in the presence of EA, when EFKK6 grew at a rate of 71 ± 13 (Table 2) minutes/generation and OG1RF grew at a rate of 134 ± 1 (Figure 4-8A, filled circles, Table 2). This disparity in growth rates decreased under aerobic conditions in the absence of EA when EFKK6 exhibited a growth rate of 69 ± 9 and OG1RF at 86 ± 8 minutes per generation (Figure 4-8B, filled circles, Table 2). However, under aerobic growth in

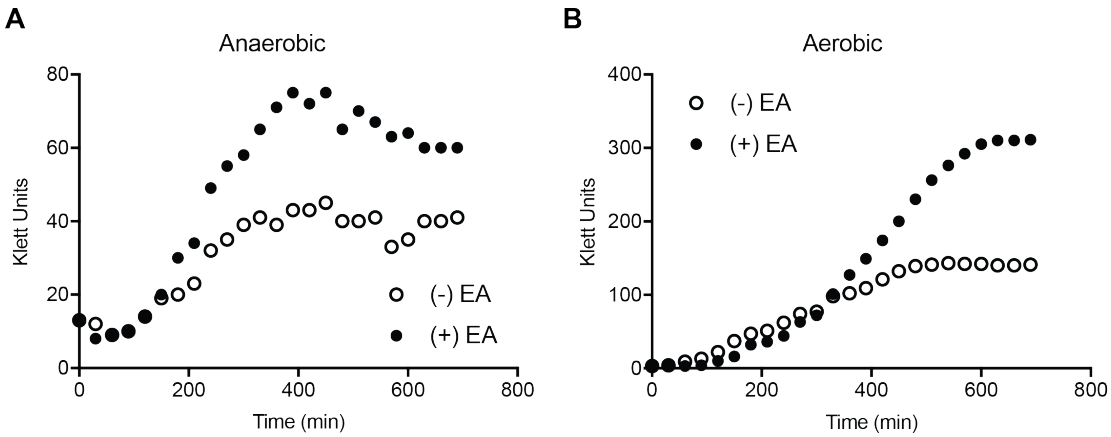

Figure 4-8. WT *E. faecalis* supports growth in the presence of EA aerobically. (A) OG1RF exhibits a moderate growth enhancement in the presence of EA anaerobically. (B) OG1RF exhibits enhanced growth in the presence of EA aerobically. (A-B) OG1RF was grown in CDM supplemented 0.2% ribose and 25 μM AdoCbl in the presence or absence of 33 mM EA.

the presence of EA the growth rate disparity diminished. OG1RF demonstrated a growth rate of 115 ± 4 minutes/doubling (Figure 4-8B, empty circles, Table 2), while EFKK6 produced a growth rate of 141 ± 3 minutes/doubling (Table 2). It was interesting that moderately faster growth rates were observed for EFKK6 under all four conditions relative to the true WT-strain, OG1RF. However, while we saw a moderate difference in growth rate, OG1RF still grew to a higher final cellular yield (Figure 4-8B). Therefore, we speculate that the *gfp* fusion protein exhibited a moderate biological effect on *E. faecalis* growth. To examine microcompartment formation, OG1RF was subjected to analysis by TEM for cells cultured in aerobic conditions and in the presence of EA. Under these conditions, the TEM micrographs revealed the presence of microcompartments. Furthermore, the cells appeared to be healthy (Figure 4-9). However, comparison of the number of Eut microcompartments that formed in the presence of EA for OG1RF versus EFKK6 revealed slightly more Eut microcompartments for EFKK6 (Table 4). One possible interpretation for this result is that the increased microcompartment formation may be due to an overall increase in shell proteins for the EFKK6 strain.

4.3.4 Decoupling EutX from the *eut* pathway enhances growth and increases Eut microcompartment formation

To further explore the utilization of EA under aerobic conditions, we investigated the role of EutX in microcompartment formation. A EutX deletion strain was inoculated into CDM supplemented with 0.2% ribose and 25 uM AdoCbl and grown in the absence or presence of EA and oxygen. Growth of the bacterial cultures were monitored every 30 minutes using a klett colorimeter.

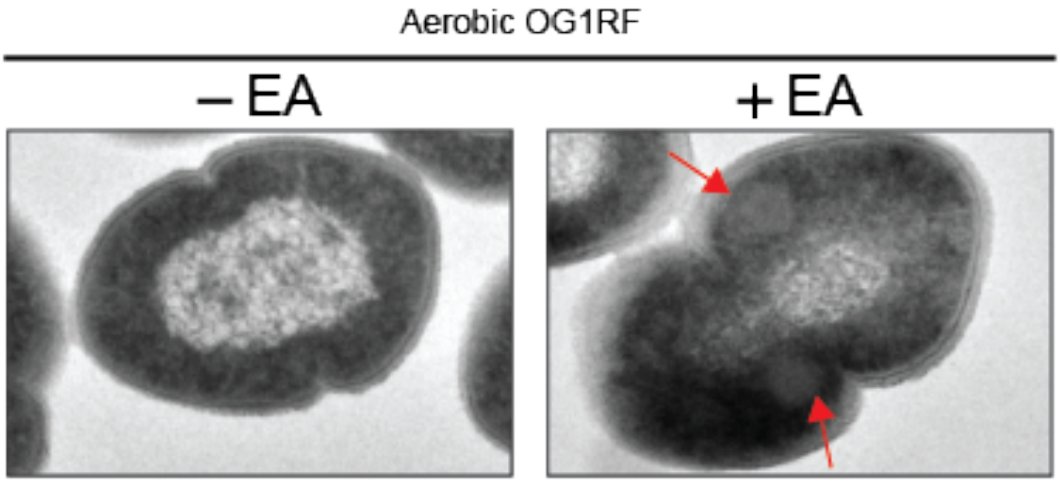

Figure 4-9. TEM images of OG1RF. OG1RF are grown in CDM supplemented 0.2% ribose and 25 μM AdoCbl in the presence of absence of 33 mM EA. Cells were collected at early stationary and washed in 1x PBS, and resuspended in 3% glutaraldehyde. TEM images collected by Dr. Karan Kaval.

The EutX deletion strain exhibited a growth rate of 109 ± 34 minutes per doubling (Figure 4-10A, filled circles, and Table 2) under anaerobic conditions in the absence of EA. Interestingly, the growth rate slowed in the presence of EA to 130 ± 30 minutes per doubling (Figure 4-10A, empty circles, and Table 2). However, while the Δ*eutX* strain exhibited a decreased growth rate, it also demonstrated a nearly two-fold increase in the total number of generations in the presence of EA relative to the absence (Table 2).

We then asked whether the presence of oxygen impacted EA utilization and Eut microcompartment formation in the Δ*eutX* strain. Under aerobic conditions the Δ*eutX* strain exhibited enhanced growth rate in the absence of EA at 94 ± 4 minutes per doubling (Figure 4-10B, filled circles, and Table 2), relative to anaerobic growth; however the addition of EA did not significantly change the growth rate (Figure 4-10B, empty circles, and Table 2). While the aerobic growth rates did not appear to change much regardless of the presence or absence of EA, we did observe an overall increase in cellular yield. The absence of EA only doubled ~2.6 times, while the presence of EA increased the number of doublings to nearly 4.

The Δ*eutX* strain was then subjected to analysis by TEM to investigate whether the enhanced growth in the presence of EA under aerobic conditions was

Table 4. Quantification of microcompartments in EFKK6, OG1RF, and *ΔeutX*

	EFKK6		OG1RF		Δ*eutX*	
	Avg.	SD	Avg.	SD	Avg.	SD
(+) EA	1.7	0.1	0.9	0.1	3.2	0.3

SD of at least 30 cells

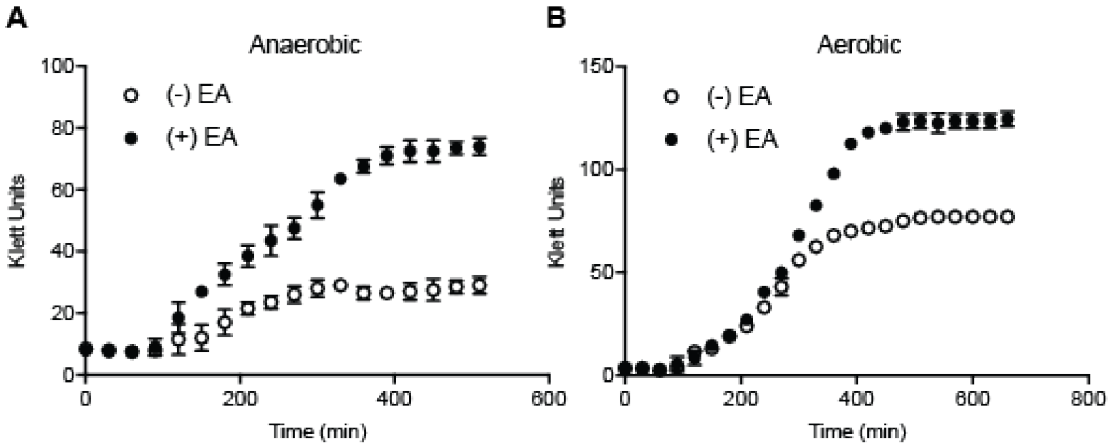

Figure 4-10. EutX deletion supports enhanced growth in EA. (A) *eutX* exhibits a moderate growth enhancement in the presence of EA anaerobically. (B) *eutX* exhibits enhanced growth in the presence of EA aerobically. (A-B) *eutX* was grown in CDM supplemented 0.2% ribose and 25 μM AdoCbl in the presence or absence of 33 mM EA.

indeed due to increased microcompartment formation. Cells were collected from early stationary phase and processed them for TEM analysis. This analysis revealed that no microcompartments were formed when cells were grown in the absence of EA. In contrast, there was an overall increase in microcompartments when the strain was cultured in the presence of EA (Figure 4-11, table 4). Quantification of the Δ*eutX* strain revealed approximately 3.2 ± 0.3 microcompartments/cell, nearly double as the EFKK6 strain. The increased number of microcompartments per cell and the enhanced overall growth in the presence of EA supports a hypothesis that in the absence of EutX, EutV can fully promote *eut* gene expression. Furthermore, when the Δ*eutX* strain and OG1RF are grown in minimal media containing EA or EA and AdoCbl and visualized by TEM, microcompartments are observed for the Δ*eutX* strain when cultured in only EA, unlike the wild type strain (Figure 4-12).

4.4 Conclusion

Previous work (111, 112) suggested a role for oxygen in EA catabolism, specifically, that *E. faecalis* required an anaerobic environment to utilize EA. We hypothesized originally that there are at least three possible explanations for the dependency on oxygen. (1) It is possible that oxygen is incorporated as a signal for genetic regulation of the *eut* gene cluster. (2) Alternatively, a reduced level of oxygen in the microcompartments could improve enzymatic reactions incorporated within them. (3) It is also possible that an overall dependency on oxygen is due to a general requirement for anoxia during *de novo* biosynthesis of AdoCbl. In data presented in

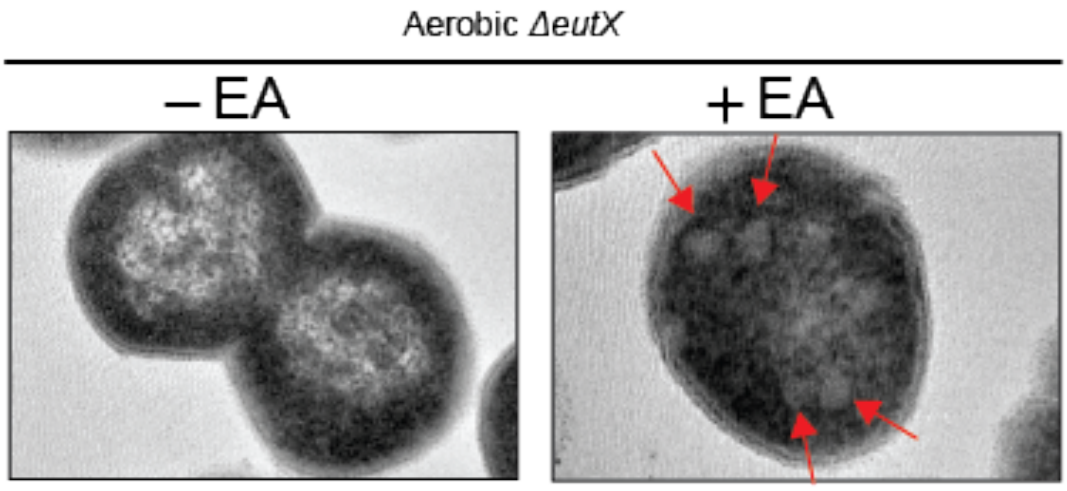

Figure 4-11. TEM images *eutX* mutant. *eutX* are grown in CDM supplemented 0.2% ribose and 25 μM AdoCbl in the presence of absence of 33 mM EA. Cells were collected at early stationary and washed in 1x PBS, and resuspended in 3% glutaraldehyde. TEM images collected by Dr. Karan Kaval.

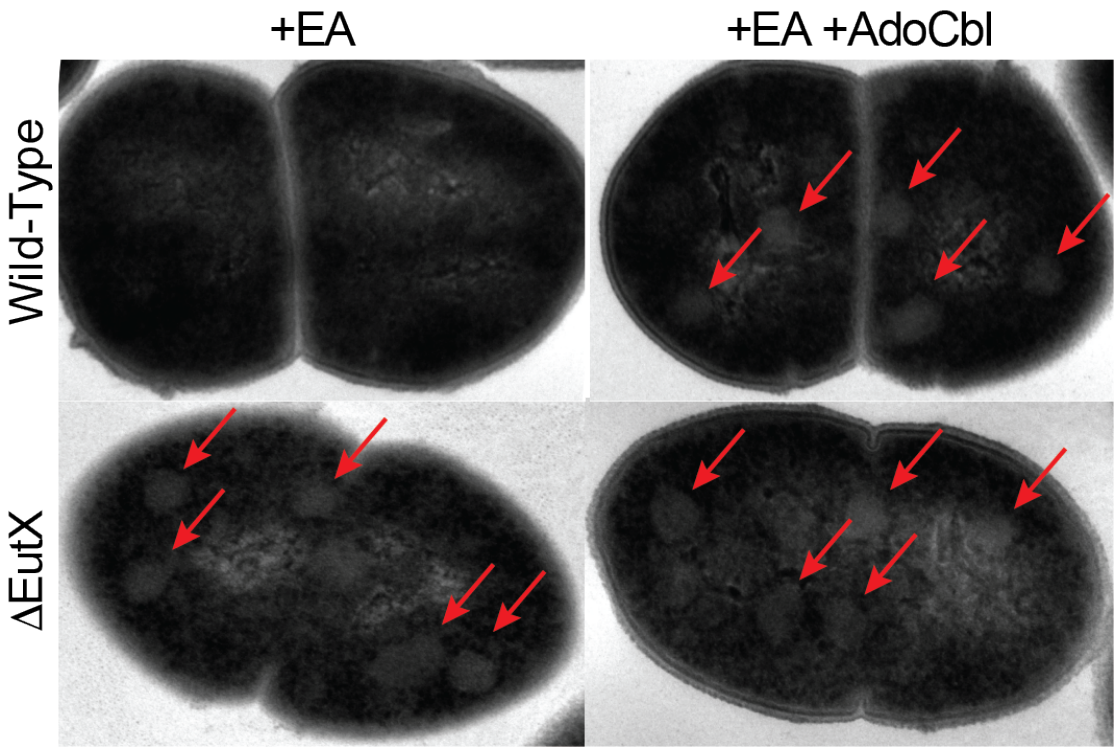

Figure 4-12. EA alone induces microcompartment formation in a *eutX* mutant. Wild type and *eutX* strains were grown in minimal media containing EA or EA and AdoCbl and visualized by TEM. Microcompartments, indicated by the red arrows were visible in the strain lacking EutX when grown in only EA, unlike the wild type strain. TEM images collected by Dr. Sruti Debroy.

this chapter, we can conclude that oxygen does not play a role in genetic regulation of the *eut* gene cluster.

Our data also suggests that oxygen is not inhibitory to microcompartment function in aerobic conditions. Instead, our data supports the hypothesis that anaerobic conditions may be required due to an effect of oxygen on AdoCbl synthesis. Therefore, our data demonstrate that oxygen is not a genetic regulatory feature of EA utilization in *E. faecalis*. Furthermore, we have established a chemically defined media (CDM) supplemented with ribose, AdoCbl, and EA, to study EA utilization in *E. faecalis*, and potentially other Gram-positive organisms.

In some species, such as *Salmonella enterica*, the organism cannot utilize EA aerobically due to the restriction of *de novo* AdoCbl synthesis occurring only under anaerobic conditions. This anaerobic requirement is due to the oxygen sensitivity of a precorrin intermediate (161, 162). The anaerobic requirement for EA utilization also ties into the hypothesis that the Eut microcompartment (78, 88) may facilitate an anaerobic environment for such sensitive reactions to take place; however, this has yet to be investigated in depth. However, *E. faecalis* does not synthesize AdoCbl *de novo*, but instead scavenges cobalamin and produces AdoCbl from a corrinoid cobalamin adenosyltransferase encoded by *eutT* (72), eliminating the oxygen sensitivity. This suggests that oxygen may not inhibit EA utilization in *E. faecalis*.

