CHARACTERIZATION OF THE S_{MK} BOX: A RIBOSWITCH THAT BINDS *S*-ADENOSYLMETHIONINE

DISSERTATION

Presented in Partial Fulfillment of the Requirements for

the Degree Doctor of Philosophy in the Graduate

School of The Ohio State University

By

Ryan T. Fuchs, B.S.

Graduate Program in Microbiology

The Ohio State University

2009

Dissertation Committee:

Professor Tina M. Henkin, Adviser

Professor Juan D. Alfonzo

Professor Charles J. Daniels

Professor Kurt Fredrick

ABSTRACT

Riboswitches are conserved RNA sequences that regulate downstream gene expression via changes in the RNA structure. Riboswitches usually respond to an environmental stimulus, typically binding of an effector molecule, that causes a structural rearrangement without additional *trans*-acting factors. The structural rearrangement can cause or prevent the formation of an intrinsic transcriptional terminator to regulate the gene at the level of premature transcription termination, or it can modulate the accessibility of the ribosome binding site to regulate at the level of translation initiation. Previous work has shown that there are multiple classes of riboswitches that bind *S*-adenosylmethionine (SAM) and regulate genes involved in biosynthesis of methionine, cysteine and SAM. The S box (or SAM-I) riboswitches are found in a wide range of bacteria while the SAM-II riboswitch is found primarily in Proteobacteria. Both riboswitches recognize SAM in a similar manner, but they are very different in terms of sequence, secondary and tertiary structure. This demonstrates that it is possible for bacteria to evolve very different regulatory RNAs to respond to the same metabolite.

We discovered a third SAM binding riboswitch motif (called the S_{MK} box or SAM-III) upstream of the *metK* gene (encoding SAM synthetase) in many Lactobacillales species. The S_{MK} box RNA was shown to bind SAM and discriminate against the closely related analog *S*-adenosylhomocysteine (SAH). Through enzymatic probing, it was determined that SAM causes a structural rearrangement in the RNA that

results in sequestration of part of the Shine-Dalgarno (SD) region. This observation was confirmed by X-ray crystallography, which also showed that part of the SD directly contacts SAM. A translational fusion of the *Enterococcus faecalis metK* leader to *lacZ* was made and introduced into *Bacillus subtilis*. When the *B. subtilis* cells were grown under conditions in which SAM pools are elevated, β-galactosidase activity decreased. In contrast, a transcriptional fusion showed no effect. This suggested that the S_{MK} box riboswitch down-regulates gene expression at the level of translation in response to SAM. This observation was supported by the fact that SAM inhibits binding of *Escherichia coli* 30S subunits to S_{MK} box RNA *in vitro*. Mutant S_{MK} box sequences that are deficient in SAM binding *in vitro* showed no SAM dependent effect on 30S subunit binding and no regulation of *lacZ* expression when SAM pools were modulated.

Reverse transcription and quantitative PCR techniques were utilized in order to analyze properties of the *metK* transcript *in vivo* in a native organism, *E. faecalis*. The abundance of *metK* transcript in *E. faecalis* cells was found to be unchanged during growth under conditions resulting in high or low SAM pools. This supported the model that regulation of *metK* by the S_{MK} box is not at the level of transcription. The half-life ($t_{1/2}$) of the *metK* transcript was determined to be ~3 min regardless of SAM pools. In contrast, the $t_{1/2}$ of the interaction between SAM and the S_{MK} box *in vitro* is 7.8 sec (A. Smith, unpublished results). Thus, the interaction of the S_{MK} box with its effector was shown to be of much shorter duration than the persistence of the transcript in the cell, suggesting that the regulatory effect and structural changes resulting from SAM binding are reversible.

Attempts were made to demonstrate S_{MK} box-mediated regulation in *E. faecalis* by utilizing a translational *metK-gusA* fusion and direct measurement of the activity of the *metK* gene product, SAM synthetase. Neither method showed regulation of gene expression due to SAM. The basis for these observations is unknown. It is possible that there are other factors involved in regulation of the *metK* gene in *E. faecalis* that are masking the effect of the S_{MK} box *in vivo*.

ACKNOWLEDGMENTS

I wish to express my thanks to my advisor, Dr. Tina Henkin, for her guidance, patience, encouragement, and support throughout the years. It has been a pleasure to work with such a gifted scientist and excellent mentor.

I would like to thank Dr. Frank Grundy for all his help and encouragement over the years and for identifying the S_{MK} box motif.

I would like to thank my committee members, Dr. Juan Alfonzo, Dr. Charles Daniels, and Dr. Kurt Fredrick for their helpful suggestions, support and time over the years.

I also wish to thank all of my colleagues in the Henkin lab for providing a helpful and pleasant working environment. It has been a privilege working with all of you.

I would also like to thank my family and friends for support and encouragement throughout my time at OSU.

VITA

September 9, 1980……...Born - Toledo, Ohio

June 2002…….B.S. Microbiology, The Ohio State University

September 2003 – present……...Graduate Research Associate, The Ohio State University

PUBLICATIONS

Research Publications

1. Fuchs, R.T., Grundy, F.G. and Henkin, T.M. 2006. The S_{MK} box is a new SAM-binding RNA for translational regulation of SAM synthetase. *Nat. Struct. Mol. Biol.* **13**: 226-33.

2. Fuchs, R.T., Grundy, F.G. and Henkin, T.M. 2007. *S*-adenosylmethionine directly inhibits binding of 30S ribosomal subunits to the S_{MK} box translational riboswitch RNA. *PNAS* **104**: 4876-80.

3. Lu, C., Smith, A.M., Fuchs, R.T., Ding, F., Rajashankar, K, Henkin, T.M. and Ke, A. 2008. Crystal structures of the SAM-III/S_{MK} riboswitch reveal the SAM-dependent translation inhibition mechanism. *Nat. Struct. Mol. Biol.* **15**: 1076-83.

4. Fuchs, R.T., Smith, A.M., Grundy, F.J., and Henkin, T.M. 2009. The *Enterococcus faecalis* S_{MK} box is a reversible riboswitch. (In preparation)

FIELDS OF STUDY

Major Field: Microbiology

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBRIEVIATIONS

AEC	aminoethylcysteine
AdoCbl	adenosylcobalamin
2-AP	2-aminopurine
ATP	adenosine triphosphate
FMN	flavin mononucleotide
GlcN6P	glucosamine-6-phosphate
NAIM	nucleotide analog interference mapping
NMR	nuclear magnetic resonance
nt	nucleotide
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
RBS	ribosome binding site
RNase	ribonuclease
RNAP	RNA polymerase
SAH	*S*-adenosylhomocysteine
SAM	*S*-adenosylmethionine
SD	Shine-Dalgarno

TBAB	tryptose blood agar base
TPP	thiamin pyrophosphate
Tris-HCl	tris-(hydroxylmethyl) aminomethane hydrochloride
UTR	untranslated region
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside

CHAPTER 1

REGULATION OF GENE EXPRESSION BY RIBOSWITCHES

Regulation of gene expression is important for all organisms because in order for cells to operate efficiently and maintain viability, they must be able to alter gene expression based on changing environmental conditions. Recent discoveries have shown that gene regulation via riboswitches is a common regulatory mechanism in bacteria (7, 41, 71). A riboswitch is a conserved RNA sequence that responds to a physiological signal and controls gene expression in the absence of *trans*-acting factors. In bacteria, riboswitch sequences are typically located in the 5' untranslated region (5'-UTR, or leader region) of a gene. Riboswitches confer regulation at the level of transcription attenuation (premature transcription termination) via modulation of the formation of an intrinsic transcriptional terminator structure, or at the level of translation by affecting sequestration of the Shine-Dalgarno (SD) region. When a riboswitch encounters its cognate physiological signal (typically, binding of a small molecule, known as an effector molecule, to a region called the aptamer), a structural rearrangement occurs that either promotes or inhibits formation of the RNA structure (called the expression platform) that confers regulation. Each class of riboswitch demonstrates specificity for its cognate signal and discriminates between closely related compounds. This is crucial in order for

the riboswitch to respond to the correct environmental stimulus. Each individual riboswitch within a class can have minor sequence and structural differences that impact the stringency of regulation. This variation in stringency tends to correlate with the physiological function of the gene that the riboswitch regulates (87). It is also possible for riboswitches to operate in tandem to confer more finely tuned regulation (56, 85).

Although most known riboswitches have been discovered in bacteria, riboswitches have been found in all three domains of life (7, 105, 170). The riboswitches identified in eukaryotes appear to regulate gene expression by regulating mRNA splicing or transcript stability. For instance, in *Neurospora crassa* three riboswitches that bind thiamine pyrophosphate (TPP) have been identified that control mRNA splicing (15). At present, riboswitches in eukaryotes have been identified by their resemblance to bacterial sequences. It is likely that eukaryotic organisms possess a wide range of novel riboswitch classes that await discovery. Given that new riboswitch classes continue to be discovered in bacteria and discovery of eukaryotic and archaeal systems has just begun, it is possible that utilization of riboswitches for gene regulation may be much more common than is currently known.

1.1 Riboswitches that respond to tRNA, metal ions, and temperature

1.1.1 The T box riboswitch

The T box riboswitch system is found in many bacteria, including all groups of Gram-positive bacteria, and regulates a wide range of genes including aminoacyl-tRNA synthetase (aaRS) genes and genes involved in amino acid biosynthesis and transport (34). More than 1000 T box sequences have been identified in 87 organisms (122). The T box

mechanism typically regulates at the level of transcription termination, although some T box RNAs in Gram-negative bacteria and members of the Actinomycetes are predicted to regulate at the level of translation initiation. Initial studies showed that expression of T box leader RNAs fused to a *lacZ* reporter gene was induced upon limitation of a particular amino acid. This indicated that T box sequences are functional gene regulators that responded to a specific effector, which was proposed to be a cognate tRNA (144). Additional studies determined that the effector is actually uncharged tRNA (159). An increase in uncharged tRNA in the cell signals a need for increased synthesis and transport of amino acids or an increase in aaRS gene expression. Thus, it is logical that many of the genes involved in those processes are regulated by the T box mechanism.

Leader sequences that contain T box motifs typically have the potential to form either a transcriptional terminator structure or a competing antiterminator structure (144). The default state of most T box RNAs is to form the terminator and terminate transcription in the absence of the uncharged tRNA effector molecule (Fig. 1.1). Uncharged $tRNA^{Gly}$ was shown to promote readthrough of the *glyQS* leader from *B. subtilis* in an *in vitro* transcription assay in the absence of any additional factors, indicating that tRNA alone is capable of generating a regulatory response (31). It was hypothesized that the promotion of readthrough by tRNA is due to stabilization of the antiterminator structure in the presence of tRNA (144). Additional *in vitro* studies supported this hypothesis by demonstrating not only that $tRNA^{Gly}$ binds *glyQS* leader RNA, but also that a bulge, consisting of the region of highest conservation among T box leaders, in the *glyQS* antiterminator structure is protected from Mg^{2+}-dependent cleavage in the presence of $tRNA^{Gly}$ (103). This protection is mediated by pairing of the CCA end

of the tRNA to a 100% conserved UGG sequence in the antiterminator, which was initially predicted by genetic analysis (159). Charged tRNA is unable to make interactions between the CCA end and T box RNA because of steric hindrance from the amino acid, and thus it cannot promote antiterminator stabilization.

In addition to the CCA end-antiterminator interaction, it was also predicted that a region called the Specifier Sequence makes a pairing interaction with the anticodon of the tRNA to determine which tRNA species is monitored by the T box leader RNA (144). Initial genetic analysis supported this prediction and additional evidence for this pairing interaction was generated when Mg^{2+}-dependent cleavage experiments showed that the Specifier Sequence of *glyQS* RNA is protected from cleavage when $tRNA^{Gly}$ is present (103). Disruption of this interaction by mutagenesis results in loss of tRNA-dependent antitermination, indicating that both the Specifier Sequence-anticodon and CCA end-antiterminator bulge interactions are required for stable tRNA binding. A $tRNA^{Gly}$ variant that mimics charged $tRNA^{Gly}$ (by having an extra C residue at the 3' end) was shown to be able to make the Specifier Sequence-anticodon interaction despite the fact that it could no longer make the CCA-antiterminator bulge interaction (33). Additionally, the charged $tRNA^{Gly}$ mimic inhibits antitermination by uncharged $tRNA^{Gly}$ unless the complete antiterminator has been formed (33). A charged $tRNA^{Gly}$ mimic that contains a mutation to disrupt Specifier Sequence pairing loses its ability to inhibit antitermination by uncharged $tRNA^{Gly}$. Together these results suggest that charged tRNA can compete with uncharged tRNA for binding the Specifier Sequence and thus the fate of the T box transcript is tied to the ratio between charged and uncharged tRNA.

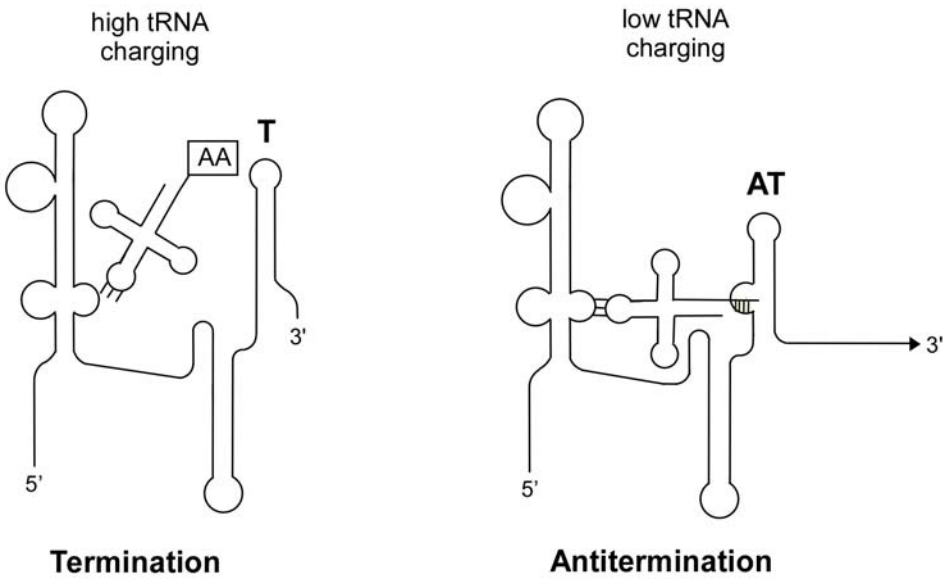

Figure 1.1. The T box riboswitch. When charged tRNA encounters T box RNA, the acceptor stem of the tRNA cannot make crucial interactions with the leader RNA. Thus, a transcriptional terminator forms and transcription terminates before the coding sequence is transcribed. When uncharged tRNA is present, its acceptor stem can bind the leader RNA and stabilize an antiterminator structure. This prevents formation of the terminator and the gene coding sequence is transcribed. T, terminator; AT, antiterminator; AA, amino acid. Figure modified from (41).

1.1.2 RNA thermosensors

Some of the simplest known riboswitches are RNA thermosensors. These RNAs respond to changing temperature rather than by binding a molecular signal. Typically, the default state of these RNAs is to be in a stem-loop conformation where the SD region is sequestered so that translation initiation is inhibited. An increase in temperature destabilizes the structure in the SD region and allows ribosome binding (Fig. 1.2).

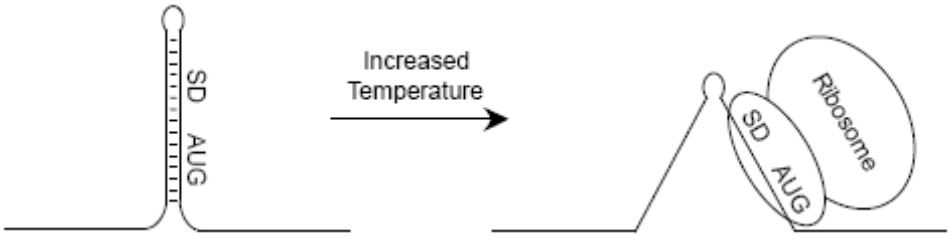

Figure 1.2. Example of a generic thermosensing riboswitch. At lower temperatures the SD region of the transcript pairs with a complementary sequence in the RNA and the SD is unavailable for translation initiation. At increased temperatures, this structure is destabilized, the ribosome can successfully compete for binding to the SD, the initiation complex can form and the mRNA can be translated. SD, Shine-Dalgarno sequence; AUG, start codon of the coding sequence.

This mechanism of regulation is utilized to increase expression of the *E. coli rpoH* gene, which encodes the heat shock sigma factor (σ^{32}), when temperature increases (108). It makes sense for the cell to upregulate this sigma factor based on temperature since σ^{32} directs transcription of heat shock genes (114). Mutational analysis of the *rpoH* leader demonstrated that the stability of a stem-loop containing the SD sequence correlates with the amount of expression of a *lacZ* fusion *in vivo* and the stringency of induction due to increased temperature (108). A ribosomal toeprint assay demonstrated that ribosomes assemble on the transcript at 42ºC, but not at 30ºC, which supports the model that the stem-loop sequestering the SD is destabilized as temperature increases (108).

Another thermosensing riboswitch element, named Repression of Heat-Shock Expression (ROSE), with a mode of action similar to the *rpoH* riboswitch, was discovered in multiple heat-shock genes in *Bradyrhizobium japonicum* (109). The ROSE element is a stem-loop conformation that contains a conserved G residue opposite the SD sequence. NMR experiments determined that this residue is involved in a weak G-G base pair that breaks as temperature increases; breakage of this pair is the first step

in destabilization of the stem-loop (109). Deletion of this G optimizes base pairing of the stem-loop and causes the structure of the RNA to be unresponsive to destabilization as temperature increases (115). This suggests that if a thermosensor contains sequence conservation, then the conservation is likely to be necessary to calibrate the response of the RNA sequence to the correct temperature range.

Thermosensor RNAs also have roles in the regulation of virulence genes. One example of this is the *prfA* gene in *Listeria monocytogenes,* which encodes a transcriptional activator that activates virulence genes and contains secondary structure in the SD region that prevents translation at 30ºC (107). Increasing the temperature to 37ºC, which occurs when the organism enters a potential mammalian host, causes the mRNA to be translated and the cells to become virulent. When the structure was destabilized by mutagenesis, *L. monocytogenes* cells were able to invade mammalian cells at 30ºC (107). Utilization of a thermosensor RNA has provided *L. monocytogenes* with a simple system to sense when it is appropriate to express virulence genes.

1.1.3 Magnesium-sensing RNAs

The first riboswitch that regulates gene expression in response to metals was discovered in the 5'-UTR of the *mgtA* gene of *Salmonella enterica* serovar Typhimurium (113). The *mgtA* gene encodes a protein that mediates Mg^{2+} influx and initiation of its transcription is regulated by an Mg^{2+}-responsive two component regulatory system (PhoP/PhoQ). Strains with a disrupted two component system retain regulation of *mgtA* expression in response to Mg^{2+}, which suggested that there is another level of regulation (139). *In vitro* transcription experiments showed that premature termination of the 5'-

UTR is promoted at a specific residue by high Mg^{2+} without the need for any additional factors (113). Additionally, the 5'-UTR RNA sequence demonstrated structural changes in response to Mg^{2+} and fusions of the 5'-UTR to *lacZ* demonstrated Mg^{2+}-dependent repression of *lacZ* expression. All of these results suggested that the 5'-UTR contains a riboswitch motif that regulates at the level of transcription termination. However, a transcription terminator structure could not be identified in the 5'-UTR. The 5'-UTR sequence does contain regions of primary and secondary structure that are conserved in the *mgtA* homologs in eight other Gram-negative bacteria species, although a terminator structure is also not evident in these species. As a result, the exact mechanism of transcription termination caused by this Mg^{2+} sensing RNA is not known.

A second class of Mg^{2+} riboswitch named the "M-box" is present in a ***B. subtilis*** Mg^{2+} influx gene called *mgtE* (112). The M-box shares no primary or secondary conservation with the *mgtA* riboswitch which suggests that these two RNAs evolved independently. M-box RNA was shown to undergo structural changes in response to Mg^{2+} and expression of a M-box-*lacZ* fusion was repressed under high Mg^{2+} conditions. Unlike the *mgtA* riboswitch, the M-box sequence has the potential to form a recognizable transcriptional terminator. These results suggest that the M-box element regulates gene expression in response to Mg^{2+} by stabilization of the transcriptional terminator structure.

The structure of M-box RNA bound to Mg^{2+} has been solved by X-ray crystallography (112). The M-box RNA has a rather large aptamer domain that forms three closely packed helices. Six Mg^{2+} atoms are located where the three helices converge, which indicates that riboswitch function may depend on cooperative binding of Mg^{2+}. This is an unusual feature of the M-box RNA because most riboswitches bind only

one molecule of their cognate signal. The significance of this binding arrangement has not been proven, although it has been proposed to allow the RNA to be more sensitive to small changes in Mg^{2+} concentration (112). The complexity of this binding arrangement probably accounts for why such a large aptamer region is utilized for recognition of such a small ligand.