It is interesting that yeast extract appears to inhibit EA utilization. Yeast extract is a complex and undefined source of carbon, therefore, we hypothesized that carbon regulation may factor into EA utilization. We sought to test this hypothesis by growing WT, *E. faecalis* OG1RF in microaerophilic conditions in CDM

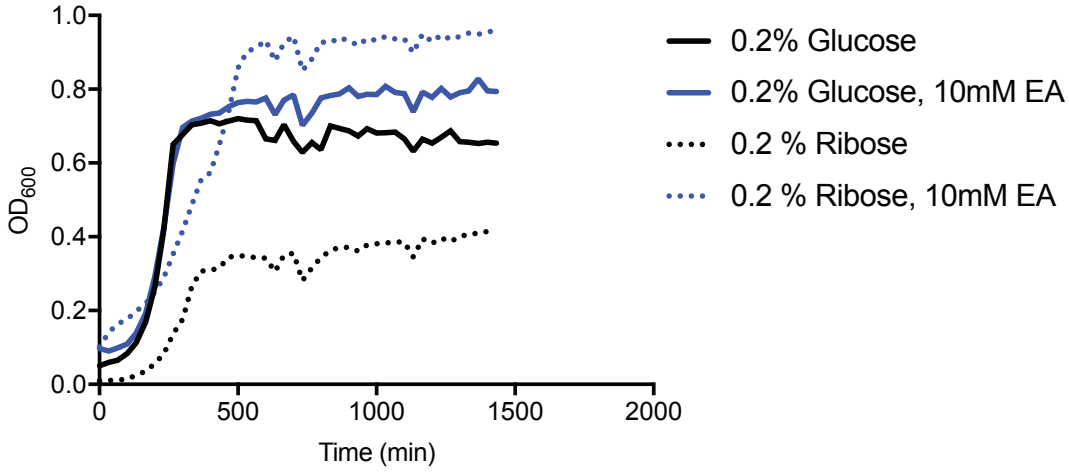

Figure 4-13. Glucose inhibites EA utilization in *E. faecalis*. WT *E. faecalis* exhibits enhanced growth in the presence of glucose regardless of the presence of EA. However, 0.2% ribose alone is not sufficient to support growth, but adding 10 mM EA enhances growth to that of the glucose containing cultures. OG1RF was grown in CDM supplemented 0.2% ribose or 0.2% glucose and 25 μM AdoCbl in the presence or absence of 10 mM EA.

supplemented with 25 uM AdoCbl, 0.2% glucose (PTS sugar) or 0.2% ribose (non-PTS sugar) in the presence or absence of 10 mM EA. OG1RF grew well in 0.2% glucose (Figure 4-13, black solid line), as expected, but exhibited no increase in growth when supplemented with 10 mM EA (Figure 4-13, blue solid line). From this observation, we hypothesized that glucose may participate in genetic regulation of the *eut* gene cluster. Furthermore, when OG1RF was grown in CDM supplemented with a non-PTS sugar, ribose, very little growth was observed (Figure 4-13, black dashed line); however, growth was greatly enhanced, to that of the glucose culture, when 10 mM EA was also present (Figure 4-13, blue dashed line). Furthermore, we previously noticed that a *eutS* promoter-*lacZ* translational fusion containing a deletion of the terminator showed high expression in the absence of EA (113). We hypothesize that this observed enhanced *lacZ* activity might be suppressed in the presence of glucose as well as the presence of EutX, suggesting a link to carbon catabolite repression (CCR).

Carbon catabolite repression generally refers to the mechanism used by bacteria to ensure that the most preferred carbon source, in most cases glucose, is catabolized first (163). In the presence of glucose, transcription of the utilization genes for alternative carbon sources in Gram-positive bacteria is repressed by the catabolite control protein CcpA, which binds to the catabolite-responsive element (*cre*) in the promoter region of the regulated operon (163). However, transcription can be autoregulated by a non-preferred carbon source when glucose is not present through a mechanism that utilizes CcpA, as shown for the sucrose utilization genes in *Staphylococcus xylosus* (163, 164). Although examples of autoregulation involve

only sugars or sugar-derivative sources, we speculate that it is possible that a gene cluster encoding for utilization of a non-sugar carbon source, such as EA, may also be subject to regulation by CcpA.

Since we hypothesize that the *eut* gene cluster may be regulated by carbon catabolite repression, it should contain a recognizable *cre* site for binding of CcpA. By searching the *E. faecalis* OG1RF *eut* gene cluster for the *cre* consensus sequence, WTGNAANCGNWNNCW, found for *B. subtilis* (165, 166), we found a putative *cre*-site just upstream of the *eutSVW* operon. Additionally, the Perego lab identified a putative -10 (TATAAT) and -35 (TTGACA) consensus sequence separated by 17 bp located in the *eutG-eutS* intergenic region (111). It is precisely within this 17 bp region that the putative *cre*-site was identified, just upstream of the P7P8 two-hairpin motif that serves to control the *eutSVW* operon. Furthermore, when searching the *L. monocytogenes eut* gene cluster for the *cre* consensus sequence, we also find a *cre*-site just upstream of a two-hairpin motif that serves to control *L. monocytogenes eutVW* operon. Therefore, these data strongly suggest that CcpA may serve to repress *eutVW* expression in the presence of glucose, but when glucose is not present, *eutVW* expression is derepressed, and the *eut* gene cluster can be upregulated.

Works described in this chapter show that *E. faecalis* not only utilizes EA aerobically in addition to anaerobically, but that it incorporates formation of Eut microcompartments to do so. Furthermore, we have supported our previous hypothesis (132) that EutX serves to sequester EutV. When we decoupled EutX from the pathway we saw an enhanced growth in EA regardless of oxygenation conditions. TEMs of this EutX deletion strain confirmed Eut microcompartments were formed,

but also indicated that more than double the microcompartments formed in this strain. In the absence of EutX, specifically P3P4, EutV can antiterminate at the ANTAR substrates and increase *eut* gene cluster expression. Furthermore, while our data in this chapter eliminated a genetic role for oxygen in *eut* gene expression, we speculate the *eut* gene cluster is still subject to additional genetic regulation. Specifically, we speculate that CCR may provide an additional regulatory layer for the *eut* pathway.

Chapter 5: Summary of primary findings and future directions

5.1 Summary of Primary Research findings

A primary motivation for this dissertation research was to investigate how *E. faecalis* incorporates AdoCbl into genetic regulation of the *eut* locus. Prior to this research, several lines of evidence pointed to an AdoCbl-sensing riboswitch located in the intergenic region of *eutT-eutG*, and suggested it may play a role in *eut* regulation. However, these reports differed in how the riboswitch might be acting. The Perego group postulated that it behaves as a canonical AdoCbl riboswitch by serving as a 5' UTR of *eutG* or a 3' UTR of *eutT*, and that binding of AdoCbl would promote termination (111, 112). Our lab previously showed evidence that was interpreted as showing the AdoCbl riboswitch functions as a 5' UTR of *eutG*, wherein AdoCbl activated antitermination (113). However, this work has resolved this question and indicated that the AdoCbl riboswitch serves to sense when cellular AdoCbl levels are low, and promotes transcription of a full-length, 300-nucleotide sRNA, EutX. EutX contains a two-hairpin motif, P3P4, at its 3' terminus that serves to sequester EutV, which in turn prevents antitermination, thereby decreasing expression of the *eut* gene cluster. Conversely, when cellular AdoCbl levels are high, the riboswitch adopts an alternative conformation and promotes synthesis of a shorter, 150-mer sRNA. P3P4 is not present in the shortened sRNA, which is therefore unable to sequester EutV (132). *In vivo* evidence of a strain containing a P3P4 deletion confirms that this RNA acts in *trans*, since its deletion allows for *lacZ* expression in a P5P6-*lacZ* fusion, while a strain that contains P3P4 either on an extra-chromosomal

plasmid or the chromosome is still sufficient to sequester EutV, and thus prevent expression of the P5P6-*lacZ* fusion. Taken together, these results indicate that for the first time a riboswitch serves to control a protein sequestration site, and, by doing so, exerts regulatory influence over an entire signal transduction pathway.

This dissertation also explores the key determinants of ANTAR proteins and their substrate binding. Previously, the ANTAR domain of EutV alone was shown to be sufficient for binding P1P2 (113). Additionally, the C-G closing base pair, A at position 1 and G at position 4, have been shown to be the key determinants of the ANTAR substrate (114). However, despite our best efforts to search for them, ANTAR substrates have not been identified for a subset of organisms that encode for proteins with ANTAR domains. This begs the question: are there other specificity determinants that occur in nucleotides that are less conserved at the primary sequence level? Furthermore, based on this hypothesis, it is likely that there will be specificity determinants in the ANTAR domain that accommodate modified substrates. We have identified one amino acid in the EutV ANTAR protein, R142, which is important for antitermination activity. R142 in the EutV ANTAR domain can be mapped to R340 in the NasR ANTAR domain. Mutating R142 to an alanine or a lysine lead to a loss of antitermination activity *in vivo* for EutV, thus suggesting an important role in antitermination activity, perhaps involved in RNA-contact. Similarly, upon purification of R340A and R340K NasR mutants, we saw loss of RNA-binding activity, further confirming an important role for this amino acid in RNA-binding. Further experiments can determine specifically which amino acids are responsible for

interacting directly with the two-hairpin motif. It would be interesting to discern correlations between amino acid residues and their nucleic acid binding partners.

Finally, our work in this document attempted to unravel the biological requirements for ethanolamine utilization in *E. faecalis*. Previous work by the Perego group had suggested that ethanolamine could only be used as a carbon/nitrogen source under anaerobic conditions. However, our data contradict these findings to show that *E. faecalis* is able to catabolize ethanolamine under both aerobic and anaerobic conditions, and that Eut microcompartments are fully formed after addition of ethanolamine and AdoCbl.

Interestingly, preliminary data from growth assays in the presence of glucose and ethanolamine versus glucose alone showed no enhanced growth with glucose and ethanolamine. This observation suggests that glucose, a PTS sugar, might be inhibitory to ethanolamine utilization. In *Firmicutes*, in additions of low glucose carbon catabolite repression (CCR) is mediated by binding of CcpA at *cre*-sites. We identified a putative *cre*-site just upstream of the *eutSVW* operon. We postulate that once glucose reserves are exhausted, *eutSVW* will be derepressed, and only then will *E. faecalis* catabolize ethanolamine.

5.2 Sensing AdoCbl and EA, and eut Gene Cluster regulation

Ethanolamine utilization has been best studied in *S. typhimurium*, where EutR, a DNA-binding transcription factor, functions as a positive regulator of the *eut* operon. Those organisms that lack EutR, such as Firmicutes, instead have a two-component system EutV/W, wherein EutW senses EA (Figure 5-1A and C).

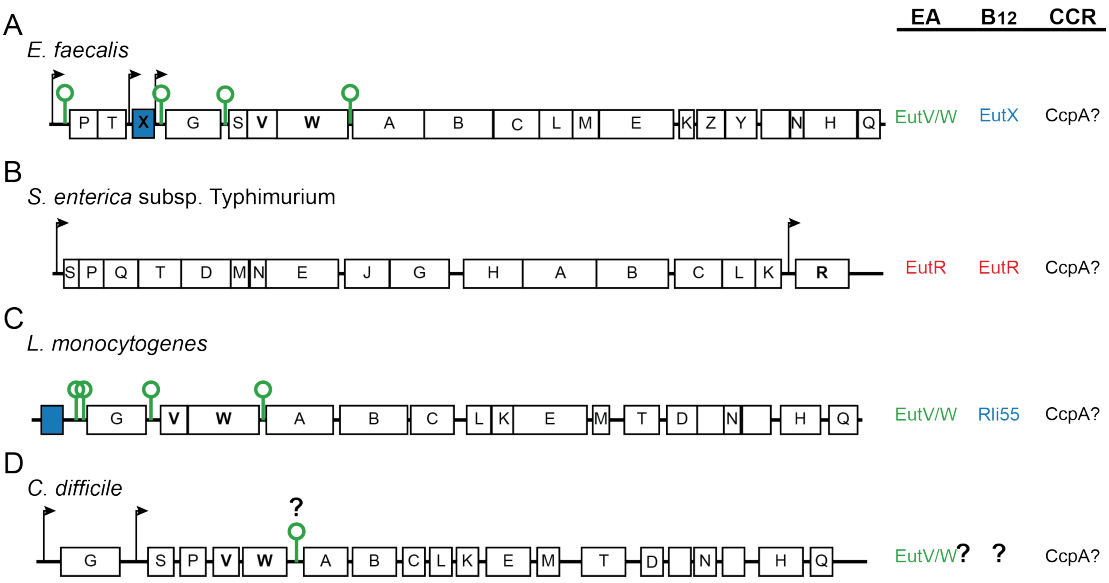

Figure 5-1. Strategies of *eut* gene regulationin multiple organisms. (A) *E. faecalis eut* gene cluster. EutW senses EA, while the AdoCbl riboswitch in EutX senses AdoCbl, possible carbon catabolite repression by CcpA. (B) *S. entericia* subsp. Typhimurium *eut* operon. The DNA-binding transcription factor, EutR, is responsible for sensing both EA and AdoCbl. (C) *L. monocytogenes eut* gene cluster. Similar to *E. faecalis*, *L. monocytogenes* EutW senses EA, while the AdoCbl riboswitch in Rli55 senses AdoCbl, possible carbon catabolite repression by CcpA. ((D) *C. difficile eut* gene cluster. The EutW presumably senses EA, but how this organism senses AdoCbl has yet to be identified.

However, the discovery of EutX adds to these regulatory features. It conceptually extends prior sRNAs to demonstrate a mechanism for how protein sequestration can be placed under signal responsive regulatory control by a riboswitch, the first of its kind (Figure 5-1A) (132). This mechanism is supported by the detection of putative EutX-like orthologs in *L. monocytogenes* (Figure 5-1C), and *Streptococcus sanguinis*. However, in *S. sanguinis*, the riboswitch and ANTAR substrate are inverted relative to the adjacent *eutG* gene and its corresponding ANTAR substrate (131). Collectively, our results indicate how organisms that lack EutR sense and respond to AdoCbl to regulate *eut* gene expression. Furthermore, our results expand the mechanisms by which riboswitches can regulate gene expression. Conversely, some organisms, such as *Clostridium difficile*, appear to encode for the two-component system EutV/W and possibly ANTAR substrates, but do not seem to have an AdoCbl sensing riboswitch (Figure 5-1D). How these organisms sense and respond to cellular concentrations of AdoCbl is not yet known. Future studies of how these organisms sense and respond to AdoCbl may reveal additional novel or modified mechanisms and add again to the repertoire of *eut* operon regulation strategies.

5.3 Biological Relevance of Ethanolamine Utilization in E. faecalis

Catabolism of ethanolamine in *E. faecalis* was once thought to occur only in an anaerobic environment. This hypothesis was based on the anaerobic requirement of *S. typhimuriums* for *de novo* synthesis of AdoCbl, a required cofactor for ethanolamine breakdown. However, our results show that a complex media containing yeast extract, which we have shown inhibits the proper use of ethanolamine, thwarted these original conclusions. We have shown that oxygen is not

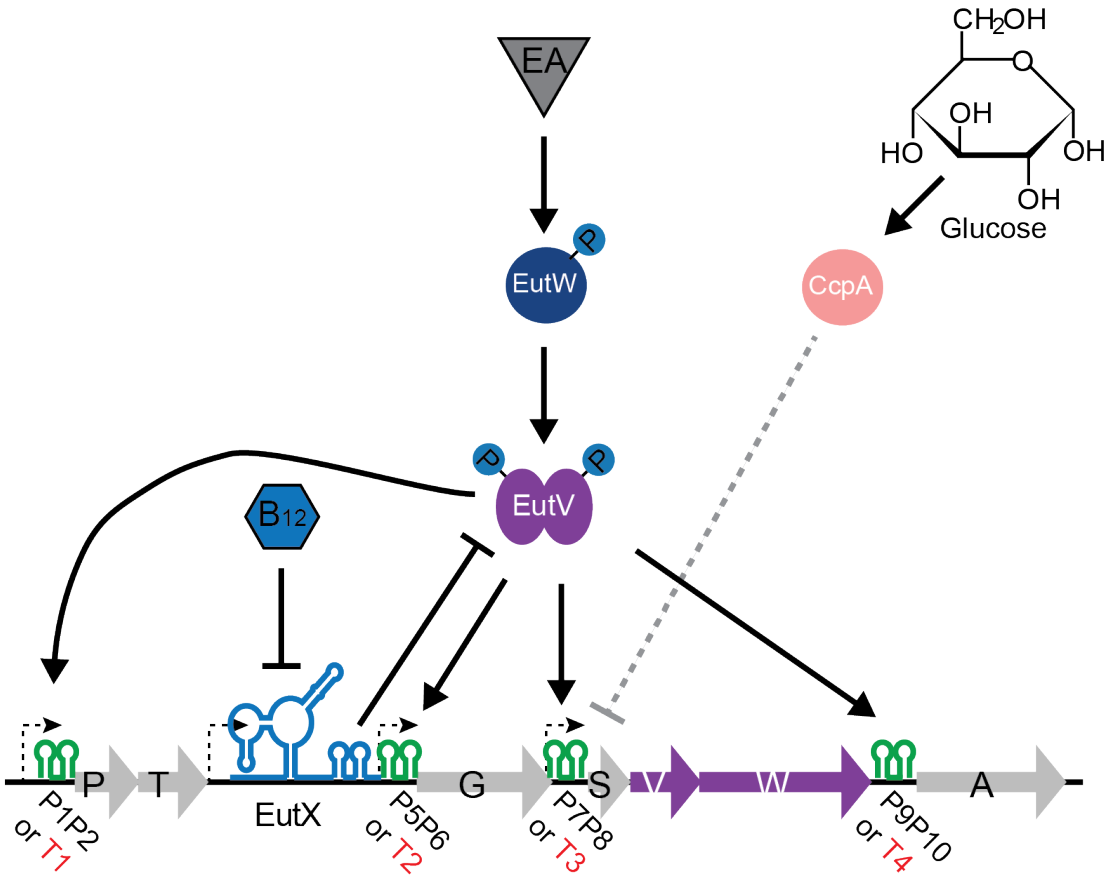

Figure 5-2. Final model for *E. faecalis eut* gene regulation. The presence of glucose activates CcpA to bind at a putative *cre*-site upstream of the *eutSVW* operon. In the absence of glucose, and presence of ethanolamine EutW senses ethanolamine, autophosphorylates, and transfers its phosphoryl group to EutV promoting dimerization and activation. EutV then promotes antiterminates at the indicated two-hairpin motifs (green). AdoCbl (B12) is sensed by the AdoCbl riboswitch found within EutX. P3P4 two-hairpin motif of EutX sequesters EutV, preventing antitermination, thus down regulating the *eut* gene cluster.

a feature of *eut* gene regulation, and that *E. faecalis* can utilize ethanolamine under both aerobic and anaerobic conditions. Furthermore, we observed that the presence of glucose in the media also inhibits the use of ethanolamine under both aerobic and anaerobic conditions. This suggests a possible role for carbon catabolite repression, and therefore CcpA in regulating the *eut* gene cluster (Figure 5-2). Strengthening our argument is a putative *cre*-site just upstream of the P7P8 two-hairpin motif regulating the *eutSVW* operon. However, this *cre*-site is untested, and therefore experiments are required to determine if the *eut* gene cluster is indeed regulated by CcpA.

It is interesting that the deletion of EutX allowed for increased Eut MCP formation, and that AdoCbl was no longer required. This observation supports our model that EutX serves to sequester EutV, preventing downstream *eut* gene expression, thus in the absence of EutX we see increased MCPs. We hypothesize that transcript analyses of the *eut* gene cluster in the absence of EutX will show increased expression.

5.4 Conclusion

Works described in this dissertation demonstrate that the *eut* gene cluster in *E. faecalis* contains several layers of regulation as can be seen in Figure 5-2. We believe that these various regulatory elements may extend beyond *E. faecalis*. For example, *L. monocytogenes* also encodes for a similar sRNA, Rli55, which behaves similar to EutX (131). Additionally, a putative *cre*-site can be found just upstream of the *L. monocytogenes* EutVW operon, suggesting its *eut* operon is also regulated by carbon catabolite repression. This dissertation has served to expand the regulatory roles of riboswitches. We have begun to address the requirements for ANTAR domain

antitermination, both *in vitro* (NasR) and *in vivo* (EutV) in a heterologous host. Finally, we have changed the way the field studies *E. faecalis* EA utilization, and potentially found that the *eut* gene cluster is regulated by carbon catabolite repression.

Chapter 6: Materials and Methods

6.1 Chemicals and oligonucleotides

All chemicals and enzymes, unless otherwise noted, were obtained from Sigma-Aldrich and New England Biolabs, respectively. DNA oligonucleotides were purchased either from Integrated DNA Technologies, Inc. or Sigma-Aldrich. Exact nucleotide sequences and brief description of DNA oligonucleotides used in these studies can be found in Table 4.

6.2 Bacterial strains and media

The *E. faecalis* strains used in the an/aerobic studies, and EutX discovery (Chapters 2 and 4) were all derivatives from strain OG1RF. All *E. faecalis* strains were obtained from Dr. Danielle Garsin's lab at the University of Texas Health Science Center at Houston. The *B. subtilis* strains in the heterologous expression and ANTAR studies (Chapter 3) were all derivatives from strain 168. All *B. subtilis* strains were obtained from Bacillus Genetic Stock Center (BGSC, Ohio).

6.2.1 Growth conditions

All *E. faecalis* strains were typically cultured at 37°C in either Brain Heart Infusion (BHI, Difco) broth or plates supplemented with 1.5% Bacto agar (Difco), Chemically Define Media (CDM, Alphabiosciences), recipe in the Table 1 supplemented with 0.2% ribose, or 1x M9HY (20 mM HEPES pH 7.0, 2g/L yeast extract, 1 mM M9 salts, 1 mM $MgSO_4$, 0.2 mM $CaCl_2$, 1 g/mL thiamin, 0.2 g/mL biotin, 0.2 g/mL pantothenic acid, 20 μg/ml nicotinic acid, 20 μg/ml riboflavin, 2

μg/ml folic acid, 0.2 g/mL arginine, 0.2 g/mL glutamate, 0.2 g/mL glycine, 0.2 g/mL histidine, 0.2 g/mL isoleucine, 0.2 g/mL leucine, 0.2 g/mL methionine, 0.2 g/mL tryptophan, and 0.2 g/mL valine).

All *B. subtilis* strains were typically cultured at 37°C in either 2xYT (16 g tryptone, 10 g yeast extract, 5 g NaCl per liter), or on Tryptone Blood Agar Base (TBAB, Difco) plates.

All *E. coli* strains were cultured at 37°C in Luria Bertani (LB, Difco) broth or agar (10 g tryptone, 5 g yeast extract, 5 g NaCl per liter)

6.2.2 Antibiotic concentrations

When appropriate, antibiotics were used at the following concentrations (μg/ml) for *E. coli*: ampicillin, 50; carbenicillin, 100. Antibiotics were used at the following concentrations (μg/ml) for *B. subtilis*: spectinomycin, 100; erythromycin, 1 and lincomycin, 25; chloramphenicol, 5; neomycin, 5. Antibiotics were added at the following concentrations (μg/ml) for *E. faecalis*: erythromycin, 50; and rifampicin, 100.

6.2.3 Amino acid concentrations for auxotrophs

B. subtilis 168 is an tryptophan auxotroph, and 50 μg/ml tryptophan was added for all experiments of this strain and its derivatives. Any strain with a disruption to the *thrC* gene creates a threonine auxotroph, thus 50 μg/ml threonine was added for all experiments of these strains and its derivatives.

6.2.4 Inducer concentrations

IPTG, bacitracin, EA, and adenosylcobalamin were used to induce specifc

gene expression and protein expression/activity with final concentrations of 1 mM, 30 μg/ml, 33 mM (*E. faecalis*) and 5 mM (*B. subtilis*), and 25 μM, respectively.

6.3 Strain construction

All bacterial strains used in this dissertation are listed in Table 7.

6.3.1 Construction of *amyE*:: P*const* P1P2T-YFP

Plasmid pDG1662 contains regions of the *amyE* gene and was thereby used to integrate a gBlock containing P_{const} P1P2T-YFP with flanking arms of homology to *amyE* within the *amyE* gene via double recombination. Oligonucleotides and gBlocks used to make this strain are listed in Table 4 and 5.

6.3.2 Construction of *thrC*:: P_{liaI} EutVW

Plasmid pDG1664 containg regions of the *thrC* gene and was thereby used to integrate a gBlock containing P_{liaI} EutVW with flanking arms of homology to *thrC* within the *thrC* gene via double recombination. Any mutations to EutV were done so using oligonucleotides specific for Q5 site-directed mutagenesis (New England Biolabs). Oligonucleotides and gBlocks used to make these strains are listed in Table 4 and 5.

6.3.3 Construction of inducible EutX

To the EutX sequence a semi-synthetic terminator (SST) was added. The semi-synthetic terminator was derived from the original termination poly-U tract found using RACE, and making a stronger terminator sequence. This semi-synthetic terminator was added to EutX via PCR amplification using oligonucleotides with the

added sequence. Next, the EutX with SST was cloned into pDG148. Oligonucleotides used to make these strains are listed in Table 4

6.3.4 Construction of inducible GlmS-P3P4

To the GlmS sequence (167) the linker region and P3P4 of EutX was added followed by the same SST. This was purchased from IDT as a gBlock with flanking arms of homology to the pDG148 vector. Next, this gBlock was cloned into pDG148. A derivative strain, GlmS (M9 mutant)-P3P4 was generated by Q5 site-directed mutagenesis (New England Biolabs). Oligonucleotides and gBlocks used to make these strains are listed in Table 4 and 5.

6.3.5 Construction of *thrC*:: P_{liaI} NasR

A gBlock containing NasR from *K. oxytoca* codon optimized for *B. subtilis* was ordered from IDT. Plasmid pDG1664 containg regions of the *thrC* gene and was thereby used to integrate a gBlock containing P_{liaI} NasR with flanking arms of homology to *thrC* within the *thrC* gene via double recombination. Any mutations to NasR were done so using oligonucleotides specific for Q5 site-directed mutagenesis (New England Biolabs). Oligonucleotides and gBlocks used to make these strains are listed in Table 4 and 5.

6.3.6 Construction of NasR expression vector

Codon optimized NasR was amplified in-frame with a His_{10}-MBP tag from pVL847 and Gibson assembled (168) into NdeI and BamHI digested pVL847.

6.3.7 Genetic transformation of *E. coli*

All Gibson assemblies, Q5 site-directed mutageneses, and subcloning reactions were transformed as follows: 5 µL of the cloning reaction was transformed into 50 µL chemically competent XL10Gold *E. coli*. The DNA was incubated with the cells for 30 minutes on ice. The mixture was heat treated at 42°C for 30 seconds followed by recovery on ice for 5 minutes. 250 µL of rich media was added to the cells for an hour outgrowth at 37°C. 100 µL of the reaction was plated on an LB plate with the appropriate antibiotic.

6.3.8 Genetic transformation of *B. subtilis*

All strains were transformed with 1-10 µg plasmid DNA using the one-step method. Briefly, the recipient strains were inoculated from a single colony into transformation media (12.5g of K_2HPO_4, 3 g of KH_2PO_4, 0.5g trisodium citrate, 0.1g $MgSO_4$, 1g Na_2SO_4, 50µM $FeCl_3$ 2µM $MnSO_4$, 0.4% glucose and 0.2% glutamate in 500 mL of water) supplemented when appropriate with 50 µg/mL tryptophan and or threonine, and placed in a 37˚C incubator standing over night. The following day they were placed in a shaking incubator and the OD_{600} was monitored until the cultures reached exponential phase (between 0.4 and 0.8) and were then inoculated with plasmid DNA. The cultures were subsequently replaced in the incubator shaking for 45 minutes. Next, 2xYT (containing 0.1 µg/mL chloramphenicol or erythromycin when the donor strain was resistant to either) was added to the culture for outgrowth and recovery, and replaced in the incubator shaking for an additional 45 minutes. Lastly, the cultures were plated onto a TBAB plate with the appropriate antibiotic(s).

6.4 RNA extraction

Cells were harvested by centrifugation at 5,700 rpm for 10 minutes at which point pellets were frozen at -80°C until ready for RNA extraction. To extract total RNA, the cell pellet was briefly thawed and resuspended in 750 µL LETS buffer (0.1 M LiCl, 10 mM EDTA pH 8.0, 10 mM Tris-HCl pH 7.4, and 1% SDS) and disrupted by continuous vortexing with 400 µL glass beads for four minutes followed by incubation at 55°C for five minutes. The suspension was centrifuged for ten minutes at 15,000 rpm and the supernatant collected and mixed with 1 mL TRI reagent (Ambion AM9738) with incubation at room temperature for five minutes. 200 µL chloroform was then added and the suspension was mixed vigorously for 15 seconds followed by a two minute incubation at room temperature. The samples were centrifuged at 15,000 rpm for 15 minutes and the top 60% (approximately 600 µL) of the phase-separated mixture was collected and precipitated with 1 mL isopropanol for 10 minutes at room temperature. Precipitated RNA was washed with 200 µL of 70% ethanol and resuspended in 20 µL purified water. If the RNA samples were to be used for qRT-PCR, the samples were treated with DNase I from Promega per manufacturer's instructions.

6.5 Northern blot analyses

Total RNA samples (10-20 µg) were heated at 65°C for 10 min in 1x gel loading buffer (45 mM Tris-borate, 4 M urea, 10% sucrose [w/v], 5 mM EDTA, 0.05% SDS, 0.025% xylene cyanol FF, 0.025% bromophenol blue) and resolved by 6% denaturing (8 M urea) polyacrylamide electrophoresis. Aliquots of an Ambion® CenturyTM RNA Marker were also resolved within neighboring lanes in order to

generate an RNA size standard curve. RNA within the polyacrylamide gels was transferred to BrightStar-Plus nylon membranes (Ambion) using a semi-dry electroblotting apparatus (Owl Scientific) in 1X NAQ buffer (4 mM MOPS, 1 mM sodium acetate, 0.1 mM EDTA) at 2 mA/cm^2 for 1 hour. The blots were UV-crosslinked at 125 mJ and hybridized overnight at 42°C in UltraHyb Oligo Buffer (Ambion) in the presence of different radiolabeled (^{32}P) antisense RNA probes or (^{32}P) DNA oligonucleotide probes. The blots were the washed twice for 15 minutes using low stringency buffer (1X SSC, 0.1% SDS, 1 mM EDTA) and then for 30 minutes using high stringency buffer (0.2X SSC, 0.01% SDS, 1 mM EDTA) Radioactive bands were visualized using ImageJ software and a FLA5000 Phosphorimager.

6.6 Rapid Amplification of cDNA Ends (RACE)

6.6.1 With 5' adenylated linkers

E. faecalis OG1RF was grown in minimal medium containing supplemental AdoCbl and EA, or EA alone. After 3.5 hours of anaerobic growth, cells were harvested and total RNA was extracted as described above. Total RNA was then DNase I-treated (Promega) and subjected to the RiboZero^{TM} Magnetic Kit (Gram-Positive Bacteria) to remove ribosomal RNA according to manufacturers' protocol. A pre-adenylated linker (NEB Universal miRNA Cloning Linker) was incubated with T4 RNA ligase KQ (NEB) for 1 hour at 25°C. The reaction was then purified using the Zymo-RNA Clean & Concentrator kit. This RNA was incubated with a DNA oligonucleotide primer complementary to the pre-adenylated linker

(ATTGATGGTGCCTACAG), and incubated for 10 minutes at 65°C. Then iScript reverse transcriptase, buffer, and 5 mM dNTPs were added and the reactions were incubated at 42°C, 50°C, 55°C, 60°C for 15 minutes each, followed by 90°C for 5 minutes. The reaction was then treated with 0.5 μg/μL RNase A for 20 minutes at 37°C and again purified by the Zymo-DNA Clean & Concentrator kit. This cDNA was used as template for PCR amplification, using oligos MG283 and MG285 (Table S2). The resulting PCR products were subcloned by TOPO TA reactions according to manufacturers' protocol (Invitrogen) and transformed into Top10 cells (Invitrogen) on LB containing 50 μg/mL ampicillin and 40 μg/mL X-Gal. White colonies were chosen and the appropriate insert sequences were analyzed by Sanger sequencing.

6.6.2 From circularized RNA transcripts

Total RNA was treated with DNase I and then incubated with TerminatorTM 5'-phosphate dependent exonuclease (Epicentre), according to manufacturers' protocol. Terminator exonuclease-treated RNA was then incubated with Tobacco Acid Pyrophosphatase (Epicentre) for full conversion of 5'-triphosphates to 5'-monophosphates. The RNA was then incubated with T4 RNA ligase (NEB) under relatively dilute conditions to promote intramolecular ligations. cDNA synthesis and PCR amplification were performed as above using oligos MG293 and MG294 (Table S2). The subsequent sub-cloning, PCR amplification and sequencing analyses were performed as above.

6.7 In vitro *transcription assays*

The double-stranded DNA template for transcription of the EutX-associated riboswitch consisted of a 308 base-pair (bp) PCR fragment. This sequence included from -135 to +1 of the *B. subtilis glyQS* promoter, followed by the EutX-associated AdoCbl riboswitch. The DNA template ended at 20 nts after the putative riboswitch intrinsic terminator, which was predicted by 3' RACE. A separate control template included a mutation (M3) of three cytosines changed to guanines, located within the L5 loop, which is known to perturb AdoCbl binding. Synchronized transcription assays were carried out as described (*16*). Template DNA (10 nM) was incubated in 1X transcription buffer (10 mM Tris pH 8.0, 20 mM NaCl, 14 mM $MgCl_2$, 14 mM β-ME, and 0.1 mM EDTA pH 8.0) with *E. coli* RNAP (Epicentre). The dinucleotide ApU (150 μM, Tri Link Biotechnologies) was also included to assist transcription initiation. ATP and GTP were added at 2.5 μM, UTP was added to 0.75 μM and [α-^{32}P] UTP (3000 Ci/mmol) was added to 0.33 μM. Transcription was initiated in the absence of CTP. The first C in the DNA template sequence was located at +30; therefore, the transcription elongation complex halted after synthesis of 29 nt under these conditions. The initiation reaction mixtures were incubated at 37°C for 10 min and were then placed on ice. Heparin (20 μg/mL, Sigma) was added to block reinitiation, and elongation was triggered by the addition of NTPs to 1 mM final. Transcription reactions were terminated with 2x urea loading buffer (8 M urea, 20% sucrose [w/v], 0.1% SDS, 0.05% bromophenol blue [w/v], 0.05% xylene cyanol FF [w/v], 0.09 M Tris-HCl pH 7.5, 0.09 M borate, and 1 mM EDTA pH 8.0) and resolved by 6% denaturing polyacrylamide gel electrophoresis and visualized and

quantified using ImageJ software and a FLA5000 phosphorimager.

6.8 Quantitative real-time RT-PCR

Total RNA from strains was isolated as described above. Approximately 5 μg RNA was then incubated in 20 μL reactions with 2 units DNase (QuantBio) per manufacturer instructions. DNA-depleted total RNA was converted to cDNA with Quanta Bio qScript cDNA Supermix. The cDNA was diluted in TE buffer and added to qPCR reactions made with Quanta Bio PerfeCTa SYBR Green FastMix. All qPCR reactions were prepared for 18 μL volumes in 96- or 384-well plates, and analyzed using a Roche Lightcycler 480 with the recommended three-step fast cycle. All oligonucleotides used for each amplicon are listed in Table ##. All experiments utilized negative RNA-only controls for DNA contamination using reference gene amplicons. C_q values and amplification efficiencies were determined using LinRegPCR (169). Relative quantification was performed relative to *gyrB*. All experiments used at minimum three biological replicates.

6.9 RNA–binding studies using electrophoretic mobility shift assays (EMSA) or differential radial capillary action of ligand assay (DRaCALA)

RNA transcripts were transcribed and radiolabeled with γ-^{32}P ATP. Similar reactions, minus γ-^{32}P ATP, were run to generate cold competitor RNA, which was added at 10-fold excess, ~3 μM, to certain controls. Radiolabeled RNA (~300 fmol) was incubated with increasing concentrations of protein in a 20 μl reaction containing 50 mM HEPES pH 7.5, 75 mM sodium chloride, 10 mM $MgCl_2$ and 2.5 ng/μl yeast tRNA for 30 minutes at 25°C. The samples were then subjected to either EMSA or

DRaCALA (130). For EMSA, 10 μL of each sample was run on a non-denaturing 6% TBE polyacrylamide gel (acrylamide: bis, 29:1). Gells were pre-run for 30 min at 15 volts, and samples were electrophoresed at 15 volts for 30 minutes to 1 hour, with 0.5X TBE running buffer, followed by drying for 45 min. For DRaCALA, ~2 μL of each sample was spotted on a nitrocellulose membrane using a pin-tool and allowed to air dry for 30 min. Gels for EMSA and nitrocellulose membranes for DRaCALA were exposed for one hour, visualized, and quantified using ImageStudioLite software and an FLA5000 phosphorimager.