1.2 The GlcN6P ribozyme

An RNA motif was recently discovered in the *glmS* gene, which encodes the enzyme glutamine-fructose-6-phosphate amidotransferase, of *B. subtilis* and other Gram-positive bacteria (101). This enzyme is responsible for the conversion of fructose-6-phosphate and glutamine into glucosamine-6-phosphate (GlcN6P). It was determined that the *glmS* leader RNA undergoes spontaneous cleavage to repress gene expression and that GlcN6P can increase the rate of this cleavage by as much as 1,000-fold (101). Additional studies showed that spontaneous cleavage in the absence of GlcN6P occurs only in Tris buffers and that GlcN6P enhances the rate of cleavage ~100,000-fold in a HEPES buffer (57). The cleavage due to Tris was attributed to a dependence on amine analogs for RNA self-cleavage. It was concluded that the RNA lacks catalytic ability on its own and that GlcN6P, specifically the amine group, acts as a coenzyme in this system rather than an effector of structural change in the RNA (57). This conclusion was supported by hydroxyl radical probing and UV crosslinking experiments that showed that the RNA maintains a similar structure in the presence or absence of GlcN6P (36). X-ray crystallography was used to solve structures of the *glmS* in the pre- and post-cleavage state (46). The crystal structures also show that there is minimal structural change in the

RNA. Furthermore, the ribozyme is able to undergo GlcN6P-dependent cleavage in the crystalline state. This supports the model that the RNA is in a preformed active state and GlcN6P binds in order to catalyze the cleavage reaction with its amine group. This represents a unique strategy for gene regulation by a metabolite sensing RNA. The *glmS* element can be considered a riboswitch since it senses a metabolite without additional *trans*-acting factors, but it is not a traditional riboswitch since it does not utilize structural rearrangements to confer its regulatory effect.

Although mutational analysis indicated that the *glmS* transcript must have the potential for self-cleavage in response to GlcN6P to repress *glmS* expression *in vivo*, the mechanism of this repression was not evident (101). There is no potential for an intrinsic terminator structure in the RNA sequence, indicating that cleavage was unlikely to promote transcription attenuation. In addition, regulation did not appear to be at the level of translation as the cleavage site in the *glmS* transcript is located ~235 nt 5' of the start codon. This suggests that the 3' post-cleavage product can be translated to make full length GlmS protein. Recent studies showed that 3' post-cleavage transcripts are not normally detectable in *B. subtilis*, but depletion of the essential ribonuclease RNase J1 causes an accumulation of 3' post-cleavage transcripts and a corresponding increase in *glmS-lacZ* expression (169). RNase J1 was shown to have no effect on mutant *glmS* transcripts that are defective in GlcN6P-dependent cleavage, indicating that the cleavage event causes susceptibility to RNase J1. Thus, it has been concluded that the mechanism for decreased gene expression due to *glmS* cleavage involves RNase J1-dependent degredation of the 3' post-cleavage transcript (169).

1.3 Metabolite binding riboswitches

The majority of riboswitches fall under the category of metabolite binding riboswitches. Metabolite binding riboswitches have been found in a large number of gene families, which indicates that they are a major factor in gene regulation. Each riboswitch class responds to a specific effector molecule, which could be a nucleotide, cofactor, amino acid, or even a metal ion.

1.3.1 Riboswitches that bind amino acids

Some bacteria use riboswitches to sense the levels of a particular amino acid in the cell. The riboswitches are typically found in genes involved in the synthesis or transport of the cognate amino acid. Binding of the amino acid typically causes repression of gene expression, thus acting as a negative feedback mechanism.

1.3.1.1 The L box

The L box riboswitch is an RNA motif found in many bacteria that regulates expression of lysine biosynthesis and transport genes by binding lysine (32, 119). The analysis of L box sequences from different organisms indicates that regulation is primarily at the level of transcription termination in Gram-positive organisms and at the level of translation initiation in Gram-negative organisms. Preliminary studies focused on the *B. subtilis lysC* L box RNA (32). The *lysC* gene encodes aspartokinase II, which catalyzes the first step in lysine biosynthesis by phosphorylating aspartate. Expression of transcriptional fusions of *lysC* to *lacZ* decreased when *B. subtilis* cells were grown in media containing high lysine levels. Also, lysine was shown to promote premature

termination of *lysC* transcription in the absence of other factors in a purified *in vitro* transcription assay. These results indicate that part of the function of the L box is as a negative feedback system where lysine inhibits its own synthesis.

The structure of an L box sequence from *Thermotoga maritima* has been solved by X-ray crystallography and it shows that the RNA envelopes lysine in a five-way junction (117). The structure also suggests that the methylene groups of lysine are not recognized by the RNA. The lack of recognition of the methylene groups is predicted to be the reason why analogs with substitutions at the C4 position have been shown to interact with L box RNA (118). Some of these C4 analogs have anti-microbial activity and, due to their ability to bind L box RNA, the anti-microbial activity was predicted to be via inhibition of lysine biosynthesis. This prediction was tested experimentally and the C4 analogs exhibit an anti-microbial effect in minimal media, but not when cells are grown in rich media (118). This suggests that lysine transport mechanisms not regulated by the L box are able to compensate for the decrease in lysine biosynthesis by scavenging lysine from the environment.

The lysine analog aminoethylcysteine (AEC), which is one of the C4 analogs with known anti-microbial activity, differs from lysine by substitution of a sulfur atom for carbon at the C4 position and is able to promote transcription termination of L box RNAs at 10-fold less efficiency than lysine (32). It has been shown that AEC-resistant mutants of *E. coli* and *B. subtilis* have mutations in their *lysC* L box sequence (32). Additional studies determined that the toxicity of AEC in cells is primarily due to its ability to bind LysRS and mischarge $tRNA^{Lys}$ rather than binding to L box RNA and inhibiting gene expression (116). Thus, it was proposed that *lysC* L box mutations confer resistance to

AEC by derepression of *lysC* gene expression. Derepression would lead to an increase in lysine levels so that lysine can outcompete AEC for $tRNA^{Lys}$ charging and eliminate the toxicity of AEC.

1.3.1.2 The glycine riboswitch

An unusual glycine binding riboswitch motif was discovered in the *B. subtilis* *gcvT* operon (56). Glycine binding to the riboswitch induces expression of the *gcvT* operon which encodes genes that allow glycine to be used as an energy source. The riboswitch is unusual because it contains two tandem glycine aptamer domains that are similar but have some variability (Fig. 1.3). Furthermore, the two aptamers participate in cooperative binding where glycine binding to one aptamer increases the affinity for glycine of the other aptamer. This cooperative mechanism is the first example to be discovered in riboswitch-mediated regulation and it is predicted to be a mechanism by which the RNA can efficiently respond to small changes in glycine levels (56).

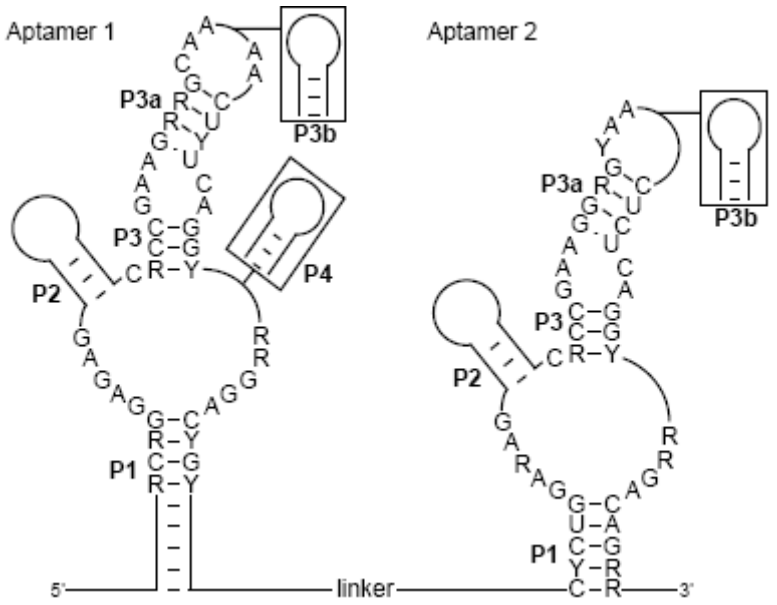

Figure 1.3. The glycine riboswitch. Secondary structure of the consensus sequence of the tandem aptamer domains of the glycine riboswitch. The P1 and P3 helices are highly conserved, but the P2 and P4 helices show no conservation. It has been proposed that the P1 helix of aptamer 1 and the P3a helix of both aptamers are involved in cooperative binding (134). The 3' side of the P1 helix of aptamer 2 is predicted to participate in the expression platform of the riboswitch and confer the regulatory effect of glycine binding. Modified from (134).

Small angle X-ray scattering was used to generate a low resolution structure of the glycine riboswitch (133). The results show that the RNA is in a more compact state when bound to glycine, but the results do not indicate what tertiary interactions are involved in cooperative binding. Nucleotide analog interference mapping (NAIM) was used to investigate what parts of the glycine aptamers are involved in cooperative binding (134). NAIM utilizes incorporation of nucleotide analogs into RNA transcripts at random locations. The RNA transcripts are then selected for activity, in this case glycine binding, and the active transcripts are cleaved at the incorporated analog positions and separated on a gel. If an analog substitution disrupts RNA function, then there will not be a band corresponding to that nucleotide position. No nucleotide analog interference was

observed in the P2 or P4 stem loops, which indicates that these unconserved regions are not involved in binding (134). Interference was observed in the P1 helix of aptamer 1 and the P3a helix of both aptamers. This indicates that these regions are involved in interactions important for binding, but the nature of these interactions is not yet known.

1.3.2 Adenine and guanine riboswitches

A small RNA motif was recently identified in *B. subtilis* located in the 5'-UTR of genes involved in purine biosynthesis and transport (120). This motif, which is found in several other Gram-positive and Gram-negative bacteria, was shown to be a small riboswitch element that binds guanine. The riboswitch element preceding the *xpt-pbuX* operon of *B. subtilis* binds guanine with a K_d of ~5 nM. The element discriminates against the closely related analogs hypoxanthine and xanthine by one order of magnitude, as these compound have a K_d of ~50 nM, but it discriminates strongly against adenine (K_d >300 μM) and many other purine analogs (120).

A secondary structure model was proposed in which the riboswitch aptamer domain folds into three stems, P1, P2 and P3 (Fig. 1.4A). It was suggested that the loops at the ends of P2 and P3 may be involved in a tertiary interaction since these regions exhibit covariation. Additional studies showed that the P2-P3 loop-loop interaction is essential for ligand binding (8). The 3' side of the P1 stem of the *B. subtilis xpt* RNA is predicted to contain sequences that could potentially be involved in formation of the 5' side of an antiterminator element (120). It has been proposed that guanine binding stabilizes the P1 stem in order to disrupt antiterminator formation, allow formation of a terminator structure, and terminate transcription of the operon. Guanine, hypoxanthine,

and xanthine all reduced Mg^{2+}-dependent RNA cleavage in an in-line probing assay consistent with the prediction that ligand binding stabilizes structural elements in the riboswitch (120).

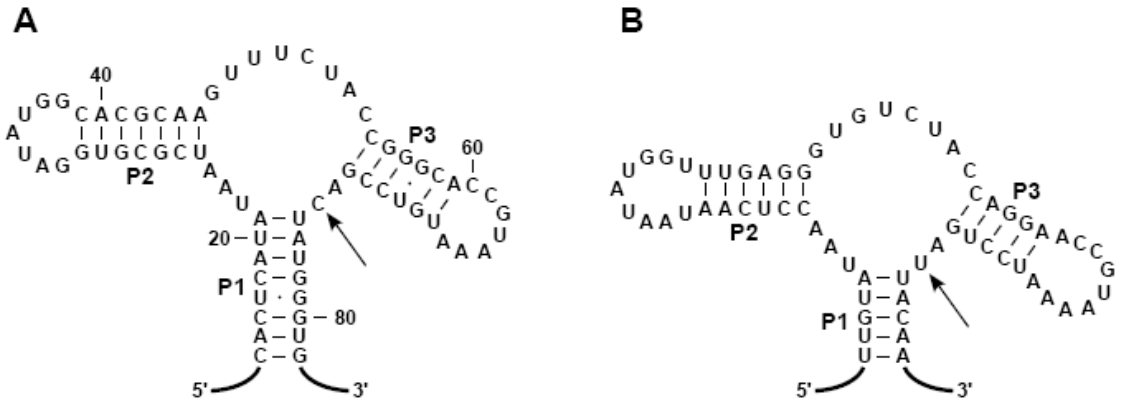

Figure 1.4. The aptamer domain of two purine binding riboswitches. Secondary structure model of the guanine binding *xpt* **(A)** and adenine binding *pbuE* **(B)** aptamer domains from *B. subtilis*. The arrow indicates the nucleotide that determines if the RNA will recognize guanine or adenine. Figure modified from (110).

A subset of predicted guanine responsive riboswitches binds adenine rather than guanine. These RNAs, which include the *ydhL* (now known as *pbuE*) gene in *B. subtilis*, which encodes a putative purine efflux pump, have a single C to U substitution that corresponds to position 74 of the *xpt* sequence (110, Fig. 1.4). The basis for binding of adenine instead of guanine was proposed to be Watson-Crick pairing between the ligand and the U residue at position 74 (110). This was supported by an experiment in which a C to U mutation was introduced into the *xpt* (guanine-responsive) riboswitch sequence; this causes the binding specificity of the RNA to switch from guanine to adenine. Similarly, introduction of the corresponding U to C mutation in the *pbuE* (adenine-responsive) riboswitch switches the binding specificity from adenine to guanine.

Additional studies determined the structure of the *B. subtilis xpt* riboswitch bound to hypoxanthine and guanine and the *Vibrio vulnificus add* riboswitch bound to adenine by X-ray crystallography (8, 81). These structures confirm that the U or C residue is responsible for ligand specificity via a Watson-Crick interaction. Another riboswitch subclass related to guanine riboswitches was discovered in *Mesoplasma florum* that responds to 2'-deoxyguanosine (45). Discovery of these specialized classes of guanine riboswitches suggests that bacteria may have evolved a wide range of variant riboswitches that respond to metabolically relevant purine derivatives.

The mechanism of action of the *pbuE* (adenine-responsive) riboswitch from *B. subtilis* has been investigated (51, 121). Most riboswitches repress gene expression in response to ligand binding, but the *pbuE* riboswitch is atypical because binding of adenine activates gene expression by preventing formation of a downstream terminator. It was determined that proper folding of the *pbuE* aptamer domain occurs in the absence of ligand with increasing metal ion concentrations, but that adenine binding is needed to stabilize the structure (51). Specifically, adenine was shown to stabilize the P1 helix whereas the stability of the P2 and P3 helices is not adenine-dependent (123). It was also determined that if the terminator structure is formed before the aptamer domain is folded then the RNA becomes unable to bind adenine (51, 121). Thus, it was proposed that functionality of the riboswitch is dependent upon the relationship between the rate of ligand binding and the rate of transcription.

The fluorescent compound 2-aminopurine (2-AP), which has the same K_d value for the *pbuE* riboswitch as adenine, was used to perform dynamic binding experiments. These experiments show that the K_d changes from 250 nM to 3 μM over a temperature

range from 15ºC to 35ºC (121). This suggests that binding of adenine to the RNA could be either kinetically or thermodynamically driven, depending on temperature. Although the K_d for adenine and 2-AP is the same, it was concluded that adenine is the likely signal *in vivo* since its association rate with the aptamer is 10-fold faster, which means adenine association is more likely to outcompete the rate of transcription to determine the fate of the transcript.

1.3.3 The RFN element

The RFN element is a riboswitch that regulates expression of riboflavin biosynthesis and transport genes in bacteria (98, 125). The element has been found in a wide range of organisms including all riboflavin biosynthetic operons in the *Bacillus/Clostridium* group (125). The RFN element binds flavin mononucleotide (FMN) and FMN promotes transcription termination of the *B. subtilis ribD* operon *in vitro* (98). The riboswitch is specific for FMN as riboflavin, which differs from FMN by a single phosphate group, was measured to have a 1,000-fold higher K_d. The RFN element is predicted to regulate at the level of translation initiation in some genes and it was shown that FMN promotes an increase in structure in the SD region in the *B. subtilis ypaA* gene (98). The crystal structures of the *Fusobacterium nucleatum* RFN element bound to FMN, riboflavin, and roseoflavin have been solved (124). The structures show that the ligand binding pocket appears to be accessible and flexible. It has been suggested that this makes the RFN element a good target for anti-microbial compounds (124).

The mechanism of action of the transcription termination controlled RFN element in *B. subtilis ribD* has been investigated (96). It was proposed that FMN must bind the

RNA after the transcription of the aptamer domain but before the transcription and formation of the antiterminator, in order for the terminator to form. Thus, the speed of transcription and kinetics of FMN binding are predicted to determine the functionality of the riboswitch. Termination of transcription was monitored in an *in vitro* transcription assay where the speed of transcription was manipulated. Manipulation of transcription speed was mediated by changes in nucleotide concentration and by addition of the NusA protein, which increases RNAP pausing. When the speed of transcription was reduced by NusA or decreased nucleotide concentration, the concentration of FMN required to attain 50% maximal termination was reduced (96). Additionally, two pause sites were identified in the transcript and disruption of these sites results in an increase in transcription rate and a corresponding decrease in responsiveness of the riboswitch to FMN. It was concluded that equilibrium is not reached between the RNA and FMN during transcription and so riboswitch function is kinetically driven by the rate of FMN association (96). This mechanism of action is cited as being the reason that the concentration of FMN required to reach 50% maximal termination is much higher than the K_d would suggest.

1.3.4 The Thi-box element

The Thi-box element was initially discovered as a conserved 38-base sequence in the 5'-UTR of thiamin (vitamin B_1) biosynthetic and transport genes in both Gram-positive and Gram-negative bacteria (61). Expression of a reporter gene fusion to the UTR from the *thiCOGE* operon of *Rhizobium etli* is repressed when cells are grown in the presence of thiamin; indicating that the Thi-box is a functional regulator. The signal

molecule for the Thi-box is thiamine pyrophosphate (TPP), the biologically active form of vitamin B_1. The *E. coli thiC* and *thiM* Thi-box RNAs have been shown to bind TPP and discriminate against thiamin (61, 99).

Multiple structures of Thi-box riboswitches bound to TPP have been solved by X-ray crystallography (82, 126, 127). The structures show that the RNA recognizes TPP in an extended conformation with two parallel helical domains. One of the helical domains interacts with the pyrimidine moiety of TPP and the other interacts with the pyrophosphate moiety. This arrangement explains the preference of the RNA for TPP over thiamin as thiamin lacks the pyrophosphate moiety and cannot make the proper interaction with the second helix. The crystal structures also reveal that the central thiazole ring is not recognized by the RNA. This observation explains why the anti-microbial compound pyrithiamine pyrophosphate (PTPP), which has a pyridine ring instead of a thiazole ring, is able to bind Thi-box RNA with an affinity similar to that of TPP (82, 129). Decreased expression of genes by binding of PTPP to Thi-box riboswitches has been proposed to be a major contributing factor to the toxicity of PTPP in bacterial cells.

In the absence of TPP the SD sequence of *E. coli thiM* RNA is available for chemical modification, but TPP binding was shown to cause a structural change in the RNA in which the SD sequence becomes protected from Mg^{2+}-dependent cleavage (77, 99). These results suggest that TPP-dependent regulation is at the level of translation initiation for *thiM* RNA (Fig. 1.5). Further analysis confirmed the transition in SD structure was responsible for the regulatory effect by incorporating mutations at key residues in *thiM* RNA that lock the RNA in either an “on” or “off” state (74). In contrast

to *E. coli thiM*, binding of TPP to the Thi-box in *B. subtilis tenA* RNA was shown to cause the formation of a transcriptional terminator (131). This demonstrates that there are examples of Thi-box riboswitches that regulate at the level of transcription termination. In general, Thi-box riboswitches appear to act primarily at the level of translation initiation in Gram-negative bacteria and at the level of transcription termination in Gram-positive bacteria (131).

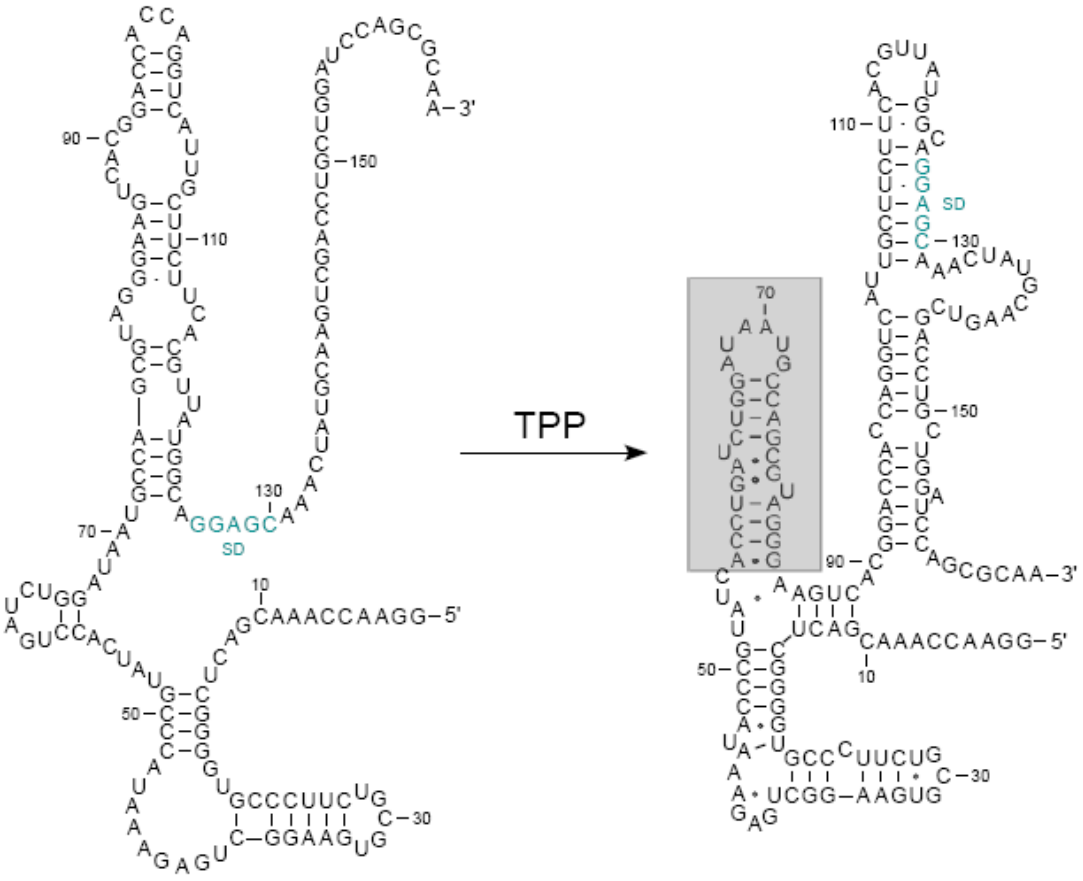

Figure 1.5. The *E. coli thiM* TPP binding riboswitch. Secondary structure model of the Thi box riboswitch located in the 5'-UTR of the *E. coli thiM* gene. In the absence of TPP, the SD (blue) is free and available for translation. TPP binding stabilizes an alternate conformation of the RNA in which the Thi-box sequence (shaded box) forms a stem-loop and the SD is sequestered and unavailable for translation initiation. Modified from (77).

Additional genome searches identified Thi box sequences that are located in archaeal and eukaryotic organisms (128, 170). Thi box sequences in archaea are located in the 5'UTR, which suggests a regulatory response similar to that in bacteria (170). In contrast, Thi box sequences in eukaryotes are found in the 3'-UTR of genes and in introns. Based on these observations, it was proposed that the Thi box may be involved in the regulation of RNA processing/stability and the regulation of transcript splicing in eukaryotes (128). The presence of Thi box riboswitches throughout all domains of life suggests that riboswitches may represent an ancient regulatory mechanism. Additionally, the discovery that riboswitch sequences are present in eukaryotes and archaea indicates that there is potential for discovery of new classes of riboswitches that may or may not resemble bacterial riboswitches.

1.3.5 The AdoCbl element

A riboswitch that binds 5'-deoxy-5'-adenosylcobalamin (AdoCbl, coenzyme B_{12}) controls the expression of an *E. coli* outer membrane cobalamin transporter gene called *btuB* (67, 70, 132). Initial studies concluded that regulation is at the level of translation initiation as a hairpin structure with the potential to sequester the SD sequence is necessary for cobalamin-dependent repression (132). In-line probing experiments show that AdoCbl causes structural changes in *btuB* RNA and mutational analysis determined that disruption of pairing in the SD region results in a loss of regulation (67). Additional work showed that AdoCbl, but not cyanocobalamin (CN-Cbl), inhibits binding of 30S ribosomal subunits to *btuB* RNA (70). Since ribosome binding to *btuB* RNA appears to require an unknown ribosome-associated protein (due to the fact that binding was

observed only when low-salt washed 30S subunits were used), it was proposed that inhibition of ribosome binding is due to AdoCbl-dependent stabilization of the SD hairpin so that the ribosome-associated factor cannot gain access to the SD. The fact that AdoCbl does not affect ribosome binding when ribosomes are pre-bound to *btuB* RNA supports the proposal that AdoCbl-dependent regulation acts at the initial stages of ribosome binding (70).

Phylogenetic analysis determined that the AdoCbl riboswitch is found in the 5'-UTR of cobalamin-related genes in many Gram-positive and Gram-negative bacteria (130). Although the structure of the AdoCbl riboswitch has not been solved, the fact that there are ~57 nt positions that are >90% conserved and at least 10 predicted base-paired elements indicates that the structure is very complex. This is perhaps not surprising as AdoCbl itself is a more complex molecule than most other riboswitch-associated metabolites. The phylogenetic analysis revealed a variant riboswitch upstream of the *B. subtilis yvrCBAK* operon that lacks several conserved helices (130). This RNA demonstrates structural changes in response to AdoCbl and regulates gene expression in response to AdoCbl levels *in vivo*. Additionally, a different variant AdoCbl responsive riboswitch that regulates at the level of transcription termination has been discovered in an ethanolamine utilization operon in *E. faecalis* (106). Although the significance of these variant structural elements is unknown, the discovery of these variants suggests that either evolution of existing AdoCbl sensing riboswitches is well-tolerated or that different RNAs have evolved independently to sense the same metabolite.

1.3.6 Cyclic di-GMP riboswitch

Cyclic di-GMP is a compound that is known to regulate a wide range of cellular functions and recent work has concluded that at least part of the regulatory effect can be attributed to a riboswitch that binds cyclic di-GMP (111). This riboswitch is present in the 5'-UTR of many types of genes including genes involved in virulence, pilus formation, and flagella expression. The riboswitch appears to form a relatively simple secondary structure with two short adjacent stem-loops that exhibit covariation. Despite this simple structure, the RNA is able to bind cyclic di-GMP tightly (K_d ~1 nM) and discriminate against related molecules. Although the mechanism of regulation of cyclic di-GMP riboswitches has not been investigated in detail, there appear to be some representatives that regulate at the level of transcription attenuation and others that regulate at the level of translation initiation. Additionally, there appear to be some representatives where cyclic di-GMP binding results in an increase in gene expression and others where binding results in a decrease in gene expression.

A cyclic di-GMP riboswitch located in the 5'-UTR of an operon encoding genes for construction of flagella in the *Clostridium difficile* genome is predicted to decrease gene expression at the level of transcription attenuation in response to cyclic di-GMP binding (111). The expression of a fusion of this riboswitch to *lacZ* is higher when cells are cotransformed with a plasmid containing a gene that hydrolyzes cyclic di-GMP to pGpG. These results suggest that cyclic di-GMP is the effector molecule for the riboswitch *in vivo*. The cyclic di-GMP riboswitch system is unusual in that most of the representative genes are not involved in the synthesis or transport of the effector

molecule. Instead, it appears that the riboswitch has evolved to regulate fundamental cellular functions by responding to a universal signal.

1.3.7 $PreQ_1$ riboswitch

A riboswitch with a very small aptamer domain was discovered in the 5'-UTR of genes involved in the synthesis of queuosine (79). The RNA binds the queuosine precursor $preQ_1$ (7-aminomethyl-7-deazaguanine) and discriminates against many related analogs. Some analogs were tolerated, as both $preQ_0$ and 7-carboxamide-7-deazaguanine bound the RNA with a K_d within one order of magnitude of the K_d for $preQ_1$. The fact that $preQ_0$ is likely a transient compound that does not accumulate in cells and that 7-carboxamide-7-deazaguanine is unlikely to be present in cells led to the conclusion that $preQ_1$ is the molecular signal for the riboswitch *in vivo*. Binding of $preQ_1$ has been proposed to reduce gene expression because mutations that disrupt the aptamer domain result in an increase in expression of reporter gene fusions. Gene regulation via this riboswitch class appears to be at the level of transcription termination in some examples and at the level of translation initiation in others.

Through phylogenetic analysis, it was initially predicted that the aptamer domain of the RNA consists of a single stem-loop with a tail that totaled 34 nt (79). NMR and X-ray crystallography data determined that the RNA actually forms an H-type pseudoknot structure in the 34 nt core region (Fig. 1.6, 137, 138). The structures show that a Watson-Crick base pair between a conserved cytidine and $preQ_1$ is involved in $preQ_1$ recognition. In fact, mutation of this residue to uridine resulted in a change in ligand specificity so that the RNA binds 2,6-diaminopurine and discriminates against $preQ_1$ (79).

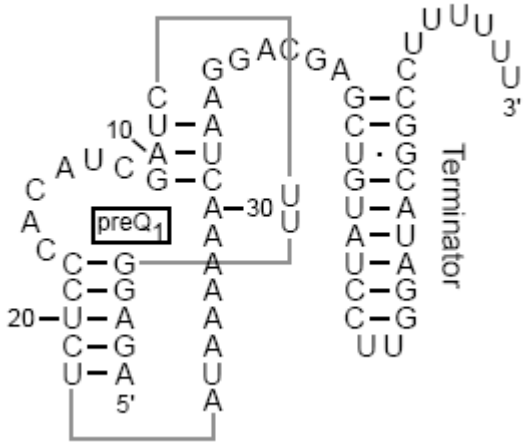

Figure 1.6. Model of the preQ$_1$-I riboswitch. The cartoon structure of the *B. subtilis queC* preQ$_1$-I riboswitch bound to preQ$_1$ derived from X-ray crystallography data (137). PreQ$_1$ binding causes formation of an H-type pseudoknot and promotes formation of an intrinsic transcriptional terminator. PreQ$_1$ interacts with G5, U6, G11, A30, and participates in a Watson-Crick base pair with C17. Modified from (137).

A second class of preQ$_1$ binding riboswitch, named preQ$_1$-II, was discovered in the Streptococcaceae family (136). The aptamer domain of preQ$_1$-II consists of about twice as many nt, lacks primary sequence conservation with the preQ$_1$-I class, and is predicted to have four helices. PreQ$_1$-II RNA exhibits differential discrimination against analogs versus preQ$_1$-I. For example, both 7-carboxamide-7-deazaguanine and preQ$_1$ bind preQ$_1$-I RNA with similar efficiency, but preQ$_1$-II RNA binds 7-carboxamide-7-deazaguanine with at least 100-fold lower efficiency than it binds preQ$_1$. This suggests that these two classes of preQ$_1$ riboswitch have different ligand binding pockets. This idea is supported by the inability to find a conserved cytosine residue in preQ$_1$-II that could be involved in Watson-Crick base pairing with preQ$_1$. Given the different sequence, structural elements, and binding pockets, it appears that these riboswitch classes have evolved separately to respond to the same molecule.

1.3.8 SAH riboswitch

A riboswitch motif that responds to *S*-adenosylhomocysteine (SAH) has been discovered in genes involved in *S*-adenosylmethionine (SAM) recycling in both Gram-positive and Gram-negative bacteria (92). Genome searches revealed 68 examples of this riboswitch. Some of these are predicted to have transcriptional terminators and regulate at the level of transcription while others are predicted to regulate at the level of translation initiation. Most examples of the riboswitch are found in genes that encode SAH hydrolase and methylenetetrahydrofolate reductase, which are the enzymes responsible for the recycling of SAH (which is a byproduct when SAM is used as a methyl donor) to methionine.

The riboswitch RNA was shown to bind SAH and binding causes structural changes in the RNA. Data from fusions of the SAH riboswitch to a *lacZ* reporter gene indicate that gene expression is induced in the presence of SAH. This makes the SAH riboswitch a rare case since most metabolite binding riboswitches repress gene expression in response to ligand binding. It is not surprising that SAH recycling genes are induced by the SAH riboswitch because the build up of SAH is toxic to cells (140).

The SAH riboswitch RNA discriminates strongly against binding to SAM and many other closely related analogs (92). Although the structure of the SAH binding pocket is unknown, analog binding data suggests that every functional group on SAH is important for ligand recognition (92). The high degree of specificity is critical for this riboswitch because in order to maintain the proper regulatory response it must differentiate between two molecules, SAM and SAH, which differ by only a single methyl group.

1.3.9 SAM riboswitches

1.3.9.1 The S box riboswitch

The S box is an RNA motif that was discovered in the 5'-UTR of 11 transcriptional units that encode genes involved in methionine and cysteine biosynthesis in *B. subtilis* (30). Phylogenetic analysis determined that this motif is highly conserved and contains regions of covariation. Additional genomic searches uncovered hundreds of S box motifs that are located in a wide range of bacterial species (7, 78). Most of the sequences have the potential to form an intrinsic transcriptional terminator and a competing antiterminator structure and thus regulate gene expression at the level of premature transcription termination (Fig. 1.7). The rarer S box leader RNAs that lack terminators appear to have the ability to occlude the SD region and regulate at the level of translation initiation. The S box is a functional regulator *in vivo* as expression of a *yitJ-lacZ* fusion is induced when cells are starved for methionine (30). The terminator structure was determined to be critical for S box regulation as disruption of the terminator structure by mutation results in constitutive *lacZ* expression. It was proposed that under high methionine conditions, an unknown regulatory factor could bind the RNA, disrupt antiterminator formation, cause terminator formation, and repress gene expression.

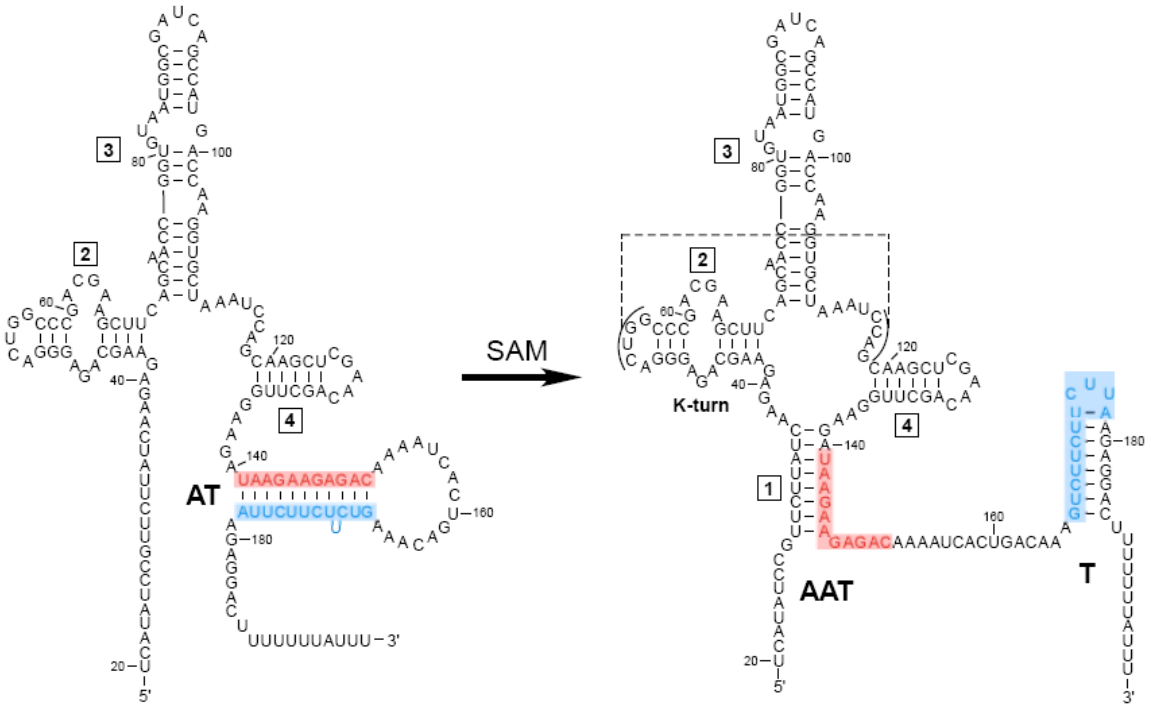

Figure 1.7. The *B. subtilis yitJ* S box leader RNA. Structural model of the *yitJ* S box RNA in the absence and presence of SAM. Helices are identified by boxed numbers; AAT, anti-antiterminator; AT, antiterminator; T, terminator; K-turn, kink-turn motif. Red and blue denotes residues that form the AT. SAM binding stabilizes helix 1 (the AAT) which prevents AT formation. This allows the T to form and terminate transcription. Modified from (58).

The identity of the unknown regulatory factor for the S box regulatory system was determined through a variety of *in vitro* experiments to be SAM (58, 100). SAM was shown to bind S box RNA and cause a structural change where an anti-antiterminator structure is stabilized that would allow formation of the terminator. Additionally, SAM promotes transcription termination when added to an *in vitro* transcription reaction of S box leader RNAs. Closely related analogs do not promote structural changes or termination which indicates that these effects are specific for SAM. The fact that the effector molecule is SAM is not surprising as it was previously proposed that SAM could

be a repressor of methionine biosynthesis due to the fact that overexpression of SAM synthetase had been shown to cause methionine auxotrophy (102). Despite these results, there was still a question of whether SAM was the effector molecule *in vivo*. When *B. subtilis* cells were mutagenized to search for *trans*-acting mutations that derepress S box expression, only one mutation was found (60). This mutation is in the coding sequence of the *metK* gene, which encodes SAM synthetase, and it causes a reduction of SAM synthetase activity that results in a reduction of SAM levels *in vivo*. The fact that the only known mutation that derepresses S box expression affects SAM levels suggests that SAM is indeed the effector molecule for the S box *in vivo*.

Analysis of S box sequences led to the development of a secondary structure model in which S box RNA was predicted to contain 4 helical domains upstream of the terminator structure (30, Fig. 1.7). Additionally, residues that covary between the loop of helix 2 and the junction between helices 3 and 4 were predicted to participate in a tertiary interaction to form a pseudoknot. These predictions were tested by a variety of *in vitro* techniques. Structure probing experiments of S box RNA structure with Mg^{2+}-dependent cleavage, RNase T1, and RNase A all support the secondary structure model (59, 100). Point mutations that disrupt pairing in the potential pseudoknot structure resulted in loss of SAM binding and SAM-dependent transcription termination (59). The effects of SAM were restored when mutations were incorporated that restored pairing, indicating that the pseudoknot is required for proper S box function. Correct formation of the pseudoknot structure is thought to be promoted by a K-turn motif located in helix 2 as K-turn motifs are able to mediate RNA tertiary interactions (59, 97, 142). The archaeal K-turn binding protein L7Ae has been shown to bind the K-turn motif of *B. subtilis yitJ in vitro* and

stabilize the pseudoknot and overall ligand bound conformation (42). This suggests that additional *trans*-acting factors may be involved in fine tuning S box regulation, but the *in vivo* significance of this result is in question since *B. subtilis* does not encode a homolog of L7Ae.

The structure of the S box aptamer domain in the presence of SAM was solved by X-ray crystallography and it confirms many structural predictions that were made from sequence analysis (62). Specifically, it confirms the existence of helices 1-4, their formation of a four-way helical junction, and formation of the pseudoknot between the loop of helix 2 and the covarying nt 5' of helix 4 (Fig. 1.7).

The crystal structure also shows the location and conformation of SAM bound to the RNA (62). SAM is enclosed in a pocket formed by helix 1, helix 3, and the junction between helices 1 and 2. The crystal structure results indicate that almost all of the functional groups of SAM are recognized by the RNA. The positive charge on the sulfur atom is recognized by favorable electrostatic interactions with two uracil residues in helix 1. This is predicted to account for the ability for the RNA to differentiate between SAM and SAH, which differ by the lack of a methyl group on SAH, since the methyl group of SAM is pointed toward a solvent region and is not recognized by the RNA. SAM is typically bound in an extended conformation by most SAM-binding proteins, but in the S box structure SAM adopts a compact conformation in which the methionine moiety stacks on the adenine ring (62, 141). Thus, the S box has evolved to bind SAM in an unusual conformation.

S box riboswitch sequences contain regions of high conservation, but they also have some regions of variability. The significance of this variability was investigated by

analyzing the properties of all 11 S box sequences present in *B. subtilis* (87). The sensitivity of these S box sequences to SAM was shown to be variable as there is a 250-fold range in K_d. In addition, the stringency of regulation of *lacZ* expression by these S box RNAs was shown to have a 250-fold range and generally correlates with the K_d. These results indicate that variability in the S box sequence can cause a significant difference in the pattern of the regulatory response. The elements that account for these differences are unknown, but it is possible that both the affinity for SAM and the relative stability of the antiterminator vs. the terminator are involved (87). It was also proposed that variability in the S box sequences provides the cell with a simple way to differentially regulate genes that respond to the same effector molecule. It was noted that genes involved in methionine biosynthesis generally display a more stringent pattern of regulation than genes involved in methionine transport. This makes logical sense as it costs the cell more energy to make methionine than to import it from the environment. Taken together, these results indicate that the S box regulon, and potentially other riboswitch systems, exhibit a more complex regulatory response than was initially appreciated.

1.3.9.2 SAM-II riboswitch

A second SAM binding riboswitch (called SAM-II) that shares no primary or secondary structural elements with the S box was identified in α-proteobacteria (19). The SAM-II element is predominantly found upstream of methionine biosynthetic genes in α-proteobacteria and some other Gram-negative species, but it is not found in Gram-positive bacteria. The SAM-II element binds SAM, but the affinity for SAM appears to

be lower than the affinity for SAM of most S box sequences. Despite the lower affinity, SAM binding was shown to be specific as the SAM-II element discriminates strongly against SAH. Some SAM-II riboswitches appear to have the ability to sequester the SD region and others appear to have putative terminator sequences, but the SAM-dependent regulatory effect of these riboswitches has yet to be characterized biochemically (19, 28).

The SAM-bound SAM-II riboswitch structure has been determined by X-ray crystallography (28). The SAM-II RNA folds into a classic H-type pseudoknot conformation consisting of two loops and two helices. SAM is bound in an extended conformation (vs. the compact conformation observed when bound to S box RNA) and contacts five consecutive base pairs or base triples. Although SAM is bound in a different conformation in S box and SAM-II RNAs, the positively charged sulfur atom is recognized in a similar manner by the carbonyl group of two uracil residues. The completely different sequences, secondary structure, and SAM conformation suggest that the S box and SAM-II RNAs are completely unrelated riboswitches that appear to have evolved separately to respond to the same metabolite and fill the same role of regulating genes related to methionine biosynthesis.

1.3.9.3 The S_{MK} box riboswitch

The S box riboswitch regulates methionine-related genes in most members of the *Bacillus/Clostridium* group, but it is not found in most members of the Lactobacillales order (78, 143). Investigation of available Lactobacillales genomic sequences led to the discovery of a new conserved element, which was named the S_{MK} box, in the 5'-UTR of the *metK* gene of many Lactobacillales species (25). The S_{MK} box exhibits no similarity

to the S box or SAM-II sequences, and has the potential for pairing of residues that overlap the SD sequence with an upstream anti-SD (ASD) region (Fig 1.8). Based on these findings, it was hypothesized that the S_{MK} box is a riboswitch that regulates at the level of translation initiation by occlusion of the SD sequence. Additionally, since the *metK* gene in many other Gram-positive organisms is regulated by the S box riboswitch in response to SAM, it was hypothesized that SAM may be the effector molecule for the S_{MK} box system.