6.10 Protein expression and purification

Deca-histidine-tagged NasR was cultured in 2xYT supplemented with 0.2% glucose, and expression induced at A_{260}= 0.5 with 1 mM IPTG at RT for 18 h. The cell pellet was resuspended in Resuspension buffer (50 mM HEPES pH 8.0, 150 mM NaCl, 1 mM $MgCl_2$, 2 mM β-ME, 10 mM Imidazole, 5% glycerol, 1 mM PMSF, 2 U DNase). Cells were lysed with 0.5 mg/mL lysozyme for 20 minutes on ice, then 0.1% (final) TritonX-100 was added, followed by $1/10^{th}$ volume 5 M NaCl. The resuspension was then sonicated on ice until no longer viscous. After the cell disruption and centrifugation, the supernatant was passed over 1 mL Ni-NTA resin (Qiagen), followed by 20 column-volumes (CV) wash buffer (50 mM HEPES pH 8.0, 500 mM NaCl, 1 mM $MgCl_2$, 2 mM β-ME, 20 mM Imidazole, 5% glycerol). The protein was eluted in five fractions of 2 CV elution buffer (50 mM HEPES pH 8.0, 150 mM NaCl, 1 mM $MgCl_2$, 2 mM β-ME, 500 mM Imidazole, 5% glycerol). Eluted protein was dialyzed against dialysis buffer (50 mM HEPES pH 8.0, 75 mM NaCl, 1 mM $MgCl_2$, 2 mM β-ME, 10% glycerol) for three 2 hour buffer exchanges. All steps

were performed either on ice or at 4°C. Purity of NasR was judged by 4-20% SDS/PAGE followed by coomassie-staining.

6.11 Growth curves

6.11.1 E. faecalis

Single colonies of *E. faecalis* strains were inoculated into BHI broth and incubated for 18 h at 37°C. The following day the overnight cultures were washed in CDM and inoculated to an OD_{600} of ~0.1 in fresh CDM supplemented with 0.2% ribose and 25 μM AdoCbl in the presence of absence of 33 mM EA. 2g/L of yeast extract was added as it was required. Aerobic or anaerobic cultures were performed as follows:

6.11.2 Aerobic

The cultures were then grown at 37°C shaking, and growth measured every 30 minutes by Klett colorimeter.

6.11.3 Anaerobic

The 10 mL of media was placed into Klett tubes, and sealed. The vials were placed under vacuum for 10 minutes, and then degassed with N_2 gas. The cultures were then grown at 37°C standing, and growth measured every 30 minutes by Klett colorimeter.

6.12 Fluorescence microscopy

For agarose pad imaging of fluorescence, strains were cultured as previously described, and resuspended in phosphate buffered saline (PBS), 3 μL was spotted

onto an agarose pad (1.5% agarose in PBS), transferred to sterile glass microscope slide, and placed in a glass bottom dish. An Axio Observer.Z microscope (Zeiss) was used to image strains at 100x magnification using phase contrast and fluorescence. Photographs were taken using Zen 2012 imaging software (Zeiss).

6.13 Flow cytometry

Single colonies of *B. subtilis* strains were inoculated into 2xYT broth with appropriate antibiotic for 18 h at 37°C. The following day the overnight cultures were inoculated into fresh 2xYT broth to an OD_{600} of ~0.1. The resulting culture was incubated at 37°C shaking allowed to reach an OD_{600} of ~ 0.5. Once the cultures reach exponential phase, the cultures were induced with 30 μg/mL bacitracin in the presence of absence of 5 mM EA, and allowed to incubate shaking at 37°C for an additional 45 minutes. Then 1 mL of each culture was centrifuged at 5700 rpm for 5 minutes, rich media decanted, and resuspended in 1x phosphate-buffered saline solution, PBS, (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 2 mM KH_2PO_4). 500 μL of washed culture was then added to 2 mL 1X PBS.

Cells were then analyzed on a FACSCanto II, where 30,000 events were counted per sample. All samples and conditions were compared to a strain that does not contain YFP, MG5183. The median value for YFP positive cells was reported in duplicate.

Appendices

Table 5. List of oligonucleotides used in the studies.

Primer	Purpose	Sequence- 5' to 3'
MG18	Probe for northern blotting, *B. subtilis* 5S RNA (Sharp & Bechhofer, 2003)	CGACTACCATCGGCGCTGAA
MG238	*E. faecalis* OG1RF 5S rRNA Northern blot probe	GTGTATCCTTCTCGCTATCGCCACCA
MG244	F. oligo for 3' of AdoCbl riboswitch (use for 3' RACE)	GTGCTTTTGGGAATTTTGGGCTGGTCTT
MG245	R. oligo for 5' of AdoCbl riboswitch (use for 3' RACE)	CAGTGGCGGGACCGTGCCAGA
MG260	Reverse transcription oligo for determining 3' end of *E. faecalis* +EA EutX	ATTGATGGTGCCTACAG
MG261	R. oligo for amplifying *E. faecalis* EutX cDNA with MG260	CTCACTGGCTTCCATTC
MG279	Probe for northern blotting in CORE of AdoCbl riboswitch	CCTGACTTAAAAGCAACCTTGCTTC
JRG174	Probe for northern blotting in EutX, inside AdoCbl riboswitch	CAAAATTCCCAAAAGCACTTCCTCGTGCC
JRG175	Probe for northern blotting in EutX, 3' of riboswitch	GTTAAGAAACATCCTGTTTTGCCTCG
MG280	Probe for northern blotting in between probes JRG174/175	CAAATATCTTGGGCTGGATTTTCAAC
MG281	Probe for northern blotting in downstream of P4	CAATAAAAAAAGGGAATATAGCAGTACGTC
MG282	R. oligo for amplifying EutX cDNA with MG260	GAATCTGGCACGGTCCCGCC
MG283	R. oligo for amplifying EutX cDNA with MG260	CAAGGTTGCTTTTAAGTCAGGTC
MG284	R. oligo for amplifying EutX cDNA with MG260	CGAGGCAAAACAGGATGTTTCTTAAC
MG285	Reverse transcription oligo for determining 3' end of *E. faecalis* +EA EutX	GCATGCATTGATGGTGCCTACAG
MG290	F. oligo for transcription of EutX with T7 promoter	TAATACGACTCACTATAGGGGAATTCGAGTACAACGCC
MG291	R. oligo for transcription of EutX	AAATTTCAGATAGATTAAACGAATATTTT

		ATG
MG293	R. oligo for 3' of AdoCbl riboswitch (use for 3' RACE)	CAACCTTGCTTCTTCACAGTGG
MG294	F. OLIGO FOR 5' OF B12 RIBOSWITCH (USE FOR 3' RACE)	CTTTCAATGTTGAAGGAGATTTCAGG
MG295	F. oligo for 5' of AdoCbl riboswitch (use for 3' RACE)	GTTAAGAAACATCCTGTTTTGCCTCG
	Third F. oligo for circularization	
MG298	F. oligo for P_{glyQS}	ATTGATTTATATTACGAAGAATATTC
MG299	R. oligo for 45 bp downstream of 3' end of AdoCbl riboswitch	TTTTCAACACAAGCGTTGTTAAG
MG301	F. oligo for gibson assembling 3'RACE w/ circularization into pUC18	ACCATGATTACGCCAAGCTTGCATGCGTT AAGAAACATCCTGTTTTGCCTCG
MG302	R. oligo for gibson assembling 3'RACE w/ circularization into pUC18	CGGCCAGTGAATTCGAGCTCGGTACCCAA CCTTGCTTCTTCACAGTGG
MG303	R. oligo for PCR amplification of AdoCbl riboswitch component of EutX	AAGAAACATCCTGTTTTGCCTCG
MG396	F. oligo for amplification of P1P2T into pJG019	GGCATAATGTGTGTGCACAAAGAATCAGA AACACAATGG
MG397	R. oligo for amplification of P1P2T into pJG019	TGTGACTTTCCTCCTTATAATGTGACTTT CCTCCTTG
MG398	R. oligo for amplification of P1P2T into pJG019 (alternative oligo)	TGTGACTTTCCTCCTTATTATTTATACAG TTCGTCCATACC
MG434	F. oligo for P_{liaI} (bacitracin inducible promoter)	GAAGGCCAAAAAACTGCTGCCTTCGGATC CATTGGCCAAAGCAGAAAG
MG435	R. oligo for P_{liaI} (bacitracin inducible promoter)	ACTTTCCTCCTTAAGCTTTCGTTTTCCTT GTCTTCATC
MG449	F. oligo for gibson assembling P_{liaI} into pDG1664 with pDG1664 at 5' and EutV at 3'	CCAAAAAACTGCTGCCTTCGATTGGCCAA AGCAGAAAG
MG450	F. oligo for gibson assembling P_{liaI} into pDG1664 with pDG1664 at 5' and EutV at 3'	CCATTTTCTTCACTCGTTTTCCTTGTCTT CATC
MG451	F. oligo for gibson assembling EutV into pDG1664 with P_{liaI} at 5' and EutW at 3'	AGGAAAACGAGTGAAGAAAATGGATGGAC
MG452	R. oligo for gibson assembling EutV into pDG1664 with P_{liaI} at 5' and EutW at 3'	ATCGTTTCATTTATTCATCATCCATTACA ATCAATTC

MG453	F. oligo for gibson assembling EutW into pDG1664 with EutV at 5' and pDG1664 at 3'	TGATGAATAAATGAAACGATTAGAGCAATTATG
MG454	R. oligo for gibson assembling EutW into pDG1664 with EutV at 5' and pDG1664 at 3'	CCATAACTTTAGGGTTATCGTCAATGAACAACATCACTGG
MG564	F. oligo for qPCR of EutV set1	TAGGCGCACTGGGTTATTTAG
MG565	R. oligo for qPCR of EutV set1	AGTAACAGCTGCGTTTGTTTG
MG566	F. oligo for qPCR of EutV set2	TCAAGACCAACTAGCCAGTAGTA
MG567	R. oligo for qPCR of EutV set2	CTAAATAACCCAGTGCGCCTAA
MG576	F. oligo for qPCR of EutW set1	GCAATCGGTGGTTTCCTTATTG
MG577	R. oligo for qPCR of EutW set1	GGCTGCAATCGCCATAATTC
MG578	F. oligo for qPCR of EutW set2	ACTAGTCAAGTCGACGCATTATT
MG579	R. oligo for qPCR of EutW set2	GCTCTTGTGCATTGTGATTGG
MG665	F. oligo for qPCR of EutX set2	TCCAGCCCAAGATATTTGTATT
MG666	R. oligo for qPCR of EutX set2	CAGTACGTCATTGTACCTGAAA
MG667	P1/P2 DNA probe for northern blot	GCCGATTGTTAAAACACGCC
MG675	F. oligo to amplify pJG019 backbone with SpeI RE-site	ACTAGTAGCCGATGCACACACATTATGCCACACCTTGTAGATAAAG
MG676	R. oligo to amplify pJG019 backbone with HindIII RE-site and P1P2T	GCGTGCTTAAGCTTAGCTTAAGGAGGAAAGTCACATTATGAG
MG681	F. oligo to Q5 SDM make EutV-Y101A on pMG1050	CGCACTGGGTGCATTAGTTAAACCTCTAG
MG682	R. oligo to Q5 SDM make EutV-Y101A on pMG1050	CCTAATTTTTTAGCTTTGTCC
MG683	F. oligo to Q5 SDM make EutV-K104A on pMG1050	TTATTTAGTTGCACCTCTAGATGAAAAATCATTAATAC
MG684	R. oligo to Q5 SDM make EutV-K104A on pMG1050	CCCAGTGCGCCTAATTTTTTAG
MG685	F. oligo to Q5 SDM make EutV-M117A on pMG1050	TACAATTGAAGCAAGCATTGAACGAGGC
MG686	R. oligo to Q5 SDM make EutV-M117A on pMG1050	GGTATTAATGATTTTTCATCTAGAG
MG687	F. oligo to Q5 SDM make EutV-Q124A on pMG1050	ACGAGGCAAAGCAACGCAGCTGTTAC
MG688	R. oligo to Q5 SDM make EutV-Q124A on pMG1050	TCAATGCTCATTTCAATTGTAGG
MG689	F. oligo to Q5 SDM make EutV-Q131A on pMG1050	GTTACTAAATGCAATCGATAAATTAAGTTTAAAATTAGAAG
MG690	R. oligo to Q5 SDM make EutV-Q131A on pMG1050	AGCTGCGTTTGTTTGCCTC

MG691	F. oligo to Q5 SDM make EutV-E140A on pMG1050	TTTAAAATTAGCAGAACGTAAAATTATCGAAAAAG
MG692	R. oligo to Q5 SDM make EutV-E140A on pMG1050	CTTAATTTATCGATTTGATTTAGTAAC
MG693	F. oligo to Q5 SDM make EutV-R142A on pMG1050	ATTAGAAGAAGCAAAAATTATCGAAAAAGCCAAAGG
MG694	R. oligo to Q5 SDM make EutV-R142A on pMG1050	TTTAAACTTAATTTATCGATTTGATTTAG
MG695	F. oligo to Q5 SDM make EutV-K149A on pMG1050	CGAAAAAGCCGCAGGTATTCTTGTAAAAG
MG696	R. oligo to Q5 SDM make EutV-K149A on pMG1050	ATAATTTTACGTTCTTCTAATTTTAAAC
MG697	F. oligo to Q5 SDM make EutV-E160A on pMG1050	TCATATATCGGCAGAAGAAGCCTAC
MG698	R. oligo to Q5 SDM make EutV-E160A on pMG1050	TTTTCTTTTACAAGAATACCTTTG
MG699	F. oligo to Q5 SDM make EutV-Y164A on pMG1050	AGAAGAAGCCGCACAAATGTTGCGTACG
MG700	R. oligo to Q5 SDM make EutV-Y164A on pMG1050	TCCGATATATGATTTTCTTTTACAAG
MG701	F. oligo to Q5 SDM make EutV-S171A on pMG1050	GCGTACGTTAGCAATGAACAAACGC
MG702	R. oligo to Q5 SDM make EutV-S171A on pMG1050	AACATTTGGTAGGCTTCTTC
MG703	F. oligo to Q5 SDM make EutV-M172A on pMG1050	TACGTTAAGTGCAAACAAACGCGCAC
MG704	R. oligo to Q5 SDM make EutV-M172A on pMG1050	CGCAACATTTGGTAGGCTTC
MG705	F. oligo to Q5 SDM make EutV-N173A on pMG1050	GTTAAGTATGGCAAAACGCGCACG
MG706	R. oligo to Q5 SDM make EutV-S173A on pMG1050	GTACGCAACATTTGGTAG
MG707	F. oligo to Q5 SDM make EutV-R175A on pMG1050	TATGAACAAAGCAGCACGTATGAGTGAAATTG
MG708	R. oligo to Q5 SDM make EutV-R175A on pMG1050	CTTAACGTACGCAACATTTG
MG709	F. oligo to Q5 SDM make EutV-M178A on pMG1050	ACGCGCACGTGCAAGTGAAATTG
MG710	R. oligo to Q5 SDM make EutV-M178A on pMG1050	TTGTTCATACTTAACGTACG
MG711	F. oligo to amplify pDG148	TGACGTACTGCTATATTCCCCAAGCTAATTCGGTGGAAACGAGG

MG712	R. oligo to amplify pDG148	TTTTCATAAAATATTCGTTTCTCTAGATC ACCTCCTTAAGCTTAATTGTTATCC
MG713	F. oligo to amplify and gibson assemble EutX into pDG148	CTTAAGGAGGTGATCTAGAGAAACGAATA TTTTATGAAAAATATGAATGGAAGCCAG
MG714	R. oligo to amplify and gibson assemble EutX into pDG148	GTTTCCACCGAATTAGCTTGGGGAATATA GCAGTACGTCATTGTACCTGAAATCTCC
MG681	F. oligo to Q5 SDM make EutV-Y101A on pMG1050	CGCACTGGGTGCATTAGTTAAACCTCTAG
MG682	R. oligo to Q5 SDM make EutV-Y101A on pMG1050	CCTAATTTTTTAGCTTTGTCC
MG683	F. oligo to Q5 SDM make EutV-K104A on pMG1050	TTATTTAGTTGCACCTCTAGATGAAAAAT CATTAATAC
MG684	R. oligo to Q5 SDM make EutV-K104A on pMG1050	CCCAGTGCGCCTAATTTTTTAG
MG685	F. oligo to Q5 SDM make EutV-M117A on pMG1050	TACAATTGAAGCAAGCATTGAACGAGGC
MG686	R. oligo to Q5 SDM make EutV-M117A on pMG1050	GGTATTAATGATTTTTCATCTAGAG
MG687	F. oligo to Q5 SDM make EutV-Q124A on pMG1050	ACGAGGCAAAGCAACGCAGCTGTTAC
MG688	R. oligo to Q5 SDM make EutV-Q124A on pMG1050	TCAATGCTCATTTCAATTGTAGG
MG689	F. oligo to Q5 SDM make EutV-Q131A on pMG1050	GTTACTAAATGCAATCGATAAATTAAGTT TAAAATTAGAAG
MG690	R. oligo to Q5 SDM make EutV-Q131A on pMG1050	AGCTGCGTTTGTTTGCCTC
MG691	F. oligo to Q5 SDM make EutV-E140A on pMG1050	TTTAAAATTAGCAGAACGTAAAATTATCG AAAAAG
MG692	R. oligo to Q5 SDM make EutV-E140A on pMG1050	CTTAATTTATCGATTTGATTTAGTAAC
MG693	F. oligo to Q5 SDM make EutV-R142A on pMG1050	ATTAGAAGAAGCAAAAATTATCGAAAAAG CCAAAGG
MG694	R. oligo to Q5 SDM make EutV-R142A on pMG1050	TTTAAACTTAATTTATCGATTTGATTTAG
MG695	F. oligo to Q5 SDM make EutV-K149A on pMG1050	CGAAAAAGCCGCAGGTATTCTTGTAAAAG
MG696	R. oligo to Q5 SDM make EutV-K149A on pMG1050	ATAATTTTACGTTCTTCTAATTTTAAAC
MG697	F. oligo to Q5 SDM make EutV-E160A on pMG1050	TCATATATCGGCAGAAGAAGCCTAC
MG698	R. oligo to Q5 SDM make EutV-E160A on pMG1050	TTTTCTTTTACAAGAATACCTTTG