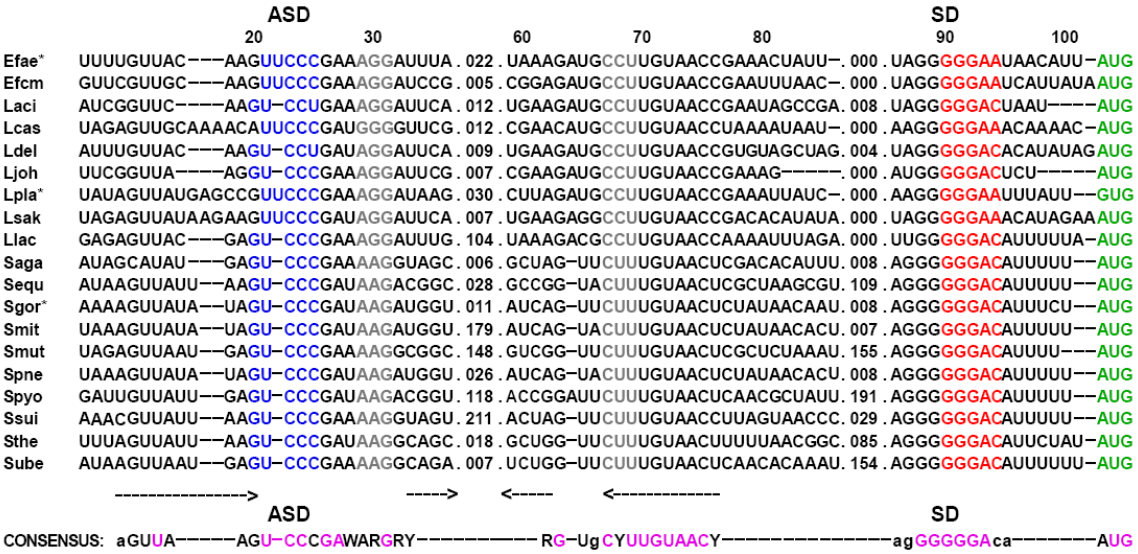

Figure 1.8. Alignment of the 5'-UTR of *metK* from 19 Lactobacillales species. Sequences are aligned to the *Enterococcus faecalis* sequence, which extends from +9 relative to the predicted transcription start site to the end of the AUG start codon (+104, green). Abbreviations are as follows: Efae, *E. faecalis*; Efcm, *E. faecium*; Laci, *Lactobacillus acidophilus*; Lcas, *L. casei*; Ldel, *L. delbrueckii*; Ljoh, *L. johnsonii*; Lpla, *L. plantarum*; Lsak, *Lactococcus sakei*; Llac, *L. lactis*; Saga, *Streptococcus agalactiae*; Sequ, *S. equi*; Sgor, *S. gordonii*; Smit, *S. mitis*; Smut, *S. mutans*; Spne, *S. pneumoniae*; Spyo, *S. pyogenes*; Ssui, *S. suis*; Sthe, *S. thermophilus*; Sube, *S. uberus*. A consensus sequence is shown at the bottom (magenta, 100% conserved residues; black capitals, ≥85% conserved; lower-case, ≥45% conserved; R, G or A; Y, C or U; W, A or U). Numbering at the top is for the *E. faecalis metK* sequence. Numbers within the alignment indicate the number of residues in the upper part of the core element (22 for *E. faecalis*) or in the region just upstream of the SD region (0 for *E. faecalis*). Potential pairing exists between residues corresponding to *E. faecalis* 90-95 in the SD region (shown in red) and an upstream ASD region comprised of residues 21-25 (shown in blue).

1.5 Goals of this study

The goals of this study were to test the hypothesis that the S_{MK} box is a riboswitch that regulates expression of the *metK* gene in many Lactobacillales species. In addition, we wanted to elucidate the mechanism of gene regulation by this riboswitch and identify its effector molecule.

CHAPTER 2

CHARACTERIZATION OF THE S_{MK} BOX RIBOSWITCH

2.1 Introduction

Although the S box mechanism (see 1.4.8.1) commonly regulates methionine-related genes in organisms in the *Bacillus/Clostridium* group, it has not been found in most members of the Lactobacillales order (78, 143). The *metK* gene (which encodes SAM synthetase, the enzyme responsible for synthesis of SAM from methionine and ATP) is usually an S box gene in organisms in which this mechanism is used. As SAM is the molecular effector controlling S box gene expression, regulation of *metK* by this mechanism provides a sensitive feedback response to the intracellular concentration of SAM. We therefore were interested in determining whether SAM is the molecular effector for *metK* regulation in bacteria that lack the S box mechanism.

We found that *metK* genes from all *Enterococcus*, *Lactococcus*, *Lactobacillus*, and *Streptococcus* spp. for which genomic sequences were available contain a conserved 5' element, which we named the S_{MK} box, that shares no similarity to the S box or SAM-II riboswitches (Fig. 1.8). The S_{MK} box has high primary sequence conservation at certain positions and conserved positioning of these residues relative to predicted helical domains that covary to maintain base-pairing. None of the sequences contain elements

resembling a transcriptional terminator, but all of them have the potential for pairing of part of the SD sequence with an upstream region called the anti-SD (ASD). The minimal upper portion of the core element was defined as a 4-bp helix (corresponding to the pairing of residues 33-36 to residues 59-62 in *E. faecalis metK*). Although the core region is highly conserved, there are regions of high variability in S_{MK} box sequences. In particular, the size of the helical domain above the core element ranges from 5 to 211 nt and the distance between the core upstream element and the SD sequence ranges from 4 to 200 nt in different species. The conservation of the core element, including the possible pairing of the ASD sequence with the SD sequence, suggested that the S_{MK} box could be a riboswitch that participates in *metK* regulation at the level of translation initiation. We hypothesized that SAM was likely to be the effector molecule for this riboswitch and developed a model in which SAM binding promotes pairing of the ASD-SD regions (Fig. 2.1). In order to test our hypothesis, we employed binding assays to test the affinity of the RNA for SAM, structure probing assays to monitor changes in RNA structure in response to SAM, and reporter gene fusions to determine if the S_{MK} box is a functional gene regulator *in vivo*.

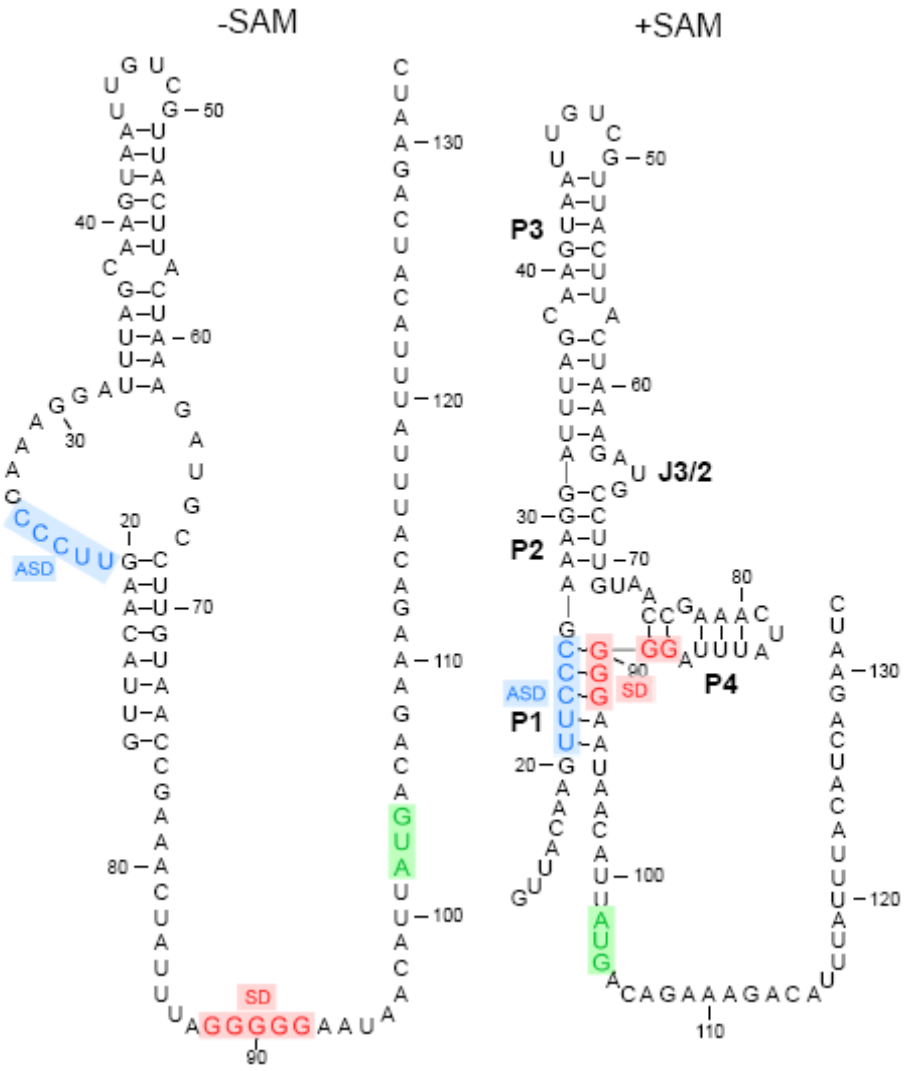

Figure 2.1. Structural model of S_{MK} box RNA from *E. faecalis*. Putative forms of *E. faecalis metK* RNA in the absence and presence of SAM. RNA is shown from residues 13-133 relative to the predicted transcription start site. Blue, ASD region; red, SD sequence; green, start codon of coding sequence. Figure modified from (25).

2.2 Materials and methods

2.2.1 Generation of DNA templates

DNA templates for construction of *lacZ* fusions and T7 RNAP transcription were generated by combination of complementary pairs of oligonucleotides via 4 nt 3' overhangs. First, the 5' ends of internal oligonucleotides were phosphorylated with T4

polynucleotide kinase (Ambion) using conditions recommended by the manufacturer. Then complementary pairs of oligonucleotides (60 pmol in 70 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 5 mM DTT, 50 μM rATP) were heated to 95ºC for 5 min and slow cooled for 1.5 h. Sets of oligonucleotides that formed the template were then mixed (10 pmol each), incubated at 45ºC for 5 min, and allowed to cool to room temperature for 30 min. T4 DNA ligase (800 U, New England BioLabs) was added and the sample was incubated overnight at 16ºC. The ligated DNA template was amplified by PCR using the outside oligonucleotides as primers. The PCR product was purified by a Qiagen PCR clean up kit and the sequence was verified by DNA sequencing by the 3700 DNA Analyzer (Applied Biosystems, Inc., Ohio State University Plant Microbe Genomics Facility). Templates for T7 RNAP transcription contained a T7 promoter and two G residues upstream of the sequence of interest. DNA templates for *lacZ* fusions replaced the *E. faecalis metK* promoter with the *B. subtilis glyQS* promoter. The *metK* constructs matched the sequences for *E. faecalis* strain V583 (NCBI AE016830), *L. plantarum* strain WCFS1 (NCBI AL935263) and *S. gordonii* Challis (The Institute for Genomic Research, http://www.tigr.org).

2.2.2 T7 RNAP transcription

RNA was generated by using a MEGA-shortscript T7 RNAP high-yield transcription kit (Ambion). RNA transcripts were purified on a 6% denaturing polyacrylamide gel, visualized by UV shadowing, and eluted in NaOAc (300 mM, pH 4.5) containing EDTA (1 mM). RNA was extracted with phenol-chloroform, precipitated in

ethanol, resuspended in water, and RNA concentration was determined by measuring absorbance at 260 nm.

2.2.3 SAM binding assay

T7 RNAP-transcribed RNA (3 μM) in 1X transcription buffer (50 mM Tris-HCl, pH 8.0, 20 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT, 10% glycerol) was heated to 65ºC and slow-cooled to 40ºC. Competitor compounds (unlabeled SAM or SAH, 400 μM) were added prior to addition of [methyl-^{14}C]SAM (8 μM, ICN; 52 mCi mmol^{-1} [1.92 GBq mmol^{-1}]). After incubation for 10 min at room temperature, samples were passed through a Nanosep 10K Omega filter and washed four times with 150 μl 1X transcription buffer. Material retained by the filter was mixed with Packard BioScience Ultima Gold scintillation fluid and counted in a Packard Tri-Carb 2100TR liquid scintillation counter. Each experiment was repeated at least twice.

2.2.4 Structural probing of S_{MK} box RNA

RNAs synthesized by T7 RNAP transcription were 5' end-labeled with [γ-^{32}P]ATP (7,000 Ci mmol^{-1}) using a KinaseMax kit (Ambion) and passed through a MicroSpin G-50 column (Amersham Biosciences). Labeled RNAs (600 nM) in 1X transcription buffer in the presence or absence of SAM (140 μM) were heated to 65ºC and slow-cooled to 40ºC. RNase T1 (0.01 U μl^{-1}; Ambion), RNase A (0.67 ng μl^{-1}; Ambion) or RNase V1 (5 x 10^{-6} to 6.7 x 10^{-5} U μl^{-1}; Ambion) was added and samples were incubated at room temperature for 5-30 min. Reactions were halted by addition of Precipitation Inactivation Buffer (Ambion) and digestion products were resuspended in

gel-loading buffer, resolved using 10% denaturing polyacrylamide gel electrophoresis (PAGE) and visualized by PhosphorImager analysis (Molecular Dynamics). Size standards were generated by RNase digestion of denatured RNAs.

2.2.5 Bacterial strains and growth conditions

B. subtilis strains BR151 (*lys-3 metB10 trpC2*), ZB449 (*trpC2 pheA1 abrB703* SPβ-cured), and ZB307A (SPβc2del2::Tn917::pSK10Δ6) were grown on tryptose blood agar base medium (TBAB; Difco) or in 2XYT broth. Spizizen minimal medium for *B. subtilis* was made with glucose as the carbon source as described (2). *E. coli* strain DH5α (φ80d*lacZΔM15 endA1 recA1 hsdR17* (r_k^- m_k^+) *thi-1 gyrA96 relA1* Δ(*lacZYA-argF*)U169) was grown on LB medium (104). Antibiotics were used at the following concentrations: ampicillin, 50 μg ml^{-1}; chloramphenicol, 5 μg ml^{-1} for selection, 0.1 μg ml^{-1} for induction. All growth was at 37ºC.

2.2.6 *lacZ* fusions and β-galactosidase assays

The *E. faecalis metK* leader region, including 15 nt downstream of the translational start site, was positioned downstream of the *B. subtilis glyQS* promoter. The resulting DNA fragment was inserted into a *lacZ* fusion vector (pFG328, 146) to generate either a *metK-lacZ* transcriptional fusion (using a variant of *lacZ* in which the *E. coli trpA* translational start region is fused to the third codon of *lacZ* to allow efficient translation initiation in *B. subtilis*, 147) or an in-frame *metK-lacZ* translational fusion (containing the first 5 codons of *metK* fused to codon 18 of *lacZ*, generated by deletion of the *trpA-lacZ* junction). Constructs were introduced into the chromosome of *B. subtilis* strain BR151

by integration into an SPβ prophage. *B. subtilis* strains containing *lacZ* fusions were grown at 37ºC in Spizizen minimal medium containing methionine (50 μg ml^{-1}) until early exponential growth. The cells were harvested by centrifugation and resuspended in Spizizen medium either with or without methionine, and samples were collected at 1 h intervals and assayed for β-galactosidase activity after toluene permeabilization of the cells (104). Growth experiments were repeated at least twice and reproducibility was ±10%.

2.3 Results

2.3.1 S_{MK} box RNA binds SAM

metK genes in many Gram-positive organisms are regulated by direct binding of SAM to the leader RNA which contains an S box riboswitch. We therefore considered SAM a likely effector for *metK* regulation via the S_{MK} box element, and we tested the *E. faecalis metK* leader RNA for SAM binding under conditions similar to those used for S box RNAs (58). RNA containing sequences from position 15 to 118 relative to the predicted transcription start site was generated by *in vitro* transcription with T7 RNAP and incubated with radiolabeled SAM. Unbound SAM was removed by filtration, and radiolabeled SAM retained by the filter was quantitated. SAM bound to the 15-118 RNA, and binding was blocked by addition of excess unlabeled SAM (Fig. 2.2); in contrast, unlabeled SAH did not compete for binding of SAM, indicating that SAM binding is specific. The efficiency and specificity of SAM binding was similar to that of S box RNAs (58). The *E. faecalis metK* RNA therefore binds SAM in the absence of any other cellular factor. Conservation of the core S_{MK} box RNA element upstream of *metK* genes

in other members of the Lactobacillales suggested that these RNAs might have a similar function. RNAs corresponding to the 5' regions of the *Lactobacillus plantarum* and *Streptococcus gordonii metK* genes also bound SAM (Fig. 2.2), demonstrating that SAM binding is a general property of this class of RNA elements.

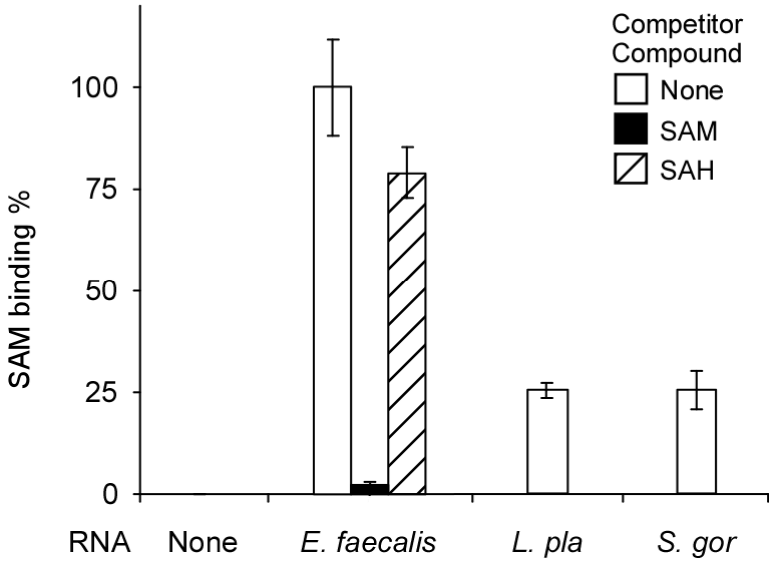

Figure 2.2. Binding of SAM to S_{MK} box RNAs. RNAs generated by T7 RNAP transcription were gel-purified, denatured and refolded in the presence of [^{14}C]SAM, and the RNA-bound SAM was separated from unbound SAM by size-exclusion filtration. Retention of SAM is expressed relative to that of the wild-type *E. faecalis metK* RNA extending from residue 15 to 118 (Fig. 2.1). No SAM was retained by the filter in the absence of RNA. Black or gray bar denotes a sample incubated with a 50-fold molar excess of unlabeled SAM or SAH, respectively. *L. pla*, *Lactobacillus plantarum metK* RNA; *S. gor*, *Streptococcus gordonii metK* RNA. Figure modified from (25).

2.3.2 Mutational analysis of S_{MK} box RNA

Deletion analysis of the *E. faecalis metK* leader was used to determine the limits of the element required for SAM binding (Fig. 2.3). Deletion of residues at the 5' end of the element, to generate an RNA extending from position 20 to 118, resulted in a slight enhancement in SAM binding. This is likely due to the disruption of the predicted

pairing in the absence of SAM between residues 15-20 and 68-73 (Fig. 2.1). The disruption of this structure is predicted to destabilize the SAM-free form of the RNA and shift the equilibrium to the SAM-bound form. Deletion of additional residues at the 5' end, to generate an RNA extending from position 32 to 118 which lacks the ASD and 5' side of helix P2, resulted in loss of detectable SAM binding (Fig. 2.3). RNAs extending from position 15 to position 95 or 105 retained 60%-85% of the binding activity of the 15-118 RNA, whereas an RNA extending to position 90, in which the 3' portion of the SD sequence was removed, showed no detectable SAM binding. These results suggested that SAM binding requires sequences extending from the ASD through the SD.

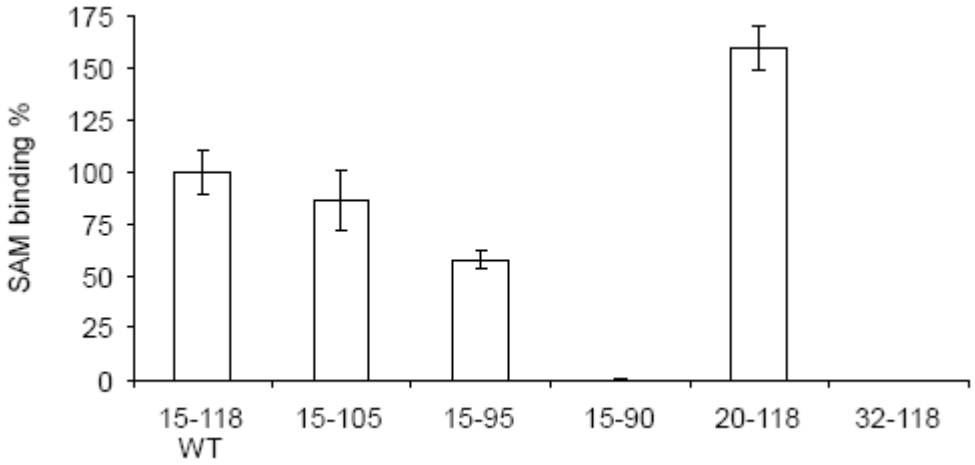

Figure 2.3. Binding of SAM to truncated *E. faecalis* S_{MK} box RNAs. RNAs generated by T7 RNAP transcription were gel-purified, denatured and refolded in the presence of [^{14}C]SAM, and the RNA-bound SAM was separated from unbound SAM by size-exclusion filtration. Retention of SAM is expressed relative to that of the wild-type *E. faecalis metK* RNA extending from residue 15 to 118 (Fig. 2.1). No SAM was retained by the filter in the absence of RNA. Figure derived from data from (25, 26).