MG699	F. oligo to Q5 SDM make EutV-Y164A on pMG1050	AGAAGAAGCCGCACAAATGTTGCGTACG
MG700	R. oligo to Q5 SDM make EutV-Y164A on pMG1050	TCCGATATATGATTTTCTTTTACAAG
MG701	F. oligo to Q5 SDM make EutV-S171A on pMG1050	GCGTACGTTAGCAATGAACAAACGC
MG702	R. oligo to Q5 SDM make EutV-S171A on pMG1050	AACATTTGGTAGGCTTCTTC
MG703	F. oligo to Q5 SDM make EutV-M172A on pMG1050	TACGTTAAGTGCAAACAAACGCGCAC
MG704	R. oligo to Q5 SDM make EutV-M172A on pMG1050	CGCAACATTTGGTAGGCTTC
MG705	F. oligo to Q5 SDM make EutV-N173A on pMG1050	GTTAAGTATGGCAAAACGCGCACG
MG706	R. oligo to Q5 SDM make EutV-S173A on pMG1050	GTACGCAACATTTGGTAG
MG707	F. oligo to Q5 SDM make EutV-R175A on pMG1050	TATGAACAAAGCAGCACGTATGAGTGAAATTG
MG708	R. oligo to Q5 SDM make EutV-R175A on pMG1050	CTTAACGTACGCAACATTTG
MG709	F. oligo to Q5 SDM make EutV-M178A on pMG1050	ACGCGCACGTGCAAGTGAAATTG
MG710	R. oligo to Q5 SDM make EutV-M178A on pMG1050	TTGTTCATACTTAACGTACG
MG711	F. oligo to amplify pDG148	TGACGTACTGCTATATTCCCCAAGCTAATTCGGTGGAAACGAGG
MG712	R. oligo to amplify pDG148	TTTTCATAAAATATTCGTTTCTCTAGATCACCTCCTTAAGCTTAATTGTTATCC
MG713	F. oligo to amplify and gibson assemble EutX into pDG148	CTTAAGGAGGTGATCTAGAGAAACGAATATTTTATGAAAAATATGAATGGAAGCCAG
MG714	R. oligo to amplify and gibson assemble EutX into pDG148	GTTTCCACCGAATTAGCTTGGGGAATATAGCAGTACGTCATTGTACCTGAAATCTCC
MG727	F. oligo to Q5 SDM make EutV-K143A on pMG1050	AGAAGAACGTGCAATTATCGAAAAAGC
MG728	R. oligo to Q5 SDM make EutV-K143A on pMG1050	AATTTTAAACTTAATTTATCGATTTGATTTAG
MG729	F. oligo to Q5 SDM make EutV-K147A on pMG1050	AATTATCGAAGCAGCCAAAGGTATTC
MG730	R. oligo to Q5 SDM make EutV-K147A on pMG1050	TTACGTTCTTCTAATTTTAAACTTAATTTATC
MG731	F. oligo to Q5 SDM make EutV-K154A on pMG1050	TATTCTTGTAGCAGAAAATCATATATCGGAAG

MG732	R. oligo to Q5 SDM make EutV-K154A on pMG1050	CCTTTGGCTTTTTCGATAATTTTAC
MG733	F. oligo to Q5 SDM make EutV-Q165A on pMG1050	AGAAGCCTACGCAATGTTGCGTAC
MG734	R. oligo to Q5 SDM make EutV-Q165A on pMG1050	TCTTCCGATATATGATTTTCTTTTAC
MG735	F. oligo to Q5 SDM make EutV-R168A on pMG1050	CCAAATGTTGGCAACGTTAAGTATGAACAAACG
MG736	R. oligo to Q5 SDM make EutV-R168A on pMG1050	TAGGCTTCTTCTTCCGATATATG
MG737	F. oligo to Q5 SDM make EutV-Q165A/R168A on pMG1050	TTGGCAACGTTAAGTATGAACAAACGC
MG738	R. oligo to Q5 SDM make EutV-Q165A/R168A on pMG1050	CATTGCGTAGGCTTCTTCTTCCGATATATG
MG739	F. oligo to Q5 SDM make EutV-Q165R/R168Q on pMG1050	TTGCAAACGTTAAGTATGAACAAACGC
MG740	R. oligo to Q5 SDM make EutV-Q165R/R168Q on pMG1050	CATACGGTAGGCTTCTTCTTCCGATATATG
MG741	F. oligo to amplify pMG1050 (backbone not including EutV/W)	GGAGAGTCACACCTAAAGAAAAGCTTATCGAATTCGATAACCCTAAAGTTATG
MG742	R. oligo to amplify pMG1050 (backbone not including EutV/W)	TTGCCAGCCATATTGTTCATAATGTGACTTTCCTCCTTAAGCTTCG
MG743	F. oligo to amplify NasR from gBlock to gibson clone into pMG1050	TTAAGGAGGAAAGTCACATTATGAACAATATGGCTGGCAACACT
MG744	R. oligo to amplify NasR from gBlock to gibson clone into pMG1050	TTATCGAATTCGATAAGCTTTTCTTTAGGTGTGACTCTCCAGAGTG
MG745	F. oligo to amplify NasF leader region containing P1/P2/Terminator to be gibson cloned into pMG1044	TGTGGCATAATGTGTGTGCAGAGTGAATAAAAGGTTTTGGGCAGC
MG746	R. oligo to amplify NasF leader region containing P1/P2/Terminator to be gibson cloned into pMG1044	GGGGCGGGGTTTTTGCTAGCTTATTTATACAGTTCGTCCATACCGTGGGT
MG747	F. oligo to amplify pMG1044 (backbone not including Efae P1P2T)	TGGACGAACTGTATAAATAAGCTAGCAAAAACC
MG748	R. oligo to amplify pMG1044 (backbone not including Efae P1P2T)	CCAAAACCTTTTATTCACTCTGCACACACATTATGCCACACCT
MG749	F. oligo to Q5 SDM make P2 C1U4 in pMG1088 to P2 C1G4	AATCGGCCAAGGAGCCCAAGA

MG750	R. oligo to Q5 SDM make P2 C1U4 in pMG1088 to P2 C1G4	TGTTAAAACACGCAATGGTGTTTC
MG751	F. oligo to amplify pMG1050 backbone (not including ANTAR domain)	AAAAGGAGGAAAGTCACATTATGAAACGATTAGAG
MG752	R. oligo to amplify pMG1050 backbone (not including ANTAR domain)	TGATTTAGTAACAGCTGCGTTTGTTTGCCTCG
MG727	F. oligo to Q5 SDM make EutV-K143A on pMG1050	AGAAGAACGTGCAATTATCGAAAAAGC
MG728	R. oligo to Q5 SDM make EutV-K143A on pMG1050	AATTTTAAACTTAATTTATCGATTTGATTTAG
MG729	F. oligo to Q5 SDM make EutV-K147A on pMG1050	AATTATCGAAGCAGCCAAAGGTATTC
MG730	R. oligo to Q5 SDM make EutV-K147A on pMG1050	TTACGTTCTTCTAATTTTAAACTTAATTTATC
MG731	F. oligo to Q5 SDM make EutV-K154A on pMG1050	TATTCTTGTAGCAGAAAATCATATATCGGAAG
MG732	R. oligo to Q5 SDM make EutV-K154A on pMG1050	CCTTTGGCTTTTTCGATAATTTTAC
MG733	F. oligo to Q5 SDM make EutV-Q165A on pMG1050	AGAAGCCTACGCAATGTTGCGTAC
MG734	R. oligo to Q5 SDM make EutV-Q165A on pMG1050	TCTTCCGATATATGATTTTCTTTTAC
MG735	F. oligo to Q5 SDM make EutV-R168A on pMG1050	CCAAATGTTGGCAACGTTAAGTATGAACAAACG
MG736	R. oligo to Q5 SDM make EutV-R168A on pMG1050	TAGGCTTCTTCTTCCGATATATG
MG737	F. oligo to Q5 SDM make EutV-Q165A/R168A on pMG1050	TTGGCAACGTTAAGTATGAACAAACGC
MG738	R. oligo to Q5 SDM make EutV-Q165A/R168A on pMG1050	CATTGCGTAGGCTTCTTCTTCCGATATATG
MG739	F. oligo to Q5 SDM make EutV-Q165R/R168Q on pMG1050	TTGCAAACGTTAAGTATGAACAAACGC
MG740	R. oligo to Q5 SDM make EutV-Q165R/R168Q on pMG1050	CATACGGTAGGCTTCTTCTTCCGATATATG
MG741	F. oligo to amplify pMG1050 (backbone not including EutV/W)	GGAGAGTCACACCTAAAGAAAAGCTTATCGAATTCGATAACCCTAAAGTTATG
MG742	R. oligo to amplify pMG1050 (backbone not including EutV/W)	TTGCCAGCCATATTGTTCATAATGTGACTTTCCTCCTTAAGCTTCG
MG743	F. oligo to amplify NasR from gBlock to gibson clone into pMG1050	TTAAGGAGGAAAGTCACATTATGAACAATATGGCTGGCAACACT

MG744	R. oligo to amplify NasR from gBlock to gibson clone into pMG1050	TTATCGAATTCGATAAGCTTTTCTTTAGGTGTGACTCTCCAGAGTG
MG745	F. oligo to amplify NasF leader region containing P1P2T to be gibson cloned into pMG1044	TGTGGCATAATGTGTGTGCAGAGTGAATAAAAGGTTTTGGGCAGC
MG746	R. oligo to amplify NasF leader region containing P1P2T to be gibson cloned into pMG1044	GGGGCGGGGTTTTTGCTAGCTTATTTATACAGTTCGTCCATACCGTGGGT
MG747	F. oligo to amplify pMG1044	TGGACGAACTGTATAAATAAGCTAGCAAAAACC
MG748	R. oligo to amplify pMG1044	CCAAAACCTTTTATTCACTCTGCACACACATTATGCCACACCT
MG749	F. oligo to Q5 SDM make P2 C1U4 in pMG1088 to P2 C1G4	AATCGGCCAAGGAGCCCAAGA
MG750	R. oligo to Q5 SDM make P2 C1U4 in pMG1088 to P2 C1G4	TGTTAAAACACGCAATGGTGTTTC
MG751	F. oligo to amplify pMG1050 backbone	AAAAGGAGGAAAGTCACATTATGAAACGATTAGAG
MG752	R. oligo to amplify pMG1050 backbone	TGATTTAGTAACAGCTGCGTTTGTTTGCCTCG
MG757	F. oligo to gibson clone *K. oxytoca* NasR into pVL847 (his10-MBP-tag)	ACTGCTAGCCATATACTGGTTCCGCGTGGATCTCAAATGAACAATATGGCTGGCAACAC
MG761	R. oligo to amplify Koxytoca NasR for gibson assembly with NdeI/BamHI pVL847	GATCTCAATTAATCAGGATCTTCTTTAGGTGTGACTCTCCAGAGTGC
MG772	F. oligo to revert R168A in EutV of pMG1124 to R168R (168)	CGCAATGTTGCGTACGTTAAGTATGAACAAAC
MG773	R. oligo to revert R168A in EutV of pMG1124 to R168R (168)	TAGGCTTCTTCTTCCGATATATG
MG774	F. oligo to Q5 SDM make EutV-R142K on pMG1050	ATTAGAAGAAAAAAAAATTATCGAAAAAGCCAAAG
MG775	R. oligo to Q5 SDM make EutV-R142K on pMG1050	TTTAAACTTAATTTATCGATTTGATTTAG
MG776	F. oligo to Q5 SDM make EutV-K143R on pMG1050	AGAAGAACGTAGAATTATCGAAAAAG
MG777	R. oligo to Q5 SDM make EutV-K143R on pMG1050	AATTTTAAACTTAATTTATCGATTTGATTTAG
MG778	F. oligo to Q5 SDM make EutV-K147R on pMG1050	AATTATCGAAAGAGCCAAAGGTATTC
MG779	R. oligo to Q5 SDM make EutV-K147R on pMG1050	TTACGTTCTTCTAATTTTAAACTTAATTTATC

MG780	F. oligo to Q5 SDM make EutV-K149R on pMG1050	CGAAAAAGCCAGAGGTATTCTTG
MG781	R. oligo to Q5 SDM make EutV-K149R on pMG1050	ATAATTTTACGTTCTTCTAATTTTAAAC
MG786	F. oligo to Q5 SDM make EutV-R168K on pMG1050	CCAAATGTTGAAAACGTTAAGTATGAACAAAC
MG787	R. oligo to Q5 SDM make EutV-R168K on pMG1050	TAGGCTTCTTCTTCCGATATATG
MG790	F. oligo to SDM make CCC to GGG on EutX present in pMG1115	CTGGCACGGTGGGGCCACTGTGA
MG791	R. oligo to SDM make CCC to GGG on EutX present in pMG1115	ATTCTCACTGGCTTCCATTC
MG792	F. oligo to amplify pDG148 for gibson assembly with EutX (+SST)	CCAATTCGTTGGGTTTTTTTCAAGCTAATTCGGTGGAAACGAGG
MG793	R. oligo to amplify pDG148 for gibson assembly with EutX (+SST)	TTTTCATAAAATATTCGTTTCTCTAGATCACCTCCTTAAGCTTAATTGTTATCC
MG794	F. oligo to amplify and gibson assemble EutX into pDG148 (with SST)	CTTAAGGAGGTGATCTAGAGAAACGAATATTTTATGAAAAATATGAATGGAAGCCAG
MG795	R. oligo to amplify and gibson assemble EutX into pDG148 (with SST)	GTTTCCACCGAATTAGCTTGAAAAAAACCCAACGAATTGGGAATATAGC
MG801	F. oligo to amplify NasF P1P2T with T7 promoter, PCR product is 138 bp	TAATACGACTCACTATAGGGAGTGAATAAAAGGTTTTGG
MG802	R. oligo to amplify NasF P1P2T with T7 promoter, PCR product is 138 bp	AACCGCTCCAGAAACGCGCAAAAAAAAAGC
MG803	F. oligo to Q5 SDM delete LINKER from pMG1133	GGGCTGGTCTTTCAATGTTG
MG804	R. oligo to Q5 SDM delete LINKER from pMG1133	GAAGATCATGTGATTTCTCTTTGTTC
MG811	F. oligo to amplify NasF P1P2 (no terminator) with T7 promoter, PCR product is 61 bp	TAATACGACTCACTATAGGGCGCGCCAATGGC
MG812	R. oligo to amplify NasF P1P2 (no terminator) with T7 promoter, PCR product is 61 bp	GGACGCCTTTATCCCTGGACATAC
MG815	F. oligo to Q5 SDM make NasR-R340A on pMG1130	TCTGGAAGAAGCAAAATTGATCGAGAAAGCC

MG847	F. oligo to amplify EutS Promoter-P7P8T with BamHI at 5'	AACCTGGATCCATTTTAAATTCTATCTTATAAGCGATTG
MG848	R. oligo to amplify EutS Promoter-P7P8T with HindIII at 5'	AACCTAAGCTTTTCGTTGTTTTTCTTCCAACC
MG864	F. oligo for NasF P1P2 with extended leader (with T7 oligo)	TAATACGACTCACTATAGGGAGTGAATAAAAGG
MG865	R. oligo for NasF P1P2 with extended leader	GGACGCCTTTATCCCTGGAC

Table 6. List of gBlocks used in these studies

Name	Purpose	Sequence- 5' to 3'
EutX	Full length EutX, sequence from *E. faecalis*	GTTTTCCCAGTCACGACGTTGTAAAACG ACGGCCAGTGCCTAATACGACTCACTAT AGGGAAGCTTAAATTTCAGATAGATTAA ACGAATATTTTATGAAAAATATGAATGG AAGCCAGTGAGAATCTGGCACGGTCCCG CCACTGTGAAGAAGCAAGGTTGCTTTTA AGTCAGGTCTTTTCATTTTTTCATTATT GGGCATGCTGTTTCGAGGCAAAACAGGA TGTTTCTTAACAACGCTTGTGTTGAAAA TCCAGCCCAAGATATTTGTATTAATCCA ATTAATGGCACGAGGAAGTGCTTTTGGG AATTTTGGGCTGGTCTTTCAATGTTGAA GGAGATTTCAGGTACAATGACGTACTGC TATATTCCCTTTTTTTATTGACAATTAA TTAAAGGCGTTGTACTCGAATTCCCCTA TAGTGAGTCGTATTAGTAATCATGGTCA TAGCTGTTTCCTGTGTGAAATTGTTAT
M3 EutX mutant	Full length EutX M3 mutant (C_3 to G_3), sequence from *E. faecalis*	GTTTTCCCAGTCACGACGTTGTAAAACG ACGGCCAGTGCCTAATACGACTCACTAT AGGGAAGCTTAAATTTCAGATAGATTAA ACGAATATTTTATGAAAAATATGAATGG AAGCCAGTGAGAATCTGGCACGGTGGGG CCACTGTGAAGAAGCAAGGTTGCTTTTA AGTCAGGTCTTTTCATTTTTTCATTATT GGGCATGCTGTTTCGAGGCAAAACAGGA TGTTTCTTAACAACGCTTGTGTTGAAAA TCCAGCCCAAGATATTTGTATTAATCCA ATTAATGGCACGAGGAAGTGCTTTTGGG AATTTTGGGCTGGTCTTTCAATGTTGAA GGAGATTTCAGGTACAATGACGTACTGC TATATTCCCTTTTTTTATTGACAATTAA TTAAAGGCGTTGTACTCGAATTCCCCTA TAGTGAGTCGTATTAGTAATCATGGTCA TAGCTGTTTCCTGTGTGAAATTGTTAT
P_{glyQS}-EutX	P_{glyQS} 5' of WT EutX. gBlock used for Transcription Termination Assay	ATTGATTTATATTACGAAGAATATTCGG GATTGTATTTAAAATCAAAGCGCTTTTT AGATCAAATGGAAAGCATGAAACATCTT ATGGGTGAAAACAAAAGTTGACATTTGG TCCATCTTTTTATATGATCATTTATTAT AAGAATATTTTATGAAAAATATGAATGG AAGCCAGTGAGAATCTGGCACGGTCCCG CCACTGTGAAGAAGCAAGGTTGCTTTTA AGTCAGGTCTTTTCATTTTTTCATTATT GGGCATGCTGTTTCGAGGCAAAACAGGA TGTTTCTTAACAACGCTTGTGTTGAAAA TCCAGCCCAAGATATTTGTATTAATCCA ATTAATGGCACGAGGAAGTGCTTTTGGG AATTTTGGGCTGGTCTTTCAATGTTGAA GGAGATTTCAGGTACAATGACGTACTGC TATATTCCCTTTTTTT