Internal deletions and point mutations of the *E. faecalis metK* leader were utilized to further define the SAM binding element (Fig. 2.4). An RNA containing residues 15-33, 45-50, and 62-118, where part of the P3 helix has been deleted, showed no significant

decrease in SAM binding. When the entire P3 stem was eliminated, represented by an RNA containing residues 15-33 and 62-118, SAM binding was reduced 10-fold. This suggests that the overall size of the P3 stem is not crucial in determining whether the RNA can bind SAM (this is supported by the high amount of variability in P3 size between species, Fig. 1.8), but at the very least a small P3 is required for SAM binding.

Disruption of the P2 helix by G30C/G31C or C67G/C68G substitutions resulted in a severe reduction in SAM binding (Fig. 2.4). Restoration of pairing by combining both sets of substitutions resulted in partial restoration of SAM binding (15-fold lower efficiency than wild-type RNA), indicating that pairing in this region is important for SAM binding. The lack of complete restoration of binding activity may be due to a requirement for primary sequence maintenance in this region.

Deletion of the P4 linker helix was achieved by making an RNA construct containing residues 15-76 and 87-118 (Fig. 2.4). This RNA showed only a slight (1.3-fold) reduction in SAM binding, indicating that this highly variable region connecting the P2 helix to the SD (Fig. 1.8) is not required for SAM binding.

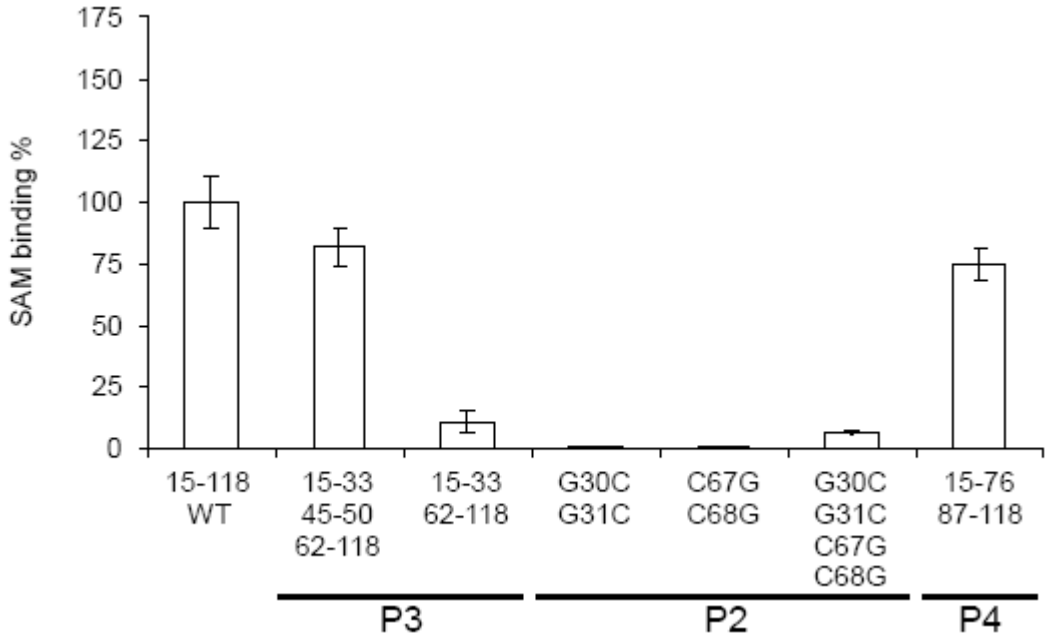

Figure 2.4. Binding of SAM to *E. faecalis* S_{MK} box RNAs with mutations in the P2, P3, or P4 helix. RNAs generated by T7 RNAP transcription were gel-purified, denatured and refolded in the presence of [^{14}C]SAM, and the RNA-bound SAM was separated from unbound SAM by size-exclusion filtration. Retention of SAM is expressed relative to that of the wild-type *E. faecalis metK* RNA extending from residue 15 to 118 and helix names are derived from Fig. 2.1. No SAM was retained by the filter in the absence of RNA. P2 data from (25).

The P1 helix was predicted to consist of pairing between residues U21-C25 in the ASD region with residues G90-A94, and this pairing potential was found in each *metK* leader sequence (Fig. 1.8). The downstream sequence overlaps the SD sequence (positions 88-92), and the residues corresponding to the predicted U21-A94 pairing covaried to G-C in 14 of 19 sequences. Point mutations were made in the ASD-SD region to assess the importance of the potential for pairing between these regions. Introduction of a U22G or A93C substitution to disrupt pairing resulted in loss of SAM binding, and combining U22G and A93C substitutions to restore pairing also restored SAM binding (Fig. 2.5). A similar pattern of results was obtained for the C24U/G91A

and C24G/G91C substitutions despite 100% conservation of the residues corresponding to C24 and G91 (Fig. 1.8). The restoration of binding activity in the C24U/G91A and C24G/G91C double mutants suggests that conservation of G91 is dictated by its participation in the SD sequence and not by any requirement for interaction with SAM; conservation of C24 is presumably dictated in turn by the requirement to retain pairing with G91. The C24U/G91A double mutant RNA showed a complete restoration of SAM binding, but the C24G/G91C double mutant showed 1.5-fold lower binding of SAM than the wild-type RNA, indicating that there is a slight primary sequence effect at these positions (Fig. 2.5). Disruption of ASD-SD pairing with a G90A substitution severely reduced SAM binding as expected, but a C25U substitution had little effect on SAM binding. This suggests that a wobble U-G pair is tolerated at this position. Surprisingly, the C25U/G90A double mutant, in which the potential for pairing is restored and a canonical GGAGG SD sequence is generated, exhibited no detectable SAM binding. This indicates that there is a primary sequence requirement at the G90 position that is independent of pairing with the ASD sequence.

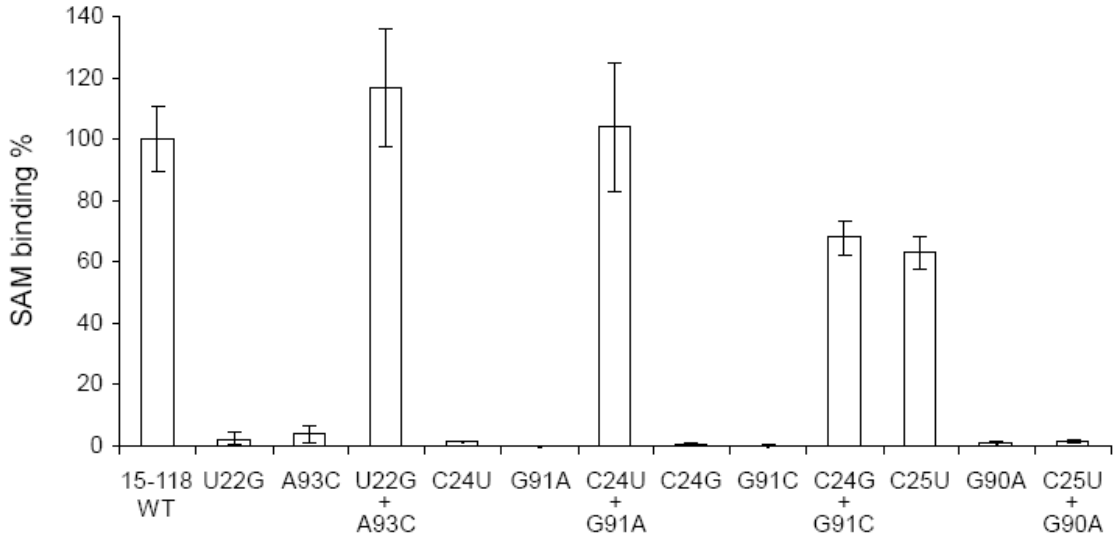

Figure 2.5. Binding of SAM to *E. faecalis* S_{MK} box RNAs with mutations in the ASD-SD regions. RNAs generated by T7 RNAP transcription were gel-purified, denatured and refolded in the presence of [^{14}C]SAM, and the RNA-bound SAM was separated from unbound SAM by size-exclusion filtration. Retention of SAM is expressed relative to that of the wild-type *E. faecalis metK* RNA extending from residue 15 to 118. No SAM was retained by the filter in the absence of RNA. Figure derived from data in (25, 26, 55).

Combined together, the mutational analysis data helped to define the SAM binding element with more precision. Specifically, the RNA must contain the ASD and SD regions and pairing between these regions (P1 helix) is required for SAM binding. Also, the P2 helix and a minimal P3 helix are required for SAM binding. The P4 region and the overall size of the P3 helix were shown not to be crucial for SAM binding, but they may play a role in fine-tuning the affinity of the RNA for SAM.

2.3.3 SAM causes structural changes in S_{MK} box RNA

Digestion of 5' end-labeled *E. faecalis* S_{MK} box RNA with RNases T1 (which cleaves adjacent to unstructured G residues) and A (which cleaves adjacent to

unstructured U and C residues) was used to identify SAM-induced structural alterations. Cleavage by RNases T1 and A was most efficient at residues predicted to be unpaired in the absence of SAM (Fig. 2.1), with extensive cleavage by RNase T1 at G47 (in the terminal loop of P3), G77, and G88 (first position of the SD) (Fig. 2.6A). Cleavage by RNase A was most efficient at U82, with additional products representing residues in the terminal loop of P3 and U65 (Fig. 2.6B). Products corresponding to cleavage at residue A64 appeared in all samples, presumably due to Mg^{2+}-dependent cleavage. Comparison to denatured RNA revealed decreased cleavage by RNase T1 at positions G50, G71 and G89-92, suggesting that these residues may be protected in the folded RNA. Similarly, residues 35, 38, 42, 52, 56, 59, and 72 showed reduced cleavage by RNase A in the folded RNA as compared with the denatured RNA, consistent with the prediction that the core element is structured even in the absence of SAM. Residues G47 and U48 (within the terminal loop of P3) and U65 (in the J3/2 bulge) showed increased cleavage in the folded RNA, suggesting that these residues become more exposed. The most prominent SAM-dependent change in the RNase T1 digestion pattern was the protection of the SD region, particularly residue G88, in the presence of SAM (Fig. 2.6A). Addition of SAM also resulted in protection of U86 (2 nt upstream of the SD) and U95 (1 nt downstream of the region predicted to pair with the ASD) from digestion by RNase A (Fig. 2.6B). The protection of the SD region in the presence of SAM is consistent with a model in which SAM binding results in sequestration of the SD region, thereby inhibiting translation initiation. In contrast, residues G104 and G108 (within and downstream of the AUG start codon) showed increased cleavage in the presence of SAM, which suggests that this region becomes less structured under these conditions.

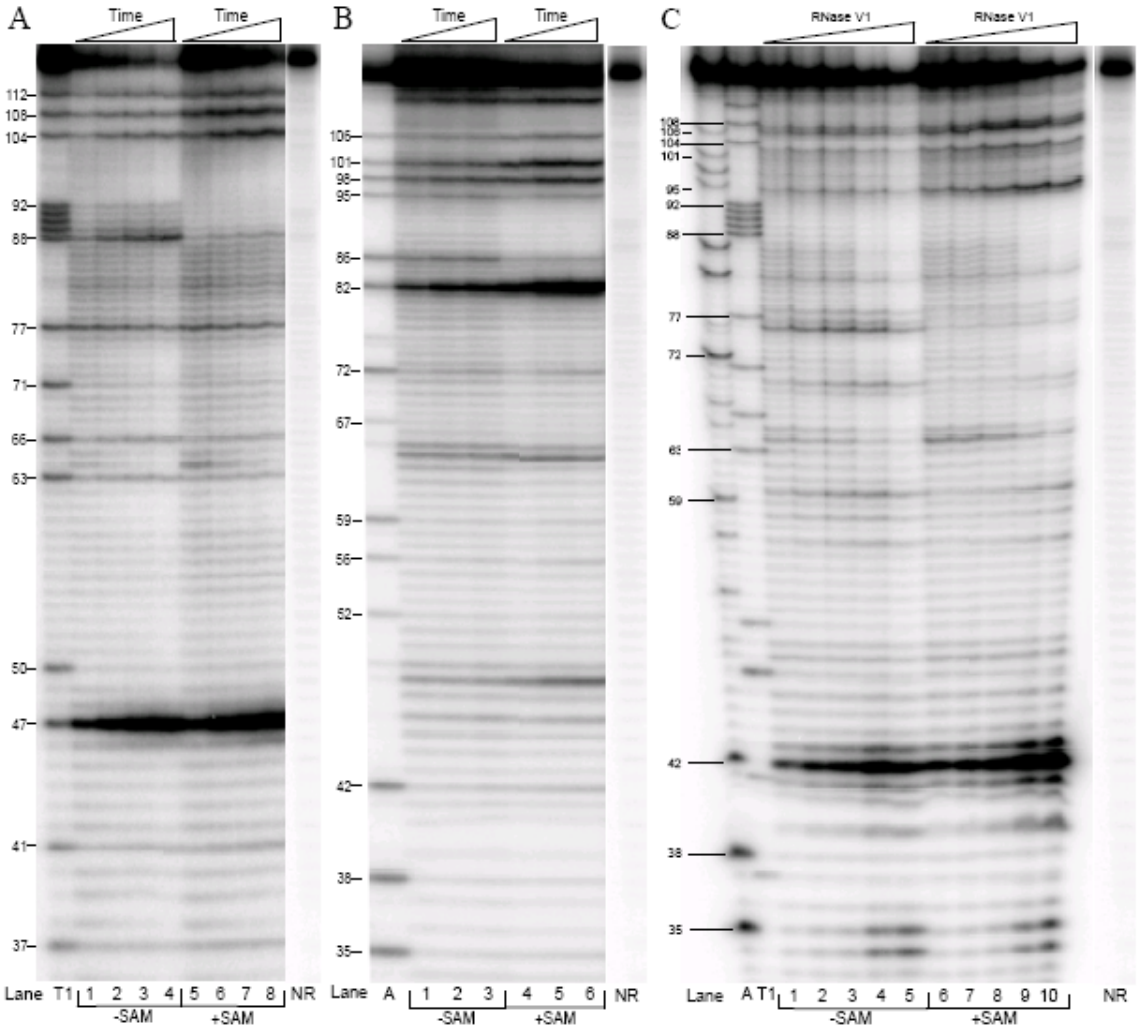

Figure 2.6. Structural mapping of the *E. faecalis* S_{MK} box RNA. RNA corresponding to positions 15-118 was generated by transcription with T7 RNAP, gel-purified, 5' end-labeled, denatured and refolded in the presence or absence of SAM, as indicated. (**A**) RNase T1 digestion. RNA was incubated with RNase T1 for 10 min (lanes 1 and 5), 20 min (lanes 2 and 6), 30 min (lanes 3 and 7), or 45 min (lanes 4 and 8). (**B**) RNase A digestion. RNA was incubated with RNase A for 10 min (lanes 1 and 4), 20 min (lanes 2 and 5), or 30 min (lanes 3 and 6). (**C**) RNase V1 digestion. RNA was incubated for 10 min with RNase V1 (5×10^{-6} U μl^{-1}, lanes 1 and 6; 8.3×10^{-6} U μl^{-1}, lanes 2 and 7; 1.3×10^{-5} U μl^{-1}, lanes 3 and 8; 2.5×10^{-5} U μl^{-1}, lanes 4 and 9; 6.7×10^{-5} U μl^{-1}, lanes 5 and 10). In **A-C**, size standards (lanes T1 and A) were generated by digestion of denatured RNA with RNases T1 and A, respectively. NR, RNA that was not nuclease-digested. Figure modified from (25).

The susceptibility of residue G88 to cleavage by RNase T1 was also used to quantitate the sensitivity of the S_{MK} box RNA to SAM. Incubation of the RNA with varying concentrations of SAM, followed by RNase T1 digestion, resulted in a concentration-dependent pattern of protection, with 50% protection of G88 at 0.9 μM SAM (Fig. 2.7). This value is within an order of magnitude of the concentration of SAM that promoted half-maximal transcription termination of most S box RNAs (87). The S box assay monitors a functional interaction with SAM, in the context of a nascent transcript emerging from *B. subtilis* RNAP, whereas the S_{MK} box assay monitors a structural change in T7 RNAP-generated RNA that has been denatured and refolded. Despite these differences, it seems that S box and S_{MK} box RNAs show a similar sensitivity to SAM concentrations.

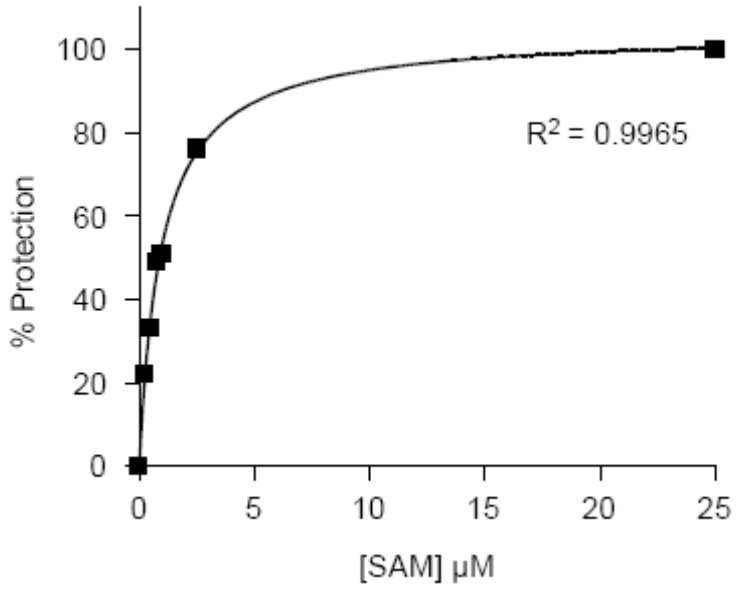

Figure 2.7. SAM-dependent protection from RNase T1 cleavage at G88 in *E. faecalis* S_{MK} box RNA. RNA corresponding to positions 15-118 was generated by transcription with T7 RNAP, gel-purified, 5' end-labeled, denatured and refolded in the presence of SAM (0, 0.25, 0.50, 0.75, 1.0, 2.5, or 25 µM) and incubated at room temperature for 5 min. The RNA was then digested with RNase T1 for 30 min at room temperature and products were resolved by PAGE. The intensity of the band at G88 in the absence of SAM was set as the 0% protection value and the intensity of the band in the presence of 25 µM SAM was set as the 100% protection value. A best fit line was generated by non-linear regression analysis to determine that the concentration of SAM at which there is 50% protection is 0.9 µM.

Digestion with RNase V1 (which preferentially cleaves adjacent to structured or stacked residues) resulted in increased cleavage in the presence of SAM at residue U95 (immediately downstream of the residues predicted to pair with the ASD sequence), consistent with a SAM-induced increase in structure in this region (Fig. 2.6C). An increase in cleavage in the region surrounding the AUG codon was also observed. Increased cleavage in the AUG region by all three RNases precludes a clear understanding of how this region is affected in the presence of SAM. Residue C75 showed higher cleavage by RNase V1 in the absence of SAM, suggesting that pairing in

this region (Fig. 2.1) occurs in the absence of SAM; a similar pattern was observed for residues U69 and U70, although cleavage was less efficient. The most efficient cleavage by RNase V1 occurred at residue U42, within the P3 helix; in conjunction with efficient RNase T1 cleavage of G47 (in the terminal P3 loop) in the presence or absence of SAM, this observation suggests that the structural arrangement of the P3 domain of the S_{MK} box is not strongly affected by SAM.

Overall, the results of structural probing with RNases support the predicted S_{MK} box model (Fig. 2.1). The results support the prediction that a portion of the internal loop of the S_{MK} box and the ASD sequence are unpaired in the absence of SAM, and binding of SAM promotes protection of the SD region, presumably through pairing with the ASD sequence.

2.3.4 Crystal structure of SAM bound S_{MK} box RNA

Using the mutational analysis data from 2.3.2 as a starting point, A. Smith in our lab developed a minimal *E. faecalis* S_{MK} box construct for X-ray crystallography (55). The ultimate goal of crystallography was to determine the location and structure of the SAM binding pocket. The minimal construct, termed $S_{MK}6$, consisted of residues 20-95 with the following modifications. Both the P3 and P4 helices were shortened and the terminal loops were replaced with GNRA tetraloops with the sequence GAAA to facilitate crystallography (Fig. 2.8A). The K_d for SAM measured by size-exclusion filtration of this minimal construct (0.57 μM) was similar to that of full-length RNA (0.85 μM) containing residues 15-118 (55). After confirmation of SAM binding, the

S_{MK}6 construct was sent to the lab of A. Ke at Cornell University for X-ray crystallography analysis.

The structure of S_{MK}6 RNA bound to SAM was determined at a resolution of 2.2 Å and the RNA was shown to fold into an inverted Y-shaped molecule (55, Fig. 2.8B). Most of the secondary-structure features match the predictions from phylogenetic analysis and RNase probing experiments, most importantly the SAM-dependent formation of the ASD-SD helix (P1 helix). SAM itself was shown to intercalate in a three-way helical junction between helices P1, P2 and P4. Of the 23 residues in the RNA that are 100% conserved across S_{MK} box sequences, 10 were found to participate in the formation of the SAM binding pocket. The P3 helix stacks on the P2 helix and the J3/2 bulge folds down to enclose the SAM binding pocket. The floor of the SAM binding pocket was determined to be defined by a crucial base-triple interaction between A73·G90-C25.

The crystal structure not only identified where SAM is located when bound to S_{MK} box RNA but also provided insights into how SAM is recognized by the RNA (55). SAM was found to make direct contacts with 5 residues that are 100% conserved (G26, G71, U72, A73, G90). The adenosine moiety of SAM stacks on top of G90 and adopts an energetically unfavorable *syn*-conformation. This conformation allows SAM to present the Watson-Crick face of the adenine moiety to form minor groove interactions with G26 and the Hoogsteen face to accept a hydrogen bond from an amine group on A73. The positive charge on the sulfonium ion of SAM is recognized through favorable electrostatic interactions with the O4 carbonyl group from U72 and a hydroxyl group from G71. A similar charge-stabilization scheme using the O4 of a uracil residue is

utilized by both the S box and SAM-II riboswitches (28, 62). In contrast to the well-defined adenosine moiety, no electron density was observed for functional groups beyond the sulfonium ion in SAM (55). In total, these results indicate that the S_{MK} box extensively recognizes the adenosine and sulfonium ion, but ignores the methionine tail. This is an unusual feature of the S_{MK} box as both the S box and SAM-II riboswitches extensively recognize the methionine tail (28, 62).

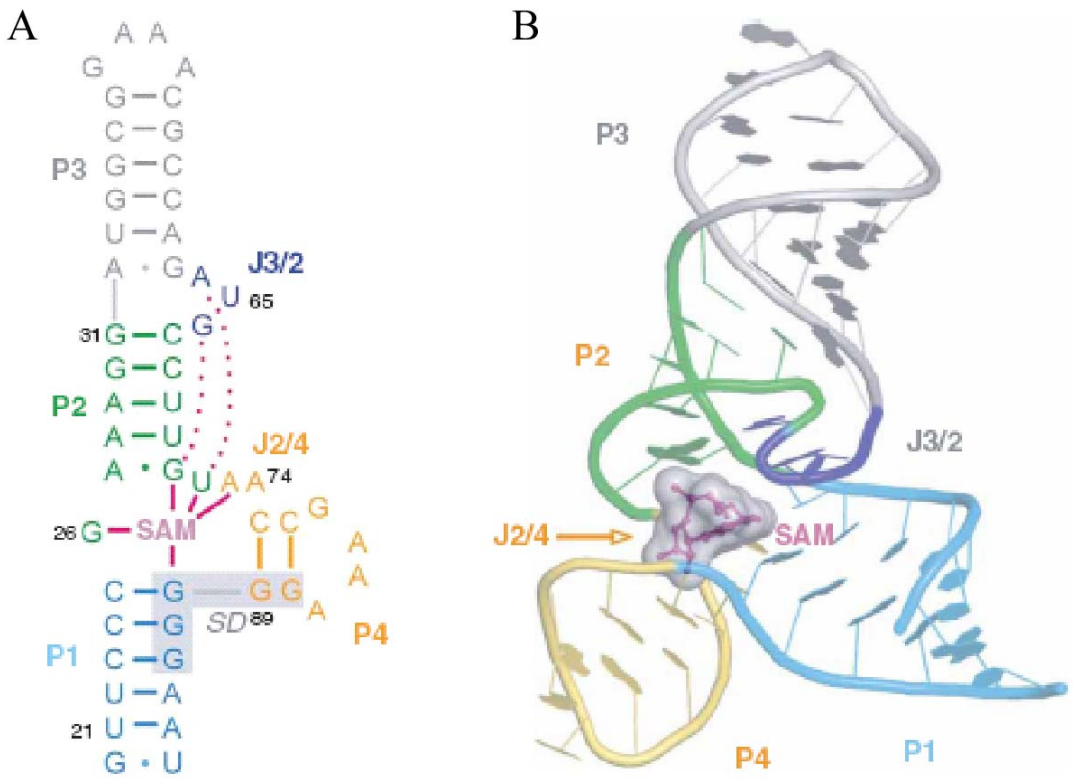

Figure 2.8. Crystal structure of the SAM-bound S_{MK} box riboswitch. (**A**) Secondary structure of the S_{MK}6 RNA construct based on the crystal structure. Helices P1 through P4 are colored in cyan, green, silver and yellow, respectively. Gray shading, SD sequence; solid magenta lines, direct contacts between the RNA and SAM; dashed magenta lines, tertiary interactions between J3/2 and P2 and J2/4. Numbering is consistent with Fig. 2.1. (**B**) Cartoon representation of the crystal structure of the S_{MK} box riboswitch. SAM is shown in overlapping CPK and surface representations in magenta and silver, respectively. The coloring scheme for the RNA is consistent with **A**. Figure and legend modified from (55).

2.3.5 Expression of *metK-lacZ* fusions in *Bacillus subtilis*

Despite the demonstrated SAM-dependent sequestration of the SD by the ASD in S_{MK} box RNA, there was still a question of whether this element could confer translational regulation *in vivo*. To investigate effects on translation independent of possible transcriptional regulation, we positioned the *E. faecalis metK* leader sequence downstream of the *B. subtilis glyQS* promoter and generated a *metK-lacZ* translational fusion in which *lacZ* expression was dependent on the *metK* RBS. The fusion was integrated as a single copy into the *B. subtilis* chromosome, and β-galactosidase levels were monitored during growth in the presence or absence of methionine. Growth in the presence or absence of methionine has been shown to result in either low or high SAM levels in *B. subtilis*, respectively (87, 151). The *B. subtilis* host was chosen for direct comparison of the S_{MK} box element to previously characterized S box elements under identical growth conditions. Expression of the *E. faecalis metK-lacZ* translational fusion was high during growth in the absence of methionine and was repressed 5-fold in the presence of methionine (Table 2.1), with a pattern similar to that observed for S box transcriptional fusions (87). The lower stringency of regulation of the *E. faecalis metK-lacZ* translational fusion as compared to most S box fusions may be due to higher activity from the *glyQS* promoter than the native promoter. Alternatively, the lower stringency may reflect the requirement for basal SAM synthetase activity even when methionine is abundant.

Table 2.1 Expression of *E. faecalis metK-lacZ* fusions in *Bacillus subtilis*

		B-galactosidase activity[a]			
Fusion[b]	**Type[c]**	**-Met**	**+Met**	**Ratio[d]**	**SAM binding[e]**
WT (15-116)	TNL	120 ± 19	24 ± 5.5	5.0	100
WT (15-116)	TNX	380 ± 28	340 ± 28	1.1	100
U22G	TNL	220 ± 27	160 ± 13	1.4	2.5
A93C	TNL	77 ± 8.7	42 ± 1.0	1.8	4.0
U22G+A93C	TNL	11 ± 3.7	2.5 ± 0.72	4.4	120
C24G	TNL	110 ± 14	95 ± 6.1	1.2	0.75
C24U	TNL	310 ± 0.71	200 ± 22	1.6	1.7
C24U+G91A	TNL	23 ± 2.9	15 ± 2.9	1.5	100
C25U	TNL	230 ± 25	80 ± 5.0	2.9	63
G90A	TNL	74 ± 5.2	74 ± 4.4	1.0	1.3
C25U+G90A	TNL	78 ± 19	82 ± 16	0.95	1.9
G30C+G31C	TNL	220 ± 33	160 ± 26	1.4	0.46
20-116	TNL	16 ± 2.8	5.0 ± 0.71	3.2	160
1-116	TNL	99 ± 14	83 ± 20	1.2	N/A

[a]Cells were grown in minimal medium (2) in the presence of methionine, then were split and grown in the presence (+Met) or absence (-Met) of methionine. β-Galactosidase activity is shown in Miller Units (104) for samples taken 4 h after the cultures were split. [b]Fusions consisted of either the wild-type (WT) or mutant *metK* leader region as indicated positioned downstream of the *B. subtilis glyQS* promoter and fused to *lacZ*. [c]TNL, translational fusion (expression dependent upon *metK* RBS); TNX, transcriptional fusion (expression dependent upon *lacZ* RBS). [d]The ratio is the β-galactosidase activity in cells grown in the absence of methionine divided by the activity in cells grown in the presence of methionine. [e]The percent SAM binding as compared to the wild-type *metK* sequence as determined in SAM binding assays (see 2.3.2). N/A, not determined.

In contrast to the *metK-lacZ* translational fusion, an analogous *metK-lacZ* transcriptional fusion, in which translation of the *lacZ* reporter is independent of the *metK* RBS, showed high constitutive expression (Table 2.1). The loss of repression of the transcriptional fusion during growth in methionine suggests that regulation of the *E. faecalis metK* gene is at the level of translation, as repression was observed only when expression was dependent upon the *metK* translation initiation region.

To verify our hypothesis that ASD-SD helix formation is responsible for translational inhibition, three sets of mutations that perturb the stability of the ASD-SD helix were evaluated. A U22G or A93C substitution is predicted to disrupt pairing between the ASD and SD and both substitutions resulted in a severe reduction in SAM binding and loss of repression during growth in methionine (Table 2.1). The U22G/A93C double mutation restores ASD-SD pairing, restores SAM binding, and was shown to restore regulation *in vivo*. Overall expression for this mutant was reduced relative to that of the wild-type construct; this may be due to an enhanced ASD-SD interaction from replacing an A-U pair with a G-C pair. Similarly, C24G or C24U substitutions in the ASD are predicted to disrupt pairing with G91 in the SD, and both substitutions resulted in a severe reduction in SAM binding and loss of regulation in response to methionine (Table 2.1). A C24U/G91A double mutation, which restores pairing at this position, restored SAM binding, but resulted in low expression and did not restore regulation *in vivo*. The low expression is likely to be due to disruption of the SD by the G91A mutation, which causes reduced affinity for 30S ribosomal subunits and therefore obscures the regulatory response. Overall, the results show that maintenance of

ASD-SD pairing is crucial for S_{MK} box-dependent gene regulation *in vivo* as disruption of pairing resulted in loss of regulation in all cases.

Mutations at two other residues predicted to participate in ASD-SD pairing, C25 and G90, were also incorporated into *lacZ* fusions. As described in 2.3.4, G90 in the SD sequence has a pivotal role in the S_{MK} box riboswitch as it organizes the SAM binding pocket through A73•G90-C25 base-triple formation and stacks directly beneath the SAM molecule. A C25U substitution creates a wobble pair with G90 that is tolerated at this position, as a *lacZ* fusion containing a C25U mutation showed partial regulation *in vivo*, consistent with a modest reduction of SAM binding *in vitro* (Table 2.1). The C25U/G90A double mutant is predicted to restore Watson-Crick pairing at this position, but not the crucial base-triple interaction with A73. This double mutant showed a complete loss of regulation *in vivo*, despite containing a canonical GGAGG SD sequence, in accordance with the loss of SAM binding *in vitro*. These data confirm the importance of G90 primary sequence conservation for SAM binding and SAM-dependent regulation.

Disruption of the P2 helix by incorporating G30C/G31C substitutions in the RNA severely affected SAM binding without changing the sequence in the ASD or SD region (Fig. 2.4). A *metK-lacZ* fusion containing the G30C/G31C substitutions was made to determine if the disruption of SAM binding conferred by the mutated sequence affected the regulatory response. These two point mutations resulted in loss of repression during growth in methionine (Table 2.1), which suggests that although ASD-SD pairing is necessary for regulation, ASD-SD pairing on its own is not sufficient. The RNA must also be able to bind SAM to be a functional regulator *in vivo*.

A fusion of a *metK* sequence truncated at the 5' end (deleting residues 15-19) to *lacZ* was also analyzed *in vivo*. Deletion of residues 15-19 is predicted to disrupt a predicted helix in the SAM-free form of the RNA, thus destabilizing the SAM-free conformation and promoting the formation of the SAM-bound conformation (Fig. 2.1). The truncated fusion produced a level of β-galactosidase expression that is at least 5-fold lower than that of the wild-type fusion under both high and low methionine conditions (Table 2.1). This supports our prediction that the truncated RNA favors the SAM-bound conformation in which the SD is sequestered by the ASD, resulting in lower expression. The truncated RNA still demonstrated regulation as expression was repressed during growth in the presence of methionine, but the stringency of regulation is lower for the truncated RNA despite the fact that this RNA exhibits higher SAM binding. The cause of this lower stringency is likely to be that the equilibrium between RNA conformations is shifted to the SAM-bound form in the truncated mutant even in the absence of SAM.

We were surprised to observe that addition of residues 1-14 to the *E. faecalis metK* fusion results in constitutive expression of *lacZ* in both high and low methionine conditions (Table 2.1). It is not clear why inclusion of these residues would cause this effect as they are not conserved, are not considered part of the S_{MK} box element, and preliminary results indicate that an RNA with these residues still binds SAM (data not shown). One possibility is that these residues serve as a binding site for an unknown factor in the heterologous *B. subtilis* host that is not present in *E. faecalis*, and binding of this factor disrupts or prevents SAM binding. Another possibility is that addition of these residues causes the S_{MK} box RNA to fold into an inactive state in *B. subtilis* and the inactive RNA is unable to sequester the SD and inhibit ribosome binding.

2.4 Discussion

This chapter reports the identification of the S_{MK} box, a new SAM-binding riboswitch RNA that regulates SAM synthetase genes in the Lactobacillales group of lactic acid bacteria. A variety of *in vitro* experiments showed that binding of SAM to S_{MK} box RNA results in a rearrangement of the RNA structure that sequesters the SD sequence. Structural mapping and mutagenesis analyses indicate that the ASD and SD regions are unpaired in the absence of SAM, and that ASD-SD pairing is required for SAM binding. No factors other than SAM were necessary for SAM binding or the SAM-dependent structural change. Also, the *E. faecalis* S_{MK} box was sufficient to confer translational repression of a *lacZ* reporter *in vivo* in *B. subtilis* under conditions when SAM pools are elevated. These data support a model in which SAM binding to the S_{MK} box RNA represses *metK* expression through inhibition of translation initiation by blocking access of the ribosome to the ribosome binding site. Additional testing of this model, including analysis of 30S subunit binding to S_{MK} box RNA, is described in Chapter 3.

The S_{MK} box riboswitch shows no similarity of primary sequence, secondary structure, or tertiary structure to the S box or SAM-II SAM-binding riboswitches. A major difference between the S_{MK} box and S box RNAs is that the SAM-binding aptamer domain of the S_{MK} box RNAs includes the SD sequence, which serves as the expression platform. In contrast, the SAM-binding domain of S box RNAs does not include the expression platform (the terminator helix) and includes only the 5' portion of the antiterminator element. Certain S box RNAs are predicted to regulate gene expression at the level of translation initiation rather than transcription termination (156). In genes of

this class, the terminator helix is replaced by a helix that sequesters the SD by pairing it with a complementary ASD sequence, and the antiterminator helix is replaced by a helix that sequesters the ASD sequence by pairing it with another competing sequence, the anti-ASD or AASD. Although no S box RNAs of this type have been characterized biochemically, it seems, by analogy with other S box RNAs, that the role of SAM binding is to prevent formation of the ASD-sequestering helix (AASD) without having a direct effect on the SD sequence itself. Other metabolite-binding RNA riboswitches that operate at the level of translation are arranged similarly to S box RNAs (41, 71, 156), so that direct participation of the regulatory target in the metabolite-binding element is an unusual feature of S_{MK} box RNAs. Although at least one member of the SAM-II riboswitch class also involves the SD sequence in part of the SAM binding domain, the regulatory mechanism of SAM-II RNAs has yet to be characterized (28).

The concentration of SAM required for protection of the *E. faecalis* S_{MK} box SD sequence from RNase digestion (<1 μM) is similar to that required for SAM-dependent transcription termination *in vitro* for most S box RNAs (59, 87). SAM concentrations have been reported to range from <50 to 400 μM for *B. subtilis in vivo*, depending on growth conditions, but it is not known what fraction of the SAM pool is in complex with other cellular components (87, 151). S_{MK} box and S box RNAs also show a similar ability to discriminate between SAM and related compounds such as SAH, which differs from SAM by a single methyl group. As SAH is the byproduct of the use of SAM as a methyl donor, discrimination between SAM and SAH is of crucial biological importance. In contrast, S_{MK} box RNA fails to discriminate against a number of SAM analogs that do not bind to S box RNA (A. M. Smith, F. J. Grundy and T. M. Henkin, in preparation).

Given that the S box and S_{MK} box riboswitches show no overall sequence or structural similarity, the major differences between them demonstrate that there are multiple ways to build RNA elements with similar biochemical and regulatory properties.

A variety of mechanisms for regulation of methionine metabolism genes have been uncovered in bacteria. Most Gram-positive bacteria with low G+C content, including certain members of the Lactobacillales, use the S box system to control expression of both methionine biosynthesis and *metK* genes at the level of transcription termination in response to SAM (30, 78, 143, 162). A T box motif predicted to monitor charging of $tRNA^{Met}$ has been identified upstream of methionine-related genes in *Enterococcus*, *Lactobacillus* and *Staphylococcus* spp., but *metK* genes are not included in this group (162, F. J. Grundy and T. M. Henkin, unpublished data). In *Bacillus* spp., the *cysE* cysteine biosynthesis gene is also a member of the T box family, responding to $tRNA^{Cys}$ (163), whereas other cysteine-related genes are regulated at the level of transcription initiation (164, 165). Cysteine and methionine biosynthesis genes in *Lactococcus lactis* are regulated at the level of transcription initiation by a LysR-type activator, FhuR (also called CmbR), which responds to *O*-acetyl serine (84, 166), and a similar mechanism using the MtaR protein has been suggested for *Streptococcus* spp. (78, 167). Regulation at the level of transcription initiation is also used in *E. coli*, with separate regulators for cysteine and methionine genes; *metK* is coordinately regulated with methionine biosynthesis genes, using SAM as the effector (168). It therefore seems that at least some of the lactic acid bacteria differ from other organisms both in coordinating regulation of cysteine and methionine biosynthesis genes and in separately regulating *metK* expression. Although the generality of these patterns requires further

investigation, it is notable that many lactic acid bacteria require an exogenous source of methionine for optimal growth and readily interconvert methionine and cysteine. SAM synthetase activity is essential for cell viability under all growth conditions, and independent regulation of *metK* in response to SAM may ensure that appropriate levels of SAM are maintained regardless of fluctuations in methionine availability. A further potential advantage of a translational control mechanism is that the switch could be reversible, allowing rapid response to variations in SAM concentration by release of the SD when SAM pools decline. The potential for the reversibility of S_{MK} box-mediated regulation is discussed in Chapter 4.

CHAPTER 3

ANALYSIS OF 30S SUBUNIT INTERACTION WITH S_{MK} BOX RNA

3.1 Introduction

In the previous chapter, the *E. faecalis* S_{MK} box was shown to bind SAM *in vitro* and confer translational repression of a *lacZ* reporter in *B. subtilis* under conditions when SAM pools are elevated. These findings suggested that the S_{MK} box RNA represents a new class of SAM-responsive riboswitch that regulates *metK* expression at the level of translation initiation by inhibiting ribosome binding to the RNA when SAM is present. However, there was no direct evidence that the apparent sequestration of the SD sequence is sufficient to inhibit ribosome binding. It is possible for binding of other regulatory molecules, including proteins and antisense RNAs, directly to or near the RBS to inhibit translation initiation by preventing ribosome binding (152, 153). Alternatively, regulatory molecules can bind to an mRNA and "trap" the 30S subunit on the mRNA in an inactive initiation complex (154, 155).

Nitrocellulose filter binding and ribosomal toeprinting assays were utilized in order to test whether SAM inhibits the binding of 30S subunits to S_{MK} box RNA in the absence of any additional factors. The results show that the *metK* RNA start codon is properly positioned in the P site of the 30S subunit in the absence of SAM.

Preincubation of the RNA with SAM reduced 30S subunit binding and resulted in inhibition of processivity of reverse transcriptase near the SD region. This supports our previous observation that this segment of the RNA becomes more structured in the presence of SAM (See 2.3.3). These findings confirm our model for S_{MK} box regulation, where SAM-promoted ASD-SD pairing directly inhibits translation initiation by blocking access of the ribosome to the RBS.

3.2 Materials and methods

3.2.1 Nitrocellulose filter binding assay

T7 RNAP-transcribed RNAs corresponding to positions 15-118 of the *E. faecalis metK* transcript were 5' end-labeled with [γ-^{32}P]ATP (7,000 Ci mmol^{-1}) by using a KinaseMax kit (Ambion) and passed through a MicroSpin G-50 column (Amersham Biosciences, Piscataway, NJ) to remove excess [γ-^{32}P]ATP. Labeled RNAs (10 nM) in 1X binding buffer (10 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 60 mM NH_4Cl, and 6 mM 2-mercaptoethanol) were heated to 65ºC and slow-cooled to 40ºC. SAM (160 μM), SAH (160 μM), or water was added, and the samples were incubated for 10 min at 37ºC. 30S ribosomal subunits (60 nM) isolated from *E. coli* MRE600 by using a high-salt wash and provided by K. Fredrick (Ohio State University) were added, followed by a 30 min incubation at 37ºC (148). $tRNA^{fmet}$ (120 nM; Sigma-Aldrich, St. Louis, MO) was included as indicated. Samples were then loaded onto a nitrocellulose filter (0.45 μM pore size; Whatman, Clifton, NJ) that had been soaked in 1X binding buffer. A vacuum was applied and filters were washed five times with 300 μl of 1X binding buffer.

Radioactivity retained by the filters was measured in a Packard Tri-Carb 2100TR liquid scintillation counter. Each experiment was repeated at least twice.