P_{glyQS}-EutX M3 mutant	P_{glyQS} 5' of pseudoknot mutation EutX. gBlock used for Transcription Termination Assay	ATTGATTTATATTACGAAGAATATTCGG GATTGTATTTAAAATCAAAGCGCTTTTT AGATCAAATGGAAAGCATGAAACATCTT ATGGGTGAAAACAAAAGTTGACATTTGG TCCATCTTTTTATATGATCATTTATTAT AAGAATATTTTATGAAAAATATGAATGG AAGCCAGTGAGAATCTGGCACGGTGGGG CCACTGTGAAGAAGCAAGGTTGCTTTTA AGTCAGGTCTTTTCATTTTTTCATTATT GGGCATGCTGTTTCGAGGCAAAACAGGA TGTTTCTTAACAACGCTTGTGTTGAAAA TCCAGCCCAAGATATTTGTATTAATCCA ATTAATGGCACGAGGAAGTGCTTTTGGG AATTTTGGGCTGGTCTTTCAATGTTGAA GGAGATTTCAGGTACAATGACGTACTGC TATATTCCCTTTTTTT
NasR	NasR DNA sequence from *K. oxytoca*, codon optimized for *B. subtilis*	ATGAACAATATGGCTGGCAACACTCCAG AGGTAGTCGATTGGTTTGCACGGGCTCG TCGTCTTCAAAAACAACAATTGCATCAA CTCGCGCAGCAGGGGACCCTTGCCGGTC AAATTTCTGCTCTCGTGCACATGCTGCA ATGCGAACGGGGGGCTTCTAACATTTGG CTTTGTTCTGGGGGGCGTCTCTATGCCG CGGAGTGTCGTGCAGGAGCGGCTCTGGT AGATGAGCAACTCACTCGCTTCTATGCA GCGCTTGAGCCAGCGCGTGACGCCGCCT CATCTGCACTGTGCTGGCGCATCGCTTG TGCAGTCTGGTATCTTCCACAGCTCGCC GCATTGCGTAAACGTGTGAGAGACCGCG AGATTGCAGCGGAGGAAGCTACCGGGCA ATTTAGCCGTATCATAAGACACTTGCTC AATATCGTTCCTCAGCTCAACGATTCTA TCGATGATCCGCAAATAGCCGGACGTAT GGTAGCACTTTACAGTTTCATGCAGGGA AAAGAGCTGGCTGGGCAAGAACGTGCTT TGGGAGCTTTGGGCTTTGCGCGCGGGCA ATTCTCCGACGAACTTAGACAGCAGCTG GTAGATAGAATTGATGGACAACAGCCTT GTTTTGACTCATTTCAAGCCCTTGCGCA ACCACCGCAAACGGCTCTTTTCGCCGAG CAATGTCAGGCCTCTCTGGAGATTGAGC AACTTCGCAGAGTCGCCTGTACGCGCCA GCCTCCTGCTGATGAGGGGGAGACAGCT TTGCGTTGGTTCTGCGCTCAGACCCAAC GTCTTGAGCAGCTGAGAGGGGTAGAGGA GTTGTTAATAGTTGACTTGTTAAATGCT GCTGACGCATTGCTGGAAGGTGAAGAGC CAGAGGCTCAGCTTCCTCCAGCAGACTG GCAAGAGGATAGTATTGCACTTAGACTC GATAAACAGCTCCTTCCACTCGTTCGCC AGCAGGCGCATGAATTACAACAGCTGTC AGGCCAACTGGCTAGTTTGAAGGATGCT CTGGAAGAACGTAAATTGATCGAGAAAG CCAAATCCGTCCTTATGACGTACCAGGG GATGCAGGAGGAACAAGCGTGGCAGGCG CTTCGGAAAATGGCAATGGATAAAAATC

AAAGAATGGTCGAGATTGCACGGGCGCT
CCTGACGGTCAAGGCACTCTGGAGAGTC
ACACCTAAAGAA

NasF P1P2T	NasF P1P2T DNA sequence from *K. oxytoca*	GAGTGAATAAAAGGTTTTGGGCAGCGCG CCAATGGCGGCGCGTATGTCCAGGGATA AAGGCGTCCAGCGGTGCGTAAGCACCGC CGGGCGCTTTTTTTTTGCGCGTTTCTGG AGCGGTTATGGAGGTAGCGATGAAACGA ATCATTTTAATGGGAGCTATCGGTGAAG CTTCAAGGAGGAAAGTCACATTAAAGCT TCAAGGAGGAAAGTCACATTATGAGCAA AGGTGAAGAACTGTTCACCGGCGTTGTG CCAATTCTGGTTGAGCTGGATGGTGACG TGAATGGCCACAAATTTTCCGTGTCTGG TGAAGGCGAGGGTGATGCTACTTATGGC AAACTGACTCTGAAACTGATCTGTACCA CCGGCAAACTGCCTGTTCCGTGGCCAAC TCTGGTCACTACTCTGGGTTACGGCCTG ATGTGTTTTGCGCGTTACCCGGATCACA TGAAACAGCATGACTTCTTCAAATCTGC CATGCCGGAAGGCTATGTCCAAGAACGT ACGATCTTTTTCAAGGACGACGGCAACT ATAAAACCCGTGCCGAAGTTAAATTCGA GGGTGACACCCTGGTTAACCGCATCGAA CTGAAAGGCATTGACTTCAAAGAGGACG GCAACATTCTGGGTCACAAGCTGGAATA CAACTACAACTCCCACAACGTTTACATT ACTGCTGACAAGCAGAAAAACGGCATCA AAGCAAACTTCAAGATCCGTCACAACAT TGAAGATGGTGGCGTACAGCTGGCAGAT CACTACCAGCAGAACACTCCAATCGGTG ATGGCCCAGTACTGCTGCCAGATAACCA TTACCTGTCCTACCAGAGCAAACTGTCT AAAGACCCGAACGAAAAACGTGACCACA TGGTACTGCTGGAATTTGTTACCGCGGC AGGCATTACCCACGGTATGGACGAACTG TATAAATAA

GlmS*-P3P4	M9 (AG to CC) mutant GlmS ribozyme fused to P3P4 (including linker region) from EutX	GAAAAATTTTGCAAAAAGTTGTTGACTT TATCTACAAGGTGTGGCATAATGTGTGG AATTGTGAGCGGATAACAATTAAGCTTA AGGAGGTGATCTAGAGGTCTTGTTCTTA TTTTCTCAATAGGAAAAGAAGACGGGAT TATTGCTTTACCTATAATTATCCCGCCC GAACTAAGCGCCCGGAAAAAGGCTTAGT TGACGAGGATGGAGGTTATCGAATTTTC GGCGGATGCCTCCCGGCTGAGTGTGCAG ATCACAGCCGTAAGGATTTCTTCAAACC AAGGGGGTGACTCCTTGAACAAAGAGAA ATCACATGATCTTCAACAACGCTTGTGT TGAAAATCCAGCCCAAGATATTTGTATT AATCCAATTAATGGCACGAGGAAGTGCT TTTGGGAATTTTGGGCTGGTCTTTCAAT GTTGAAGGAGATTTCAGGTACAATGACG TACTGCTATATTCCCAATTCGTTGGGTT TTTTTCAAGCTAATTCGGTGGAAACGAG GTCATCATTTCCTTCCGAAAAAACGGTT GCATTTAAATCT

Table 7. List of plasmids used in these studies.

Name	Description	Reference
pMG1040	EutX	(132)
pMG1041	M3 EutX	(132)
pMG1044	*amyE::* P_{const} P1P2T-YFP	This Study
pMG1050	*thrC::* P_{liaI} EutV/EutW	This Study
pMG1100	*thrC::* P_{liaI} EutV Y101A/EutW	This Study
pMG1101	*thrC::* P_{liaI} EutV K104A/EutW	This Study
pMG1102	*thrC::* P_{liaI} EutV M117A/EutW	This Study
pMG1103	*thrC::* P_{liaI} EutV Q124A/EutW	This Study
pMG1104	*thrC::* P_{liaI} EutV Q131A/EutW	This Study
pMG1105	*thrC::* P_{liaI} EutV E140A/EutW	This Study
pMG1106	*thrC::* P_{liaI} EutV R142A/EutW	This Study
pMG1107	*thrC::* P_{liaI} EutV K149A/EutW	This Study
pMG1108	*thrC::* P_{liaI} EutV E160A/EutW	This Study
pMG1109	*thrC::* P_{liaI} EutV Y164A/EutW	This Study
pMG1110	*thrC::* P_{liaI} EutV S171A/EutW	This Study
pMG1111	*thrC::* P_{liaI} EutVM172A/EutW	This Study
pMG1112	*thrC::* P_{liaI} EutV N173A/EutW	This Study
pMG1113	*thrC::* P_{liaI} EutV R175A/EutW	This Study
pMG1114	*thrC::* P_{liaI} EutV M178A/EutW	This Study
pMG1115	EutX with semi-synthetic terminator in pDG148	This Study
pMG1119	*thrC::* P_{liaI} EutV K143A/EutW	This Study
pMG1120	*thrC::* P_{liaI} EutV K147A/EutW	This Study
pMG1121	*thrC::* P_{liaI} EutV K154A/EutW	This Study
pMG1122	*thrC::* P_{liaI} EutV Q165A/EutW	This Study
pMG1123	*thrC::* P_{liaI} EutV R168A/EutW	This Study
pMG1124	*thrC::* P_{liaI} EutV Q165A R168A/EutW	This Study
pMG1125	*thrC::* P_{liaI} EutV Q165R R168Q/EutW	This Study
pMG1126	*thrC::* P_{liaI} NasR (from *K. oxytoca*)	This Study
pMG1127	*amyE::* P_{const} P1P2$_{nasF}$-YFP	This Study
pMG1130	His_{10}-MBP- NasR (from *K. oxytoca*)	This Study
pMG1136	*thrC::* P_{liaI} EutV R142K/EutW	This Study
pMG1137	*thrC::* P_{liaI} EutV K143R/EutW	This Study
pMG1138	*thrC::* P_{liaI} EutV K147R/EutW	This Study
pMG1139	*thrC::* P_{liaI} EutV K149R/EutW	This Study
pMG1140	*thrC::* P_{liaI} EutV R168K/EutW	This Study
pMG1141	EutX* with semi-synthetic terminator in	This Study

	pDG148 (* denotes C1U4 mutations in loops)	
pMG1142	His_{10}-MBP- NasR R340A (from *K. oxytoca*)	This Study
pMG1143	His_{10}-MBP- NasR R340K (from *K. oxytoca*)	This Study
pMG1156	*amyE::* P_{eutS} P7P8T-YFP	This Study

Table 8. List of strains used in these studies.

Name	Description	Organism	Reference
168	Wild-type	*B. subtilis*	(Kunst et al. 1997)
OG1RF	Wild-type	*E. faecalis*	(Bourgogne et al. 2008)
ΔEutX	Deletion of EutX from chromosome	*E. faecalis*	(DebRoy et al. 2014)
ΔEutV/W	Deletion of EutV/W from chromosome	*E. faecalis*	(DebRoy et al. 2014)
EFKK6	EutM-GFP in ectopic location, under endogenous promoter control	*E. faecalis*	This Study
ΔEutX plasmid	Deletion of EutX from extrachromosomal plasmid	*E. faecalis*	(DebRoy et al. 2014)
ΔRb plasmid	Deletion of AdoCbl Rb from extrachromosomal plasmid	*E. faecalis*	(DebRoy et al. 2014)
ΔP3P4 plasmid	Deletion of P3P4 from extrachromosomal plasmid	*E. faecalis*	(DebRoy et al. 2014)
MG5141	*amyE*:: P_{const} P1P2-Term-YFP	*E. coli*	This Study
MG5147	*amyE*:: P_{const} P1P2-Term-YFP	*B. subtilis*	This Study
MG5182	*thrC*::P_{liaI} -EutV/W	*E. coli*	This Study
MG5183	*thrC*:: P_{liaI} -EutV/W	*B. subtilis*	This Study
MG5187	*amyE*:: P_{const} P1P2-Term-YFP, *thrC*:: P_{liaI} -EutV/W	*B. subtilis*	This Study
MG5195	*amyE*::P1P2-Term-YFP (no promoter)	*E. coli*	This Study
MG5196	*amyE*::P1P2-Term-YFP (no promoter), *thrC*:: P_{liaI} -EutV/W	*B. subtilis*	This Study
MG5197	*amyE*:: P_{const} P1P2-Term-YFP with C1U4 in P1P2 loops	*E. coli*	This Study
MG5198	*amyE*:: P_{const} P1P2-Term-YFP with C1U4 in P1P2 loops, *thrC*:: P_{liaI} -EutV/W	*B. subtilis*	This Study
MG5200	*amyE*::P_{const} P1P2-Term-YFP with C1U4 in P1 loop (WT P2)	*E. coli*	This Study
MG5201	amyE:: P_{const} P1P2-Term-YFP with C1U4 in P1 loop (WT P2), thrC:: PliaI -EutV/W	*B. subtilis*	This Study
MG5225	*thrC*:: P_{liaI} EutV Y101A /EutW	*E. coli*	This Study
MG5226	*thrC*:: P_{liaI} EutV K104A /EutW	*E. coli*	This Study
MG5227	*thrC*:: P_{liaI} EutV M117A /EutW	*E. coli*	This Study
MG5228	*thrC*:: P_{liaI} EutV Q124A /EutW	*E. coli*	This Study
MG5229	*thrC*:: P_{liaI} EutV Q131A /EutW	*E. coli*	This Study

MG5230	*thrC*:: P_{liaI} EutV E140A /EutW	*E. coli*	This Study
MG5231	*thrC*:: P_{liaI} EutV R142A /EutW	*E. coli*	This Study
MG5232	*thrC*:: P_{liaI} EutV K149A /EutW	*E. coli*	This Study
MG5233	*thrC*:: P_{liaI} EutV E160A /EutW	*E. coli*	This Study
MG5234	*thrC*:: P_{liaI} EutV Y164A /EutW	*E. coli*	This Study
MG5235	*thrC*:: P_{liaI} EutV S171A /EutW	*E. coli*	This Study
MG5236	*thrC*:: P_{liaI} EutV M172A /EutW	*E. coli*	This Study
MG5237	*thrC*:: P_{liaI} EutV N173A /EutW	*E. coli*	This Study
MG5238	*thrC*:: P_{liaI} EutV R175A /EutW	*E. coli*	This Study
MG5239	*thrC*:: P_{liaI} EutV M178A /EutW	*E. coli*	This Study
MG5240	EutX with SST in pDG148	*E. coli*	This Study
MG5241	*amyE*:: P_{const} P1P2-Term-YFP, *thrC*:: P_{liaI} EutV Y101A /EutW	*B. subtilis*	This Study
MG5242	*amyE*:: P_{const} P1P2-Term-YFP, *thrC*:: P_{liaI} EutV K104A /EutW	*B. subtilis*	This Study
MG5243	*amyE*:: P_{const} P1P2-Term-YFP, *thrC*:: P_{liaI} EutV M117A /EutW	*B. subtilis*	This Study
MG5244	*amyE*:: P_{const} P1P2-Term-YFP, *thrC*:: P_{liaI} EutV Q124A /EutW	*B. subtilis*	This Study
MG5245	*amyE*:: P_{const} P1P2-Term-YFP, *thrC*:: P_{liaI} EutV Q131A /EutW	*B. subtilis*	This Study
MG5246	*amyE*:: P_{const} P1P2-Term-YFP, *thrC*:: P_{liaI} EutV E140A /EutW	*B. subtilis*	This Study
MG5247	*amyE*:: P_{const} P1P2-Term-YFP, *thrC*:: P_{liaI} EutV R142A /EutW	*B. subtilis*	This Study
MG5248	*amyE*:: P_{const} P1P2-Term-YFP, *thrC*:: P_{liaI} EutV E149A /EutW	*B. subtilis*	This Study
MG5249	*amyE*:: P_{const} P1P2-Term-YFP, *thrC*:: P_{liaI} EutV E160A /EutW	*B. subtilis*	This Study
MG5250	*amyE*:: P_{const} P1P2-Term-YFP, *thrC*:: P_{liaI} EutV Y164A /EutW	*B. subtilis*	This Study
MG5251	*amyE*:: P_{const} P1P2-Term-YFP, *thrC*:: P_{liaI} EutV S171A /EutW	*B. subtilis*	This Study
MG5252	*amyE*:: P_{const} P1P2-Term-YFP, *thrC*:: P_{liaI} EutV M172A /EutW	*B. subtilis*	This Study
MG5253	*amyE*:: P_{const} P1P2-Term-YFP, *thrC*:: P_{liaI} EutV N173A /EutW	*B. subtilis*	This Study
MG5254	*amyE*:: P_{const} P1P2-Term-YFP, *thrC*:: P_{liaI} EutV R175A /EutW	*B. subtilis*	This Study
MG5255	*amyE*:: P_{const} P1P2-Term-YFP, *thrC*:: P_{liaI} EutV M178A /EutW	*B. subtilis*	This Study
MG5257	*amyE*:: P_{const} P1P2-Term-YFP, *thrC*:: P_{liaI} EutV/W, EutX with SST in pDG148	*B. subtilis*	This Study