3.2.2 Primer extension inhibition (toeprint) assay

A DNA primer complementary to positions 180-203 of the *E. faecalis metK* transcript was 5' end-labeled as described above, except that a G-25 column was used to remove excess [γ-^{32}P]ATP. The labeled oligonucleotide (10 nM) was annealed to T7 RNAP-transcribed RNA (10 nM) corresponding to positions 15-208 of the *metK* transcript in 1X binding buffer by heating to 65ºC and slow cooling to 40ºC. SAM (160 μM), SAH (160 μM), or water was added, and the samples were incubated for 10 min at 37ºC. 30S ribosomal subunits (60 nM; obtained as described above) were added, with or without $tRNA^{fmet}$ (120 nM; Sigma-Aldrich), followed by a 30 min incubation at 37ºC. Avian myeloblastosis virus reverse transcriptase (1 unit per rxn, Thermoscript RT-PCR; Invitrogen, Carlsbad, CA) and dNTPs (375 μM) were added followed by 10 min incubation at 37ºC. Reaction products were resolved by using 10% denaturing PAGE, visualized by using PhosphorImager analysis (Molecular Dynamics, Sunnyvale, CA) and quantified by using ImageQuant 5.2 software. A DNA sequencing ladder was generated by using a DNA Sequenase 2.0 Kit (USB), a DNA template containing positions 15-220 of the *E. faecalis metK* gene, and the same downstream primer as was used in the primer extension assays.

3.2.3 SAM binding assay

SAM binding assays were performed as described in Chapter 2.

3.3 Results

3.3.1 Effect of SAM on 30S subunit binding to S_{MK} box RNA

Nitrocellulose filter binding assays were performed to determine whether the proposed SAM-promoted pairing of the ASD to the SD in *E. faecalis* S_{MK} RNA is sufficient to inhibit binding of 30S ribosomal subunits. This assay took advantage of the ability of nitrocellulose filters to bind 30S subunits but not free nucleic acids. Highly purified *E. coli* 30S subunits were utilized in the assays and these heterologous subunits were selected because the S_{MK} box is absent in *E. coli*, reducing the probability that any additional cellular factors potentially required for regulation would co-purify with the ribosomal subunits. End-labeled *in vitro*-generated S_{MK} box RNA was heated, slow-cooled, and incubated in the presence or absence of SAM. *E. coli* 30S subunits were added in the presence or absence of $tRNA^{fmet}$, the samples were passed through a nitrocellulose filter, and the amount of radiolabeled RNA retained by the filter was determined by scintillation counting.

Incubation of RNA corresponding to positions 15-118 relative to the *E. faecalis metK* transcription start-site with 30S ribosomal subunits resulted in ~30% retention of the RNA, whereas <0.03% retention was observed when ribosomal subunits were not added (Fig. 3.1A). Preincubation of the RNA in the presence of SAM resulted in a 4-fold reduction in retention of the RNA by 30S subunits, whereas the addition of SAH had no effect. Similar results were obtained when the RNA was preincubated with 30S subunits before the addition of SAM (Fig. 3.1B), suggesting that the RNA-ribosome complex can be disrupted by the addition of SAM. Retention of the RNA increased 2-fold when $tRNA^{fmet}$ was included, presumably because the tRNA stabilizes the interaction of the

30S subunit with *metK* RNA. Incubation of the RNA with SAM before the addition of 30S subunits and $tRNA^{fmet}$ reduced retention of RNA ~2-fold, but SAM did not significantly reduce retention of the RNA when both 30S subunits and $tRNA^{fmet}$ were added prior to the addition of SAM (Fig. 3.1B). The loss of the effect of SAM is not a result of SAM preferentially binding to $tRNA^{fmet}$ over the *metK* RNA because the tRNA did not retain a measurable amount of SAM in a binding assay (Fig. 3.2A). These results suggest that SAM can compete with binding of 30S subunits to the RNA, but the addition of the initiator tRNA results in a stable complex that is resistant to challenge by SAM and is therefore committed to translation.

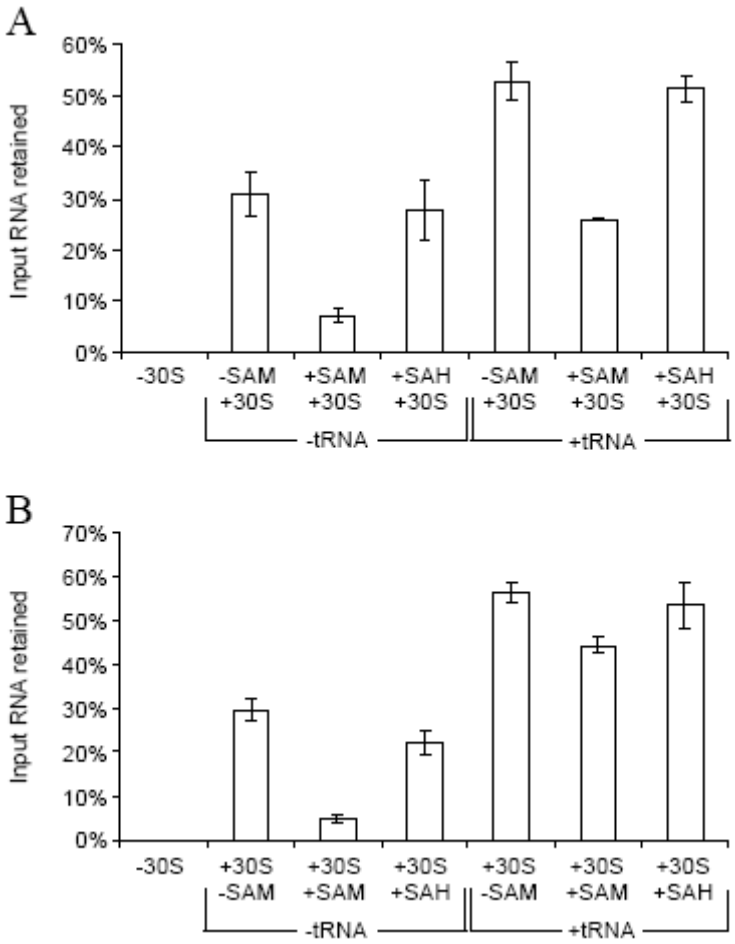

Figure 3.1. Nitrocellulose filter binding assays. *E. faecalis metK* RNAs containing positions 15-118 were generated by T7 RNAP transcription, gel-purified, and end-labeled with [γ-^{32}P]ATP. Labeled RNAs were incubated with SAM or SAH followed by tRNA and 30S subunits (**A)** or tRNA and 30S subunits followed by SAM or SAH (**B**) as indicated. Samples were passed through a nitrocellulose filter, and radioactive material retained by the filter was quantified by scintillation counting. Less than 0.03% of input RNA was retained by the filters in the absence of 30S subunits.

The effect on ribosome binding of mutations in the ASD-SD pairing region was also tested. The U22G substitution in the ASD, which was predicted to disrupt pairing with residue A93 immediately downstream of the SD (positions 88-92), resulted in a 40-fold reduction of SAM binding *in vitro* (Fig. 3.2A). The U22G mutation also resulted in complete loss of SAM-dependent inhibition of 30S subunit binding (Fig. 3.2B). The

A93C mutation, which is predicted to disrupt ASD-SD pairing without affecting the SD sequence itself, similarly resulted in loss of SAM binding *in vitro* (Fig. 3.2A). Combination of the U22G and A93C substitutions, which is predicted to restore pairing, restored both SAM binding and the sensitivity to SAM in the filter binding assays (Fig. 3.2), indicating that the A93C mutation could suppress the defect in response to SAM conferred by the U22G substitution. These results show that the reduction of 30S subunit binding in the presence of SAM occurs only when the RNA can bind SAM. The increase in 30S subunit binding exhibited by the U22G mutant RNA in the absence of SAM compared with either the wild-type or U22G/A93C double mutant (Fig. 3.2B) suggests that ASD-SD pairing occurs to some extent even in the absence of SAM and that SAM binding shifts the equilibrium to the paired form. These experiments support the model that SAM-promoted ASD-SD pairing can inhibit 30S subunit binding to S_{MK} box RNA.

Lactobacillus plantarum metK RNA, which also contains an S_{MK} element (Fig. 1.8), showed the same pattern of 30S subunit binding as the *E. faecalis metK* RNA, although the binding efficiency was somewhat lower (Fig. 3.2B). This result indicates that other S_{MK} box RNAs exhibit a similar effect of SAM on ribosome binding.

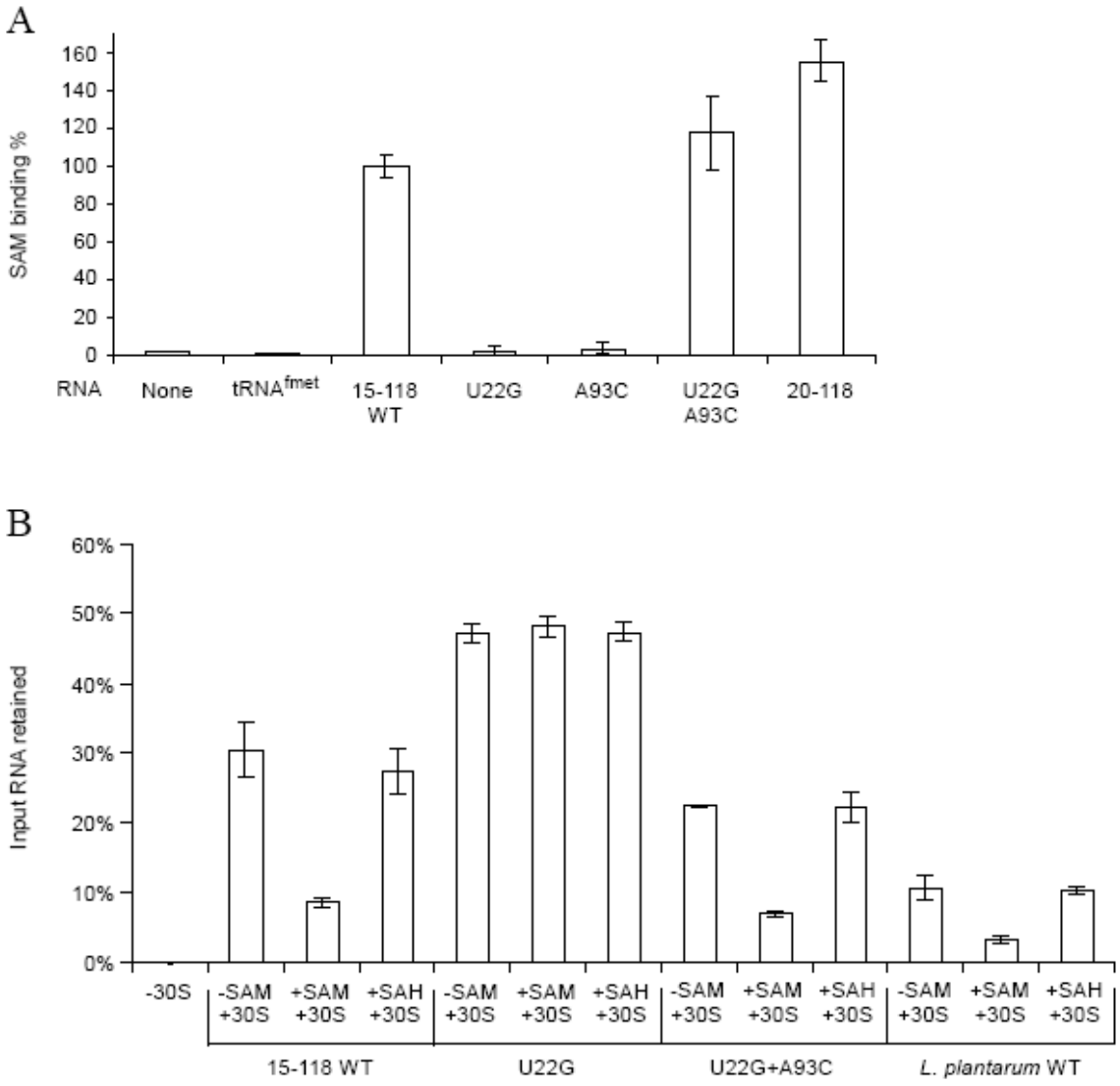

Figure 3.2. Effect of S_{MK} box mutations on SAM binding and 30S subunit binding. (**A**) SAM binding. RNAs generated by T7 RNAP transcription were gel-purified and then incubated in the presence of [methyl-^{14}C]SAM. The RNA-bound SAM was separated from unbound SAM by size-exclusion filtration. Retention of SAM is expressed relative to that of wild-type *E. faecalis metK* RNA extending from residues 15-118. (**B**) Nitrocellulose filter binding assays. RNAs end-labeled with [γ-^{32}P]ATP were incubated with SAM or SAH followed by tRNA and 30S subunits. Samples were passed through a nitrocellulose filter, and radioactive material retained by the filter was quantified by scintillation counting. Less than 0.03% of input RNA was retained by the filters in the absence of 30S subunits.

3.3.2 Position of S_{MK} box RNA within a bound 30S subunit

Although the filter binding studies indicated that 30S subunits bound to *E. faecalis metK* RNA, there was still a question as to whether binding was localized to the ribosome binding site (RBS). Primer extension inhibition ("toeprint") assays were used to localize the position of the 30S subunit on the RNA and assess the effect of SAM (149). A DNA primer complementary to a region of the *E. faecalis metK* RNA downstream of the SD was annealed to the RNA and then extended by reverse transcriptase. Extension by reverse transcriptase is inhibited when the enzyme encounters a bound obstacle (e.g., the 30S subunit) or significant secondary structure. The extension products were resolved by PAGE and identified by comparison to a DNA sequencing ladder.

Reactions containing the *E. faecalis* wild-type S_{MK} RNA (containing residues 15-208 relative to the transcription start site) resulted in two extension products, one corresponding to a halt at position G77 and the other resulting from full-length extension to the 5' end of the RNA (Fig. 3.3, lane 1). The G77 product was observed in all reactions and is probably due to secondary structure in the core region of the RNA (Fig. 2.1). A 5' truncated RNA that began with residue 20, which retained SAM binding (Fig. 3.2A), also exhibited the G77 product (data not shown); this indicates that the conserved predicted pairing between positions 15-20 and 68-75 (Fig. 2.1), which is absent in the 20-208 RNA, is dispensable for SAM binding and is not responsible for the G77 extension product.

Incubation of the 15-208 RNA in the presence of SAM resulted in a new extension product corresponding to position A94 (Fig. 3.3, lane 2), which is two nt 3' of

the end of the SD (residues 88-92) and comprises the first position of the predicted ASD-SD pairing (Fig. 2.1). Studies with avian myeloblastosis virus reverse transcriptase have shown that termination can occur when the enzyme encounters secondary structure, and the majority of these truncated products correspond to the first paired base in the structure or the last position before a cross-link (11, 13). Incubation of the RNA in the presence of SAH instead of SAM did not result in formation of the A94 product (Fig. 3.3, lane 3). The A94 extension product was also absent in reactions using the U22G mutant RNA (which did not bind SAM) and was present in the U22G/A93C double mutant in which both ASD-SD pairing and SAM binding were restored (Fig. 3.3, lanes 10, 13). These observations support the conclusion that the A94 product is dependent upon the RNA-SAM binding interaction and that the A94 product is likely to correspond to ASD-SD pairing.

The addition of 30S subunits to a reaction containing the 15-208 RNA resulted in a set of low-abundance products clustered around U116-U122 (Fig. 3.3, lane 4). If both 30S subunits and $tRNA^{fmet}$ were included, ~90% of the primer extension products corresponded to position U118, which is 16 nt downstream of the A in the AUG start codon. This distance is representative of the position at which reverse transcriptase stops when the start codon of an mRNA is positioned in the P site of the 30S subunit (150), and the predominance of this product in reactions containing both 30S subunits and $tRNA^{fmet}$ is consistent with stabilization of the initiation complex in the presence of the tRNA. The abundance of extension products corresponding to U118 was reduced ~2-fold in the presence of SAM, whereas SAH had no effect. These results indicate that *E. coli* 30S

subunits bind to the *E. faecalis metK* RBS and that ribosome binding is inhibited in the presence of SAM.

In contrast to what was observed for the wild-type RNA, the addition of SAM had no effect on the abundance of the U118 product for the U22G mutant RNA (Fig. 3.3, lanes 9-11), which is defective in SAM binding. Introduction of the compensatory A93C mutation, which suppressed the SAM binding defect of the U22G mutation (Fig. 3.2A), resulted in restoration of sensitivity to SAM (Fig. 3.3, lanes 12-14), providing further support for the model that ASD-SD pairing is required for both SAM binding and SAM-dependent inhibition of ribosome binding. This confirms the results from the filter binding assays where SAM inhibited binding of 30S subunits to S_{MK} box RNA only when the RNA could bind SAM.

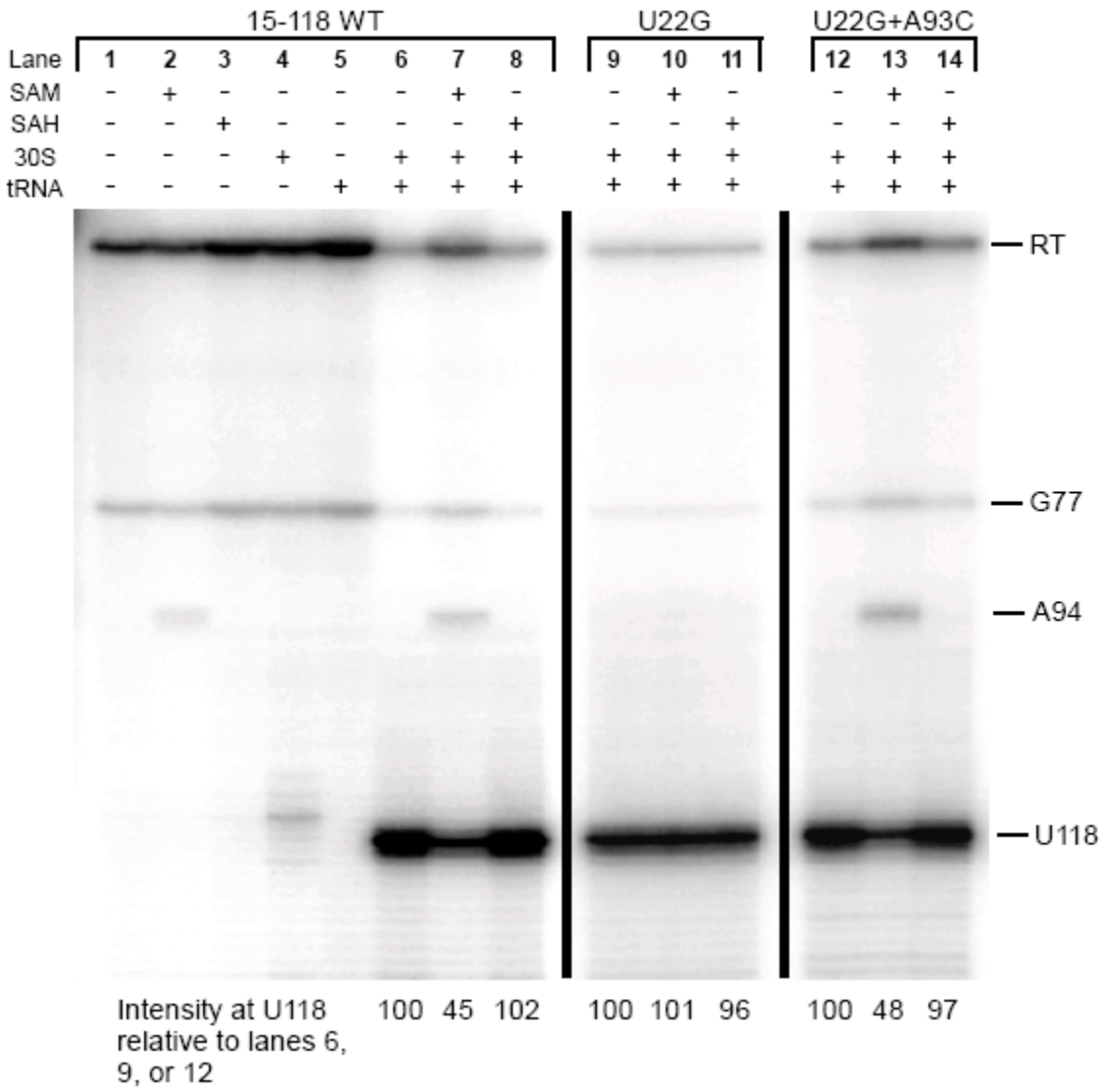

Figure 3.3. Primer extension inhibition analysis of S_{MK} box RNA. *E. faecalis metK* RNAs containing positions 15-208 were generated by T7 RNAP transcription, gel-purified, and then annealed to a [γ-^{32}P]ATP-labeled DNA primer complementary to positions 180-203. Annealed RNAs were incubated in the presence of SAM or SAH followed by $tRNA^{fmet}$ and 30S subunits. dNTPs and reverse transcriptase were added, and reactions were stopped by the addition of gel-loading buffer. Radioactivity at the position of the U118 stop in lanes 7 and 8 was compared with the value for lane 6, which was normalized to 100. Lanes 10 and 11 were compared with lane 9, and lanes 13 and 14 were compared with lane 12 in a similar manner. RT, readthrough to 5' end.

3.4 Discussion

The data presented in this chapter show that *E. coli* 30S subunits bind to S_{MK} box RNA in the absence of other factors and that SAM specifically inhibits ribosome binding to the RNA. Primer extension analysis demonstrated that in the absence of SAM, the 30S subunit is positioned with the AUG start codon of the *E. faecalis metK* RNA within the ribosomal P site, whereas the addition of SAM resulted in a significant reduction of the ribosomal toeprint. Under these conditions, the extension proceeded through the ribosomal toeprint site and yielded a new product corresponding to the 3' end of the ASD-SD pairing (P1 helix). These findings further support the model that SAM binding promotes pairing of the ASD to the SD, which prevents translation initiation by blocking access of the ribosome to the RBS.