MG5269	*thrC*:: P_{liaI} EutV K143A /EutW	*E. coli*	This Study
MG5270	*thrC*:: P_{liaI} EutV K147A /EutW	*E. coli*	This Study
MG5271	*thrC*:: P_{liaI} EutV K154A /EutW	*E. coli*	This Study
MG5272	*thrC*:: P_{liaI} EutV Q165A /EutW	*E. coli*	This Study
MG5273	*thrC*:: P_{liaI} EutV R168A /EutW	*E. coli*	This Study
MG5274	*thrC*:: P_{liaI} EutV Q165AR168A /EutW	*E. coli*	This Study
MG5275	*thrC*:: P_{liaI} EutV Q165RR168Q /EutW	*E. coli*	This Study
MG5276	*thrC*:: P_{liaI} NasR (from *K. oxytoca*)	*E. coli*	This Study
MG5277	*amyE::* P_{const} P1P2$_{nasF}$- YFP	*E. coli*	This Study
MG5278	*amyE*:: P_{const} P1P2-Term-YFP with C1U4 in P1 loop and C1G4 in P2 loop	*E. coli*	This Study
MG5280	His_{10} MBP NasR (From *K. oxytoca*) in pVL847CbR	*E. coli*	This Study
MG5281	*amyE::* P_{const} P1P2-Term-YFP, *thrC::* P_{liaI} EutV K143A /EutW	*B. subtilis*	This Study
MG5282	*amyE::* P_{const} P1P2-Term-YFP, *thrC::* P_{liaI} EutV K147A /EutW	*B. subtilis*	This Study
MG5283	*amyE::* P_{const} P1P2-Term-YFP, *thrC::* P_{liaI} EutV K154A /EutW	*B. subtilis*	This Study
MG5284	*amyE::* P_{const} P1P2-Term-YFP, *thrC::* P_{liaI} EutV Q165A /EutW	*B. subtilis*	This Study
MG5285	*amyE::* P_{const} P1P2-Term-YFP, *thrC::* P_{liaI} EutV R168A /EutW	*B. subtilis*	This Study
MG5286	*amyE::* P_{const} -P1P2-Term-YFP, thrC:: P_{liaI} EutV Q165A R168A /EutW	*B. subtilis*	This Study
MG5287	*amyE::* P_{const} P1P2-Term-YFP, *thrC::* P_{liaI} EutV Q165RR168Q /EutW	*B. subtilis*	This Study
MG5288	*amyE::* P_{const} P1P2-Term-YFP, *thrC::* P_{liaI} NasR (from *K. oxytoca*)	*B. subtilis*	This Study
MG5289	*amyE::* P_{const} P1P2$_{nasF}$- YFP, thrC:: P_{liaI} EutV/EutW	*B. subtilis*	This Study
MG5290	*amyE::* P_{const} P1P2-Term-YFP with C1U4 in P1 loop and C1G4 in P2 loop, *thrC::* P_{liaI} EutV/EutW	*B. subtilis*	This Study
MG5292	His_{10} MBP NasR (From K. oxytoca) in pVL847CbR	*E. coli*	This Study
MG5293	*amyE::* P_{const} P1P2-Term-YFP,	*B. subtilis*	This Study

	thrC::PLiaI-EutV/W, empty vector pDG148		
MG5295	GlmS*P3P4 with SST (GlmS is M9 mutant) in pDG148	*E. coli*	This Study
MG5296	GlmS*P3P4 with SST (GlmS is M9 mutant) into NEB PCR cloning vector	*E. coli*	This Study
MG5297	EutX w/ SST and CCC->GGG mutant in Riboswitch pseudoknot in pDG148	*E. coli*	This Study
MG5298	GlmS-P3P4 with SST (GlmS is WT) in pDG148	*E. coli*	This Study
MG5299	*thrC::* P_{liaI} EutV R142K /EutW	*E. coli*	This Study
MG5300	*thrC::* P_{liaI} EutV K143R /EutW	*E. coli*	This Study
MG5301	thrC:: P_{liaI} EutV K147R /EutW	*E. coli*	This Study
MG5302	*thrC::* P_{liaI} EutV K149R /EutW	*E. coli*	This Study
MG5303	*thrC::* P_{liaI} EutV R168K /EutW	*E. coli*	This Study
MG5304	*amyE::* P_{const} P1P2$_{nasF}$- YFP	*B. subtilis*	This Study
MG5305	*amyE::* P_{const} P1P2$_{nasF}$- YFP, thrC::P_{liaI}-NasR (from *K. oxytoca)*	*B. subtilis*	This Study
MG5306	*amyE::* P_{const} P1P2-Term-YFP, *thrC::* P_{liaI} EutV/W, EutX (CCC->GGG mutant) with SST in pDG148 (self-replicating plasmid)	*B. subtilis*	This Study
MG5307	*amyE::* P_{const} P1P2-Term-YFP, *thrC::* P_{liaI} EutV R142K /EutW	*B. subtilis*	This Study
MG5308	*amyE::* P_{const} P1P2-Term-YFP, *thrC::* P_{liaI} EutV K143R /EutW	*B. subtilis*	This Study
MG5309	*amyE::* P_{const} P1P2-Term-YFP, *thrC::* P_{liaI} EutV K147R /EutW	*B. subtilis*	This Study
MG5310	amyE:: P_{const} P1P2-Term-YFP, thrC:: PliaI EutV K149R /EutW	*B. subtilis*	This Study
MG5311	*amyE::* P_{const} P1P2-Term-YFP, *thrC::* P_{liaI} EutV R168K /EutW	*B. subtilis*	This Study
MG5312	*amyE::* P_{const} P1P2-Term-YFP, *thrC::* P_{liaI} EutV/W, GlmS*P3P4 with SST (GlmS is M9 mutant) in pDG148	*B. subtilis*	This Study
MG5314	*thrC::* P_{liaI} NasR (from *K. oxytoca*)	*B. subtilis*	This Study
MG5321	His_{10} MBP NasR R340A in pVL847CbR	*E. coli*	This Study
MG5322	His_{10} MBP NasR R340K in	*E. coli*	This Study

	pVL847CbR		
MG5335	*amyE::*P_{eutS}-P7P8T-YFP	*E. coli*	This Study
MG5336	BKE29740 (delta-ccpA) pDR224	*B. subtilis*	This Study
MG5347	*amyE::*P_{eutS}-P7P8T-YFP, *thrC::*P_{LiaI}-EutV/W	*B. subtilis*	This Study
MG5349	delta-ccpA (BKE29740), *amyE::*P_{eutS}-P7P8T-YFP, *thrC::*P_{liaI}-EutV/W	*B. subtilis*	This Study

Bibliography

1. Browning DF, Busby SJ (2004) The regulation of bacterial transcription initiation. *Nat Rev Microbiol* 2:57–65.
2. Waters LS, Storz G (2009) Regulatory RNAs in Bacteria. *Cell* 136(4):615–628.
3. Breaker RR (2012) Riboswitches and the RNA world. *Cold Spring Harb Perspect Biol* 4(2).
4. Babitzke P, Romeo T (2007) CsrB sRNA family: sequestration of RNA-binding regulatory proteins. *Curr Opin Microbiol* 10(2):156–163.
5. Görke B, Vogel J (2008) Noncoding RNA control of the making and breaking of sugars. *Genes Dev* 22(21):2914–2925.
6. Wassarman KM (2012) 6S RNA: A regulator of transcription. *Regulatory RNAs in Prokaryotes*, pp 109–130.
7. Coppins RL, Hall KB, Groisman EA (2007) The intricate world of riboswitches. *Curr Opin Microbiol* 10(2):176–181.
8. Henkin TM (2009) RNA-dependent RNA switches in bacteria. *Methods Mol Biol* 540:207–214.
9. Hershberg R, Altuvia S, Margalit H (2003) A survey of small RNA-encoding genes in *Escherichia coli*. *Nucleic Acids Res* 31(7):1813–1820.
10. Busch A, Richter AS, Backofen R (2008) IntaRNA: Efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* 24(24):2849–2856.
11. Storz G, Vogel J, Wassarman KM (2011) Regulation by small RNAs in

bacteria: expanding frontiers. *Mol Cell* 43(6):880–891.

12. Gottesman S, Storz G (2011) Bacterial small RNA regulators: Versatile roles and rapidly evolving variations. *Cold Spring Harb Perspect Biol* 3(12).

13. Bartel DP (2009) MicroRNAs: Target Recognition and Regulatory Functions. *Cell* 136(2):215–233.

14. Boisset S, et al. (2007) *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev* 21(11):1353–1366.

15. Heidrich N, Chinali A, Gerth U, Brantl S (2006) The small untranslated RNA SR1 from the *Bacillus subtilis* genome is involved in the regulation of arginine catabolism. *Mol Microbiol* 62(2):520–536.

16. Heidrich N, Moll I, Brantl S (2007) *In vitro* analysis of the interaction between the small RNA SR1 and its primary target *ahrC* mRNA. *Nucleic Acids Res* 35(13):4331–4346.

17. Geissmann T, et al. (2009) A search for small noncoding RNAs in *Staphylococcus aureus* reveals a conserved sequence motif for regulation. *Nucleic Acids Res* 37(21):7239–7257.

18. Huntzinger E, et al. (2005) *Staphylococcus aureus* RNAIII and the endoribonuclease III coordinately regulate spa gene expression. *EMBO J* 24(4):824–35.

19. Licht A, Preis S, Brantl S (2005) Implication of CcpN in the regulation of a novel untranslated RNA (SR1) in *Bacillus subtilis*. *Mol Microbiol* 58(1):189–206.

20. Morfeldt E, Taylor D, von Gabain A, Arvidson S (1995) Activation of alpha-toxin translation in *Staphylococcus aureus* by the trans-encoded antisense RNA, RNAIII. *EMBO J* 14(18):4569–4577.

21. Novick RP, et al. (1993) Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J* 12(10):3967–75.

22. Wagner EGH, Altuvia S, Romby P (2002) Antisense RNAs in bacteria and their genetic elements. *Adv Genet* 46:361–398.

23. Brantl S (2009) Bacterial chromosome-encoded small regulatory RNAs. *Futur Microbiol* 4(1):85–103.

24. Weaver KE (2007) Emerging plasmid-encoded antisense RNA regulated systems. *Curr Opin Microbiol* 10(2):110–116.

25. Brantl S (2007) Regulatory mechanisms employed by *cis*-encoded antisense RNAs. *Curr Opin Microbiol* 10(2):102–109.

26. Gerdes K, Wagner EGH (2007) RNA antitoxins. *Curr Opin Microbiol* 10(2):117–124.

27. Fozo EM, et al. (2008) Repression of small toxic protein synthesis by the Sib and OhsC small RNAs. *Mol Microbiol* 70(5):1076–1093.

28. Repoila F, Darfeuille F (2009) Small regulatory non-coding RNAs in bacteria: physiology and mechanistic aspects. *Biol Cell* 101(2):117–31.

29. Massé E, Salvail H, Desnoyers G, Arguin M (2007) Small RNAs controlling iron metabolism. *Curr Opin Microbiol* 10(2):140–145.

30. Papenfort K, Vogel J (2009) Multiple target regulation by small noncoding RNAs rewires gene expression at the post-transcriptional level. *Res Microbiol*

160(4):278–287.

31. Guillier M, Gottesman S, Storz G (2006) Modulating the outer membrane with small RNAs. *Genes Dev* 20(17):2338–2348.

32. Valentin-Hansen P, Johansen J, Rasmussen AA (2007) Small RNAs controlling outer membrane porins. *Curr Opin Microbiol* 10(2):152–155.

33. Toledo-Arana A, Repoila F, Cossart P (2007) Small noncoding RNAs controlling pathogenesis. *Curr Opin Microbiol* 10(2):182–188.

34. Vogel J, Group RNAB, Biology I (2009) A rough guide to the non-coding RNA world of *Salmonella*. *Mol Microbiol* 71(1):1–11.

35. Romby P, Charpentier E (2010) An overview of RNAs with regulatory functions in Gram-positive bacteria. *Cell Mol Life Sci* 67(2):217–237.

36. Babitzke P, Baker CS, Romeo T (2009) Regulation of translation initiation by RNA binding proteins. *Annu Rev Microbiol* 63:27–44.

37. Kugel JF, Goodrich JA (2007) An RNA transcriptional regulator templates its own regulatory RNA. *Nat Chem Biol* 3(2):89–90.

38. Trotochaud AE, Wassarman KM (2005) A highly conserved 6S RNA structure is required for regulation of transcription. *Nat Struct Mol Biol* 12(4):313–9.

39. Barrick JE, Sudarsan N, Weinberg Z, Ruzzo WL, Breaker RR (2005) 6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter. *RNA* 11(5):774–84.

40. Møller T, et al. (2002) Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol Cell* 9(1):23–30.

41. Brennan RG, Link TM (2007) Hfq structure, function and ligand binding. *Curr*

Opin Microbiol 10(2):125–133.

42. Panja S, Woodson SA (2012) Hfq proximity and orientation controls RNA annealing. *Nucleic Acids Res* 40(17):8690–8697.

43. Panja S, Woodson SA (2012) Hexamer to monomer equilibrium of *E. coli* Hfq in solution and its impact on RNA annealing. *J Mol Biol* 417(5):406–412.

44. Jousselin A, Metzinger L, Felden B (2009) On the facultative requirement of the bacterial RNA chaperone, Hfq. *Trends Microbiol* 17(9):399–405.

45. Santiago-Frangos A, Kavita K, Schu DJ, Gottesman S, Woodson SA (2016) C-terminal domain of the RNA chaperone Hfq drives sRNA competition and release of target RNA. *Proc Natl Acad Sci U S A* 113(41):E6089–E6096.

46. Massé E, Majdalani N, Gottesman S (2003) Regulatory roles for small RNAs in bacteria. *Curr Opin Microbiol* 6(2):120–124.

47. Bohn C, Rigoulay C, Bouloc P (2007) No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*. *BMC Microbiol* 7:10.

48. Dambach M, Irnov I, Winkler WC (2013) Association of RNAs with *Bacillus subtilis* Hfq. *PLoS One* 8(2).

49. Silvaggi JM, Perkins JB, Losick R (2005) Small untranslated RNA antitoxin in *Bacillus subtilis*. *J Bacteriol* 187(19):6641–6650.

50. Gaballa A, et al. (2008) The *Bacillus subtilis* iron-sparing response is mediated by a Fur-regulated small RNA and three small, basic proteins. *Proc Natl Acad Sci U S A* 105(33):11927–32.

51. Nielsen JS, et al. (2009) Defining a role for Hfq in Gram-positive bacteria: Evidence for Hfq-dependent antisense regulation in *Listeria monocytogenes*.

Nucleic Acids Res 38(3):907–919.

52. Irnov, Kertsburg A, Winkler WC (2006) Genetic Control by *cis*-Acting Regulatory RNAs in *Bacillus subtilis*: General Principles and Prospects for Discovery. *Cold Spring Harb Symp Quant Biol* 71:239–249.

53. Toledo-Arana A, et al. (2009) The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* 459(7249):950–956.

54. Barrick JE, et al. (2004) New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc Natl Acad Sci U S A* 101(17):6421–6.

55. Weinberg Z, et al. (2007) Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. *Nucleic Acids Res* 35(14):4809–4819.

56. Weinberg Z, et al. (2010) Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol* 11(3):R31.

57. Winkler WC, Breaker RR (2005) Regulation of bacterial gene expression by riboswitches. *Annu Rev Microbiol* 59(1):487–517.

58. Serganov A, Nudler E (2014) A Decade of Riboswitches. *Cell* 152(1):17–24.

59. Dann CE (2011) RNA structure: Riboswitch strikes a chord. *Nat Chem Biol* 7(10):660–1.

60. Grundy FJ, Winkler WC, Henkin TM (2002) tRNA-mediated transcription antitermination *in vitro*: codon-anticodon pairing independent of the ribosome. *Proc Natl Acad Sci U S A* 99(17):11121–11126.

61. Wakeman CA, Ramesh A, Winkler WC (2009) Multiple Metal-Binding Cores Are Required for Metalloregulation by M-box Riboswitch RNAs. *J Mol Biol* 392(3):723–735.

62. Winkler WC, Breaker RR (2003) Genetic control by metabolite-binding riboswitches. *ChemBioChem* 4(10):1024–1032.

63. Mellin JR, Cossart P (2015) Unexpected versatility in bacterial riboswitches. *Trends Genet* 31(3):150–156.

64. Caron M-P, et al. (2012) Dual-acting riboswitch control of translation initiation and mRNA decay. *Proc Natl Acad Sci U S A* 109(50):E3444-53.

65. Winkler WC, Nahvi A, Roth A, Collins J a, Breaker RR (2004) Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* 428(6980):281–286.

66. Roth A, Breaker RR (2009) The structural and functional diversity of metabolite-binding riboswitches. *Annu Rev Biochem* 78:305–34.

67. Gollnick P, Babitzke P (2002) Transcription attenuation. *Biochim Biophys Acta - Gene Struct Expr* 1577(2):240–250.

68. Wickiser JK, Winkler WC, Breaker RR, Crothers DM (2005) The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. *Mol Cell* 18(1):49–60.

69. Delfosse V, et al. (2009) Riboswitch structure: An internal residue mimicking the purine ligand. *Nucleic Acids Res* 38(6):2057–2068.

70. Wickiser JK, Cheah MT, Breaker RR, Crothers DM (2005) The kinetics of ligand binding by an adenine-sensing riboswitch. *Biochemistry* 44(40):13404–

13414.

71. Winkler WC (2005) Metabolic monitoring by bacterial mRNAs. *Arch Microbiol* 183(3):151–159.

72. Garsin DA (2010) Ethanolamine utilization in bacterial pathogens: roles and regulation. *Nat Rev Microbiol* 8(Box 1):290–295.

73. Garsin DA (2012) Ethanolamine: A signal to commence a host-associated lifestyle? *MBio* 3(4). doi:10.1128/mBio.00172-12.

74. Anderson CJ, Clark DE, Adli M, Kendall MM (2015) Ethanolamine Signaling Promotes *Salmonella* Niche Recognition and Adaptation during Infection. *PLoS Pathog* 11(11). doi:10.1371/journal.ppat.1005278.

75. Jett BD, Huycke MM, Gilmore MS (1994) Virulence of Enterococci. *Clin Microbiol Rev* 7(4):462–478.

76. Randle CL, Albro PW, Dittmer JC (1969) The phosphoglyceride composition of Gram-negative bacteria and the changes in composition during growth. *Biochim Biophys Acta - Lipids Lipid Metab* 187(2):214–220.

77. Roof DM, Roth JR (1988) Ethanolamine utilization in *Salmonella typhimurium*. *J Bacteriol* 170(9):3855–3863.

78. Kerfeld CA, Heinhorst S, Cannon GC (2010) Bacterial microcompartments. *Annu Rev Microbiol* 64:391–408.

79. Axen SD, Erbilgin O, Kerfeld CA (2014) A Taxonomy of Bacterial Microcompartment Loci Constructed by a Novel Scoring Method. *PLoS Comput Biol* 10(10):e1003898.

80. Roof DM, Roth JR (1989) Functions required for vitamin B_{12}-dependent

ethanolamine utilization in *Salmonella typhimurium*. *J Bacteriol* 171(6):3316–3323.

81. Stojiljkovic I, Baumler AJ, Heffron F (1995) Ethanolamine utilization in *Salmonella typhimurium*: Nucleotide sequence, protein expression, and mutational analysis of the *cchA cchB eutE eutJ eutG eutH* gene cluster. *J Bacteriol* 177(5):1357–1366.

82. Kofoid E, Rappleye C, Stojiljkovic I, Roth J (1999) The 17-gene ethanolamine (eut) operon of *Salmonella typhimurium* encodes five homologues of carboxysome shell proteins. *J Bacteriol* 181(17):5317–5329.

83. Roof DM, Roth JR (1992) Autogenous regulation of ethanolamine utilization by a transcriptional activator of the eut operon in *Salmonella typhimurium*. *J Bacteriol* 174(20):6634–6643.

84. Sheppard DE, Roth JR (1994) A rationale for autoinduction of a transcriptional activator: Ethanolamine ammonia-lyase (EutBC) and the operon activator (EutR) compete for adenosyl- cobalamin in *Salmonella typhimurium*. *J Bacteriol* 176:1287–1296.

85. Luzader DH, Clark DE, Gonyar LA, Kendall MM (2013) EutR is a direct regulator of genes that contribute to metabolism and virulence in enterohemorrhagic *Escherichia col*i o157: H7. *J Bacteriol* 195(21):4947–4953.

86. Brinsmade SR, Paldon T, Escalante-Semerena JC (2005) Minimal functions and physiological conditions required for growth of *Salmonella enterica* on ethanolamine in the absence of the metabolosome. *J Bacteriol* 187(23):8039–8046.

87. Penrod JT, Roth JR (2006) Conserving a volatile metabolite: A role for carboxysome-like organelles in *Salmonella enterica*. *J Bacteriol* 188(8):2865–2874.

88. Kinney JN, Axen SD, Kerfeld CA (2011) Comparative analysis of carboxysome shell proteins. *Photosynthesis Research*, pp 21–32.

89. Buckel W, Golding BT (2006) Radical enzymes in anaerobes. *Annu Rev Microbiol* 60:27–49.

90. O'Brien JR, et al. (2004) Insight into the Mechanism of the B_{12}-Independent Glycerol Dehydratase from *Clostridium butyricum*: Preliminary Biochemical and Structural Characterization. *Biochemistry* 43:4635–4645.

91. Scott KP, Martin JC, Campbell G, Mayer CD, Flint HJ (2006) Whole-genome transcription profiling reveals genes up-regulated by growth on fucose in the human gut bacterium *Roseburia inulinivorans*. *J Bacteriol* 188:4340–4349.

92. Wagner AF, Frey M, Neugebauer FA, Schäfer W, Knappe J (1992) The free radical in pyruvate formate-lyase is located on glycine-734. *Proc Natl Acad Sci U S A* 89:996–1000.

93. Berg IA, et al. (2010) Autotrophic carbon fixation in archaea. *Nat Rev Microbiol* 8(6):447–60.

94. Thauer RK (2007) Microbiology. A fifth pathway of carbon fixation. *Science (80-)* 318(5857):1732–1733.

95. Abdul-Rahman F, Petit E, Blanchard JL (2013) The Distribution of Polyhedral Bacterial Microcompartments Suggests Frequent Horizontal Transfer and Operon Reassembly. *Phylogenetics Evol Biol* 1:1000118.

96. Kerfeld CA, Melnicki MR (2016) Assembly, function and evolution of cyanobacterial carboxysomes. *Curr Opin Plant Biol* 31:66–75.

97. Saier MH (2013) Microcompartments and protein machines in prokaryotes. *J Mol Microbiol Biotechnol* 23:243–269.

98. Cheng S, Liu Y, Crowley CS, Yeates TO, Bobik TA (2008) Bacterial microcompartments: Their properties and paradoxes. *BioEssays* 30(11–12):1084–1095.

99. Tsai SJ, Yeates TO (2011) Bacterial microcompartments: Insights into the structure, mechanism, and engineering applications. *Prog Mol Biol Transl Sci* 103:1–20.

100. Huseby DL, Roth JR (2013) Evidence that a metabolic microcompartment contains and recycles private cofactor pools. *J Bacteriol* 195:2864–2879.

101. Sheppard DE, Penrod JT, Bobik T, Kofoid E, Roth JR (2004) Evidence that a B_{12}-adenosyl transferase is encoded within the ethanolamine operon of *Salmonella enterica*. *J Bacteriol* 186(22):7635–7644.

102. Buan NR, Suh SJ, Escalante-Semerena JC (2004) The *eutT* gene of *Salmonella enterica* encodes an oxygen-labile, metal-containing ATP: Corrinoid adenosyltransferase enzyme. *J Bacteriol* 186(17):5708–5714.

103. Tsoy O, Ravcheev D, Mushegian A (2009) Comparative genomics of ethanolamine utilization. *J Bacteriol* 191(23):7157–7164.

104. Laub MT, Goulian M (2007) Specificity in two-component signal transduction pathways. *Annu Rev Genet* 41:121–145.

105. Galperin MY (2006) Structural classification of bacterial response regulators:

Diversity of output domains and domain combinations. *J Bacteriol* 188(12):4169–4182.

106. Shu CJ, Zhulin IB (2002) ANTAR: An RNA-binding domain in transcription antitermination regulatory proteins. *Trends Biochem Sci* 27(1):3–5.

107. Chai W, Stewart V (1998) NasR, a novel RNA-binding protein, mediates nitrate-responsive transcription antitermination of the *Klebsiella oxytoca* M5al *nasF* operon leader *in vitro*. *J Mol Biol* 283(2):339–351.

108. Wilson S a, Wachira SJ, Norman R a, Pearl LH, Drew RE (1996) Transcription antitermination regulation of the *Pseudomonas aeruginosa* amidase operon. *EMBO J* 15(21):5907–5916.

109. Chai W, Stewart V (1999) RNA sequence requirements for NasR-mediated, nitrate-responsive transcription antitermination of the *Klebsiella oxytoca* M5al *nasF* operon leader. *J Mol Biol* 292(2):203–216.

110. Del Papa MF, Perego M (2011) *Enterococcus faecalis* virulence regulator FsrA binding to target promoters. *J Bacteriol* 193(7):1527–1532.

111. Del Papa MF, Perego M (2008) Ethanolamine activates a sensor histidine kinase regulating its utilization in *Enterococcus faecalis*. *J Bacteriol* 190(21):7147–7156.

112. Baker KA, Perego M (2011) Transcription antitermination by a phosphorylated response regulator and cobalamin-dependent termination at a B_{12} riboswitch contribute to ethanolamine utilization in *Enterococcus faecalis*. *J Bacteriol* 193(10):2575–86.

113. Fox KA, et al. (2009) Multiple posttranscriptional regulatory mechanisms

partner to control ethanolamine utilization in *Enterococcus faecalis*. *Proc Natl Acad Sci*.

114. Ramesh A, et al. (2012) The mechanism for RNA recognition by ANTAR regulators of gene expression. *PLoS Genet* 8(6).

115. Chai W, Stewart V (1998) NasR, a novel RNA-binding protein, mediates nitrate-responsive transcription antitermination of the *Klebsiella oxytoca* m5al *nasF* operon leader *in vitro*. *J Mol Biol* 283(2):339–351.

116. Lin JT, Stewart V (1996) Nitrate and nitrite-mediated transcription antitermination control of *nasF* (nitrate assimilation) operon expression in *Klebsiella pheumoniae* M5al. *J Mol Biol* 256(3):423–35.

117. Hollands K, et al. (2012) Riboswitch control of Rho-dependent transcription termination. *Proc Natl Acad Sci U S A* 109(14):5376–5381.

118. Fröhlich KS, Vogel J (2009) Activation of gene expression by small RNA. *Curr Opin Microbiol* 12(6):674–682.

119. Altuvia S (2007) Identification of bacterial small non-coding RNAs: experimental approaches. *Curr Opin Microbiol* 10(3):257–261.

120. Sonnleitner E, Romeo A, Bläsi U (2012) Small regulatory RNAs in *Pseudomonas aeruginosa*. *RNA Biol* 9(4):364–371.

121. Blomberg P, Nordström K, Wagner EG (1992) Replication control of plasmid R1: RepA synthesis is regulated by CopA RNA through inhibition of leader peptide translation. *EMBO J* 11(7):2675–2683.

122. Loh E, et al. (2009) A trans-Acting Riboswitch Controls Expression of the Virulence Regulator PrfA in *Listeria monocytogenes*. *Cell* 139(4):770–779.

123. Johnson Jr JE, Reyes FE, PolaJohnson Jr, J. E., Reyes, F. E., Polaski, J. T., & Batey, R. T. (2012). B12 cofactors directly stabilize an mRNA regulatory switch. *Nature*, 492(7427), 133–137.

124. Nahvi A, Barrick JE, Breaker RR (2004) Coenzyme B_{12} riboswitches are widespread genetic control elements in prokaryotes. *Nucleic Acids Res* 32(1):143–150.

125. Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR (2003) Rfam: An RNA family database. *Nucleic Acids Res* 31(1):439–441.

126. Nahvi A, et al. (2002) Genetic control by a metabolite binding mRNA. *Chem Biol* 9(9):1043–1049.

127. Nahvi A, Green R (2013) Structural analysis of RNA backbone using in-line probing. *Methods Enzymol* 530:381–397.

128. Landick R, Wang D, Chan CL (1996) Quantitative analysis of transcriptional pausing by *Escherichia coli* RNA polymerase: his leader pause site as paradigm, pp 334–353.

129. Landick R, Wang D, Chan CL (1996) Quantitative analysis of transcriptional pausing by *Escherichia coli* RNA polymerase: his leader pause site as paradigm. *Methods Enzymol* 274:334–353.

130. Roelofs KG, Wang J, Sintim HO, Lee VT (2011) Differential radial capillary action of ligand assay for high-throughput detection of protein-metabolite interactions. *Proc Natl Acad Sci* 108(37):15528–15533.

131. Mellin JR, et al. (2014) Sequestration of a two-component response regulator by a riboswitch-regulated noncoding RNA. *Science* 345:940–943.

132. DebRoy S, et al. (2014) A riboswitch-containing sRNA controls gene expression by sequestration of a response regulator. *Science* 345(6199):937–940.

133. West AH, Stock AM (2001) Histidine kinases and response regulator proteins in two-component signaling systems. *Trends Biochem Sci* 26(6):369–376.

134. Gao R, Stock AM (2009) Biological insights from structures of two-component proteins. *Annu Rev Microbiol* 63:133–54.

135. Boudes M, et al. (2012) The structure of the NasR transcription antiterminator reveals a one-component system with a NIT nitrate receptor coupled to an ANTAR RNA-binding effector. *Mol Microbiol* 85(3):431–444.

136. Stewart V, van Tilbeurgh H (2012) Found: The elusive ANTAR transcription antiterminator. *PLoS Genet* 8(6). 18(19):5175–5186.

138. Toymentseva A a, Schrecke K, Sharipova MR, Mascher T (2012) The LIKE system, a novel protein expression toolbox for *Bacillus subtilis* based on the *liaI* promoter. *Microb Cell Fact* 11(1):143.

139. Butcher BG, Lin YP, Helmann JD (2007) The *yydFGHIJ* operon of *Bacillus subtilis* encodes a peptide that induces the LiaRS two-component system. *J Bacteriol* 189(23):8616–8625.

140. Wolf D, et al. (2010) In-depth profiling of the LiaR response of *Bacillus subtilis*. *J Bacteriol* 192(18):4680–4693.

141. Makita Y, Nakao M, Ogasawara N, Nakai K (2004) DBTBS: database of transcriptional regulation in *Bacillus subtilis* and its contribution to comparative genomics. *Nucleic Acids Res* 32(suppl 1):D75–D77.

142. Yoshida KI, et al. (2000) Systematic study of gene expression and transcription organization in the *gntZ-ywaA* region of the *Bacillus subtilis* genome. *Microbiology* 146(3):573–579.

143. McKenzie T, Hoshino T, Tanaka T, Sueoka N (1986) The nucleotide sequence of pUB110: Some salient features in relation to replication and its regulation. *Plasmid* 15(2):93–103.

144. Stragier P, Bonamy C, Karmazyn-Campelli C (1988) Processing of a sporulation sigma factor in *Bacillus subtilis*: how morphological structure could control gene expression. *Cell* 52(5):697–704.

145. Smith K, Youngman P (1992) Use of a new integrational vector to investigate compartment-specific expression of the *Bacillus subtilis spollM* gene. *Biochimie* 74(7–8):705–711.

146. Paintdakhi A, et al. (2016) Oufti: An integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. *Mol Microbiol* 99(4):767–777.

147. Chen TJ, Kotecha N (2014) Cytobank: Providing an analytics platform for community cytometry data analysis and collaboration. *Curr Top Microbiol Immunol* 377:127–157.

148. Yang J, et al. (2015) The I-TASSER Suite: protein structure and function prediction. *Nat Meth* 12(1):7–8.

149. Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 5(4):725–738.

150. Pettersen EF, et al. (2004) UCSF Chimera—A Visualization System for

Exploratory Research and Analysis. *J Comput Chem* 25:1605–1612.

151. Morth JP, Feng V, Perry LJ, Svergun DI, Tucker PA (2004) The crystal and solution structure of a putative transcriptional antiterminator from *Mycobacterium tuberculosis*. *Structure* 12(9):1595–1605.

152. Wu Q, Stewart V (1998) NasFED proteins mediate assimilatory nitrate and nitrite transport in *Klebsiella oxytoca (pneumoniae)* M5al. *J Bacteriol* 180(5):1311–1322.

153. Albright FR, White DA, Lennarz WJ (1973) Studies on Enzymes Involved in the Catabolism of Phospholipids in Escherichia coli. *J Biol Chem* 248(11):3968–3977.

154. Scarlett FA, Turner JM (1976) Ethanolamine catabolism mediated by coenzyme B_{12}-dependent ethanolamine ammonia-lyase in *Escherichia coli* and *Klebsiella aerogenes*.

155. Jones PW, Turner JM, A- C (1984) Interrelationships between the enzymes of ethanolamine metabolism in *Escherichia coli*. *J Gen Microbiol* 130(2):299–308.

156. Roof DM, Roth JR (1988) Ethanolamine utilization in *Salmonella typhimurium*. *J Bacteriol* 170(9):3855–63.

157. Roof DM, Roth JR (1992) Autogenous regulation of ethanolamine utilization by a transcriptional activator of the *eut* operon in *Salmonella typhimurium*. *J Bacteriol* 174(20):6634–43.

158. Gera K, McIver KS (2013) Laboratory growth and maintenance of *Streptococcus pyogenes* (The Group A Streptococcus, GAS). *Curr Protoc*

Microbiol (SUPPL.30).

159. Chowdhury C, Sinha S, Chun S, Yeates TO, Bobik TA (2014) Diverse bacterial microcompartment organelles. *Microbiol Mol Biol Rev* 78(3):438–68.

160. Cameron JC, Wilson SC, Bernstein SL, Kerfeld CA (2013) Biogenesis of a bacterial organelle: The carboxysome assembly pathway. *Cell* 155.

161. Moore SJ, Warren MJ (2012) The anaerobic biosynthesis of vitamin B_{12}. *Biochem Soc Trans* 40(3):581 LP-586.

162. Moore SJ, et al. (2013) Elucidation of the anaerobic pathway for the corrin component of cobalamin (vitamin B_{12}). *Proc Natl Acad Sci* 110(37):14906–14911.

163. Brückner R, Titgemeyer F (2002) Carbon catabolite repression in bacteria: Choice of the carbon source and autoregulatory limitation of sugar utilization. *FEMS Microbiol Lett* 209(2):141–148.

164. Egeter O, Brückner R (1996) Catabolite repression mediated by the catabolite control protein CcpA in *Staphylococcus xylosus*. *Mol Microbiol* 21(4):739–49.

165. Opsata M, Nes IF, Holo H (2010) Class IIa bacteriocin resistance in *Enterococcus faecalis* V583: the mannose PTS operon mediates global transcriptional responses. *BMC Microbiol* 10:224.

166. Miwa Y, Nakata a, Ogiwara a, Yamamoto M, Fujita Y (2000) Evaluation and characterization of catabolite-responsive elements (cre) of Bacillus subtilis. *Nucleic Acids Res* 28(5):1206–10.

167. Wakeman CA, Winkler WC, Dann CE (2007) Structural features of metabolite-sensing riboswitches. *Trends Biochem Sci* 32(9):415–424.

168. Sargent F, et al. (2013) A synthetic system for expression of components of a bacterial microcompartment. *Microbiol (United Kingdom)* 159:2427–2436.

169. Ramakers C, Ruijter JM, Lekanne Deprez RH, Moorman AFM (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett* 339(1):62–66.