Most studies on the mechanism of action of riboswitches have focused on transcription termination systems. In systems of this type, binding of the effector to the riboswitch RNA causes the RNA to form an intrinsic transcriptional terminator or a competing antiterminator element. Structural probing, *in vitro* transcription, and mutational analysis have provided support for this mode of action for several of these riboswitch systems (reviewed in 156, 157). By contrast, most translational riboswitches have been analyzed in less detail, with most of the experimental work focused on regulation *in vivo* and effector-dependent modulation of RNA structure *in vitro*. Primer extension assays of the *E. coli btuB* riboswitch showed that the ability of 30S subunits to bind *btuB* RNA is reduced when adenosylcobalamin is present, providing a clear indication that regulation occurs at the level of translation initiation (70). One caveat is that the ribosomes used in the *btuB* experiments were prepared with a low-salt wash,

which does not remove all ribosome-associated factors. Binding was not observed with ribosomal subunits prepared with a high-salt wash, which was interpreted by the authors to indicate that other factors could be involved in translational control of *btuB* RNA (70). The work presented in this chapter demonstrates a SAM-dependent effect on binding of 30S subunits prepared with a high-salt wash from a heterologous host in which the S_{MK} box mechanism is not found. This suggests that for S_{MK} box RNA, the observed effect of SAM on ribosome binding does not require additional cellular factors. The S_{MK} box is therefore the first translational control riboswitch for which the translational mechanism is clearly established by biochemical analysis.

Multiple mechanisms of translational regulation have been identified in bacteria. For example, the *E. coli* CsrA protein regulates the expression of *cstA* by directly binding the *cstA* RBS and preventing ribosome binding (152). A similar mechanism is used by the *trans*-acting OxyS RNA, which pairs with the RBS of the *fhlA* mRNA (153). Thermosensor RNAs use intramolecular interactions that sequester the RBS at low temperatures and expose the RBS at higher temperatures (See 1.1.2). In contrast, regulation of the ribosomal proteins encoded by the α operon in *E. coli* results from "entrapment" of the 30S subunit in an unproductive complex by binding of the S4 repressor protein to the mRNA (155). Two examples of a "typical" metabolite-binding riboswitch translational mechanism are found in the *ypaA* gene in *B. subtilis* and the *thiM* gene in *E. coli* (See 1.4.2 and 1.4.3, respectively). In each case, the effector molecule, flavin mononucleotide for *ypaA* and thiamine pyrophosphate for *thiM*, binds to an aptamer domain in the 5' untranslated region of the mRNA. This binding causes a downstream structural shift that sequesters the RBS (the expression platform) and

sequestration of the RBS was proposed to block ribosomal access; biochemical analyses were later performed to prove the proposed mechanism for *thiM* (74). The S_{MK} box also uses effector-dependent occlusion of the RBS to inhibit translation initiation but differs from these other systems in that the expression platform is an intrinsic part of the aptamer domain, so that disruption of ASD-SD pairing results in loss of SAM binding (See 2.3.2). The S_{MK} riboswitch therefore represents a special class of metabolite-binding regulatory RNAs.

The majority of riboswitches identified in low G+C Gram-positive organisms appear to regulate gene expression by premature transcription termination, whereas riboswitches from high G+C Gram-positive and Gram-negative organisms tend to regulate at the level of translation (34, 90, 156, 158). This divergence is readily observed with riboswitch elements that are found in both groups of organisms but regulate by different mechanisms depending on the host (e.g., the lysine-binding L box riboswitch; 32). The S_{MK} box, which is found only in the Lactobacillales group of low G+C Gram-positive bacteria, is unusual in that it appears to function only at the level of translation. The experiments in this chapter were performed with *E. coli* components, whereas *in vivo* experiments in Chapter 2 were performed in *B. subtilis*. These results show that S_{MK} box riboswitches have the potential to function as gene regulators in a variety of non-native organisms, including both Gram-positive and Gram-negative species.

Experiments in this chapter showed that inhibition of 30S subunit binding to S_{MK} box RNA was observed regardless of whether SAM was added before or after the ribosomal subunits. However, the addition of the initiator tRNA to the complex resulted in a stable complex that was resistant to SAM, indicating that the complex was now

committed to translation. This suggests that competition between the ribosome and SAM for binding to the mRNA occurs at the initial stages of formation of the initiation complex. We hypothesize that one possible benefit of regulating in this manner is that regulation could be reversible as the cell could quickly respond to a sudden drop in SAM levels by release of SAM from S_{MK} box mRNAs, allowing access to the RBS and expression of the gene. The importance of a rapid response may reflect the fact that *metK* is an essential gene for which expression is required at some level during all growth conditions. Our hypothesis that the S_{MK} box represents a true reversible switch required further analysis because reversibility depends on the maintenance of a stable pool of the *metK* transcript. Experiments to test this hypothesis are described in Chapter 4.

CHAPTER 4

ANALYSIS OF THE S_{MK} BOX IN *Enterococcus faecalis*

4.1 Introduction

The work presented in previous chapters supports the model that the S_{MK} box riboswitch regulates *metK* gene expression at the level of translation in response to SAM. In Chapter 2, it was shown that S_{MK} box RNA binds SAM and that SAM binding causes sequestration of the SD region. Additionally, the S_{MK} box sequence was shown to regulate the expression of a *lacZ* reporter gene at the level of translation. In Chapter 3, it was shown that SAM alone is sufficient to inhibit binding of 30S ribosomal subunits to S_{MK} box RNA *in vitro*. All of the work in the previous chapters was performed either *in vivo* in *B. subtilis* or *in vitro* with *E. coli* components; neither of these organisms utilize the S_{MK} box mechanism. As a result, it was not known if the S_{MK} box functions in a similar manner in a native organism. It is possible that a different host environment could significantly modify the regulatory response (e.g., the stringency of regulation). Thus, it is important to investigate the properties of the S_{MK} box in a native organism in order to better understand its role in regulation of *metK* expression.

In this chapter we investigate the regulatory properties of the *E. faecalis metK* S_{MK} box riboswitch in the native organism. Since the S_{MK} box is predicted to regulate

gene expression in response to SAM, the cellular pool of SAM was measured to determine if SAM pools decrease when *E. faecalis* cells are starved for methionine, as was previously observed in *B. subtilis* (87). The ability to manipulate SAM pools allowed us to monitor the effect of changing SAM levels on *metK* transcript abundance and stability, and *metK* expression. The stability of the *metK* transcript was measured *in vivo* to investigate potential regulation at the level of transcript stability. Experiments were also performed that monitored the levels of SAM synthetase and expression of a *metK-gusA* fusion in *E. faecalis* under methionine starvation conditions. The results of these experiments allow us to make observations about the properties of the *metK* transcript *in vivo* and the regulatory response of the S_{MK} box in a native organism.

4.2 Materials and methods

4.2.1 Bacterial strains, plasmids and growth conditions

The *E. faecalis* strains OG1 (160), OG1RF (Rif^R, Fua^R; 161), and CK111 (pCF10-101, Spc^R, Tet^R; 48) were used in these experiments. *E. faecalis* cells were grown in BHI (Difco) or CDM minimal medium (32 mM K_2HPO_4, 17 mM KH_2PO_4, 6 mM $(NH_4)_2SO_4$, 160 µM $MgSO_4$, 0.2% glucose, 230 µM methionine, 35 µg/mL of the 19 other common L-amino acids, 100 µg/mL thiamine, 20 µg/mL adenosine, 20 µg/mL cytosine, 20 µg/mL uracil, 20 µg/mL biotin, 20 µg/mL Ca-pantothenate, 20 µg/mL pyridoxine, 5 µg/mL guanosine, 2 µg/mL nicotinic acid, 2 µg/mL riboflavin, 0.2 µg/mL folic acid, and 0.005% yeast extract; 66). Prior to growth of *E. faecalis* in CDM, an overnight starter culture was made that consisted of 1 part BHI and 4 parts CDM. Overnight cultures were diluted 10-fold in fresh CDM and were incubated at 37ºC with shaking. *E. coli* strain EC1000

(Kan^R; 50) was grown in BHI at 37ºC. When required for selective growth of *E. faecalis*, erythromycin (Erm) and tetracycline (Tet) were used at 10 μg/mL; rifampicin (Rif) at 200 μg/mL; spectinomycin (Spc) at 1 mg/mL. When required for selective growth of *E. coli*, Erm was used at 100 μg/mL. Plasmid pCJK141, a derivative of pCJK47 (Erm^R; 48), was used for markerless genetic exchange.

4.2.2 Measurement of SAM pools

E. faecalis OG1 cells were grown until mid-log phase in CDM minimal medium containing methionine (230 μM) and yeast extract (0.005%). Cells were harvested by centrifugation and resuspended in CDM containing low methionine (1.8 μM). Samples were collected by filtration at the indicated timepoints and extracted with formic acid (0.5 M), and the formic acid was removed from the samples by lyophilization as described (72). Cell extracts were tested in a single round *in vitro* transcription termination assay using a P_{gly}-*yitJ* S box riboswitch template and compared to a SAM standard curve as previously described (87). Each experiment was repeated at least twice.

4.2.3 Total RNA extraction

E. faecalis OG1 cells were grown until mid-log phase in CDM minimal medium containing methionine (230 μM) and yeast extract (0.005%). Cells were harvested by centrifugation and resuspended in fresh CDM containing either low (1.8 μM) or high (230 μM) methionine. Samples were taken at 1 h intervals and added to 2 volumes of RNAprotect™ bacteria reagent (Qiagen). Following 5 min incubation at room temperature, samples were centrifuged and the cell pellet was resuspended in 200 μL of

treatment buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 30 mg/mL lysozyme, and 500 u/mL mutanolysin). After incubation at 37ºC for 10 min, total RNA was extracted with the RNeasy® mini kit (Qiagen). RNA samples were treated with RNase-free DNase I (0.06 U/µL; Ambion) to minimize genomic DNA contamination and the total RNA concentration was determined by using a ND-1000 spectrophotometer (NanoDrop Technologies, Inc.).

For RNA stability (half-life) experiments, aliquots of cells in CDM were taken at desired timepoints and rifampicin was added (0.5 mg/mL; Sigma). Growth was continued, and samples were taken at 0, 2, 5, and 10 min post rifampicin addition and added to 2 volumes of RNAprotect™ bacteria reagent. RNA was then extracted as described above.

4.2.4 Reverse transcription and quantitative PCR

Reverse transcription reactions were carried out with total RNA (~100 ng) extracted from *E. faecalis*, a primer complementary to the transcript of interest (3 µM) and ThermoScript™ reverse transcriptase (0.75 U/µL; Invitrogen). The resulting cDNA was used as a template for quantitative real-time PCR. Formation of PCR product was measured by iQ™ SYBR® Green fluorescent dye (Bio-Rad) in duplicate. Data sets were collected with a Bio-Rad iQ real-time PCR system and analyzed using iCycler version 3.1 software (Bio-Rad). Each data point was compared to a standard curve to determine the copy number of the target gene in each sample. The copy number in each sample was normalized based on the amount of RNA that was present in the initial reverse transcription reaction. Fluorescence data were collected for melt curves from 56ºC to

95ºC after the last PCR cycle to show that only one product was produced in each reaction. Each experiment was repeated at least three times.

4.2.5 Incorporation of gene fusions into *E. faecalis*

The *E. faecalis metK* sequence containing residues -80 to +116 (relative to the predicted transcription start site) was fused in frame to the *E. coli gusA* coding sequence. The resulting constructs were incorporated into plasmid pCJK141, which was designed for markerless exchange in *E. faecalis* at an intergenic region, by restriction digestion and ligation. Ligated plasmids were introduced into *E. coli* EC1000 by transformation. Plasmids were purified from transformants and sequenced to confirm the presence of the correct gene fusion. Purified plasmids containing gene fusions were introduced by electroporation into *E. faecalis* CK111 as described (5) except that the concentration of lysozyme was 500 μg/mL. Transformants were used as donors to incorporate the gene fusion into *E. faecalis* OG1RF by conjugation, which was carried out as described (48). Transconjugate colonies arose within 24-36 h at a frequency of ~10^{-3} per donor. Transconjugates were purified at least twice on selective medium and then chromosomal DNA was isolated to confirm the integration of the gene fusion at the correct site by PCR. Isolates containing the correct gene fusion were then subjected to methionine starvation and β-glucuronidase assays.

4.2.6 β-Glucuronidase assay

E. faecalis OG1RF cells containing *metK-gusA* fusions were grown in CDM minimal medium containing Erm (100 μg/mL), methionine (230 μM) and yeast extract

(0.005%). Cells were harvested by centrifugation and resuspended in fresh CDM lacking yeast extract with either low (4.6 μM) or high (1.4 mM) methionine. Samples were taken at 1 h intervals for 6 h, the OD_{595} was measured, and the samples were stored at -70ºC. A β-glucuronidase assay was developed based on that of Bae et al. (4). Briefly, samples were resuspended in 500 μL of a treatment buffer (10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 10 mM EDTA and 30 mg/mL lysozyme) and incubated at 37ºC for 15 min. Cells were collected by centrifugation, resuspended in 1 mL of GUS buffer (12.5 mM NaH_2PO_4, 37.5 mM Na_2HPO_4, 10 mM β-mercaptoethanol, 1 mM EDTA, and 0.1% Triton X-100), and then incubated at 37ºC for 10 min. p-Nitrophenyl-β-D-glucuronide was added to each sample at a final concentration of 0.5 mg/mL, the samples were incubated at room temperature until a yellow color became visible, Na_2CO_3 (300 mM) was added to stop the reaction, and the reaction time was recorded. The absorbance of each sample was measured at 405 nm and β-glucuronidase activity was calculated and expressed in units similar to Miller Units (104). Each experiment was repeated at least twice.

4.2.7 SAM synthetase assay

An overnight culture of *E. faecalis* OG1 was used to inoculate fresh CDM medium containing all amino acids. The culture was incubated at 37ºC until mid-log phase and the cells were collected by centrifugation. The cells were then split in fresh CDM containing either low (1.8 μM) or high (230 μM) methionine and the cultures were incubated for 2.5 h at 37ºC. The OD_{595} was measured for each culture and an equal number of cells from each were collected by centrifugation. In order to make cell

extracts, cells were resuspended in a treatment buffer (100 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 100 mM KCl, 1 mM DTT, 20% glycerol, 10 mg/mL lysozyme, 500 U/mL mutanolysin, and protease inhibitor mix [Roche]), incubated at 37ºC for 10 min, and then 0.1 mm zirconia/silica beads (1.25 g; BioSpec) were added and cells were disrupted by vortexing for 5 min. Cell extracts were assayed for SAM synthetase activity by measuring incorporation of [methyl-^{3}H]methionine (1 Ci/mmol; MP Biomedicals) into [^{3}H]SAM using a protocol adapted from the work of Ochi and Freese (73). Samples were filtered through P81 phosphocellulose paper (Upstate Biotechnology) and washed with distilled water. Bound [^{3}H]SAM was quantitated in Packard Bioscience Ultima Gold scintillation fluid using a Packard Tri-Carb 2100TR liquid scintillation counter. The amount of [^{3}H]SAM generated in the reaction mixture was measured over time and used to determine the activity of SAM synthetase in the crude cell extract by calculating the amount of [^{3}H]SAM (in pmol) generated per minute.

4.3 Results

4.3.1 Manipulation of SAM pools in *E. faecalis*

It has been shown that intracellular SAM concentrations decrease in a *B. subtilis* methionine auxotroph when the cells are starved for methionine (87, 151). This is expected as SAM is synthesized from methionine and ATP. We therefore measured SAM pools in *E. faecalis* cells grown in the presence or absence of methionine, to identify conditions under which SAM pools vary. Growth of *E. faecalis* strain OG1 in CDM minimal medium was dependent on addition of methionine, consistent with the observation that *E. faecalis* strains require methionine for growth (58). Cells were grown

in CDM containing methionine (230 μM), and resuspended in CDM containing either high (230 μM) or low (1.8 μM) methionine, and samples were taken at intervals and extracted with formic acid. The extracts were added to an *in vitro* transcription assay using a *B. subtilis yitJ* template, which contains an S box riboswitch that exhibits SAM-dependent transcription termination (58). The concentration of SAM in each extract was determined by comparing the ratio of terminated and readthrough transcription products to a standard curve generated using known amounts of SAM. The intracellular concentration of SAM was then calculated based on the A_{595} of each culture sample.

SAM levels in *E. faecalis* were observed to increase during the first 40 min after dilution into fresh medium (Fig. 4.1A). This is likely to be due to the small amount of methionine in the fresh medium. The SAM concentration then decreased rapidly, with a 5-fold reduction at the 53 min time point followed by a slow decline over the rest of the 180 min time course (~8-fold total reduction of SAM concentration). This result is similar to what was observed in *B. subtilis,* where SAM pools decreased 6-fold at the 75 min time point and showed a 10-fold total reduction over the time course (87). Addition of methionine to the *E. faecalis* culture reversed the decrease in SAM pools; when methionine was added to 230 μM after the 60 min time point, SAM levels completely recovered to the maximal value within 10 min (Fig. 4.1B). These results show that SAM levels in *E. faecalis* can be easily manipulated, which suggests that regulatory responses due to SAM can be studied by growth in CDM medium containing varying concentrations of methionine.

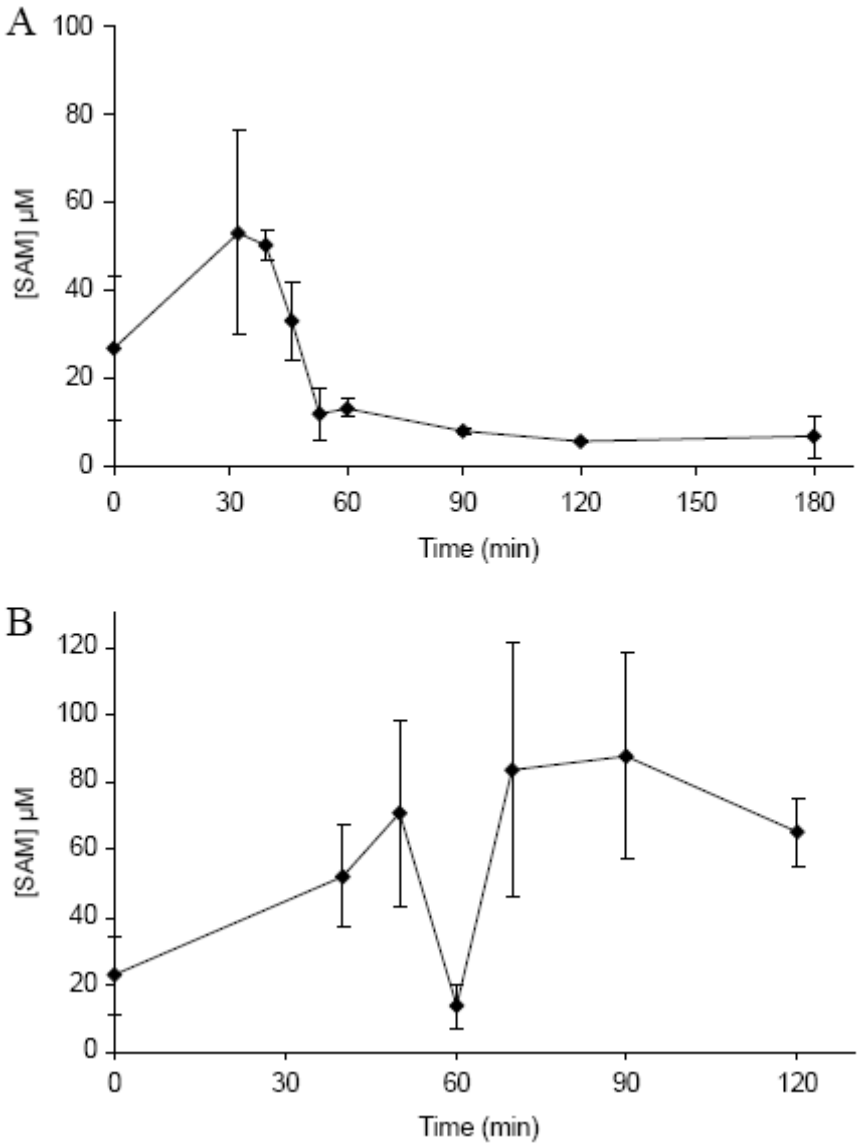

Figure 4.1. Measurement of SAM pools in *E. faecalis*. (**A**) *E. faecalis* cells were grown in minimal medium containing high methionine (230 µM) until mid-log phase. Cells were collected by centrifugation and resuspended in fresh media containing low methionine (1.8 µM). Samples were taken at timepoints as indicated and extracted with formic acid. The concentration of SAM in the extract was determined in an *in vitro* transcription reaction with a SAM-responsive S box RNA (*yitJ* from *B. subtilis*). (**B**) Cells were treated as in (**A**) except that the methionine concentration was adjusted back to 230 µM after the 60 min sample was taken.

4.3.2 *metK* transcript abundance in *E. faecalis*

The abundance of the *metK* transcript in *E. faecalis* grown in high or low methionine was measured by using a quantitative reverse transcription PCR assay to determine if transcript levels change in response to changing methionine/SAM levels. It was expected that transcript levels would remain the same since the S_{MK} box riboswitch is predicted to regulate *metK* expression at the level of translation initiation and not at the level of transcription. Cells were grown in CDM medium in the presence of high (230 μM) or low (1.8 μM) concentrations of methionine and total RNA was extracted at 1 h intervals. The RNA samples were treated with DNase to remove DNA contamination and the RNA was used as a template for reverse transcription. Separate reverse transcription reactions were carried out using primers that correspond to the coding region of *metK*, the coding region of a putative methionine synthase gene (which is predicted to be regulated by the T box mechanism in response to a reduction in aminoacylated $tRNA^{Met}$; 122), and 5S rRNA. After reverse transcription, each cDNA sample was used as a template in a quantitative real-time PCR reaction. The amount of the target template was determined and then normalized based on the starting amount of total RNA. The change in transcript levels was determined by dividing the value for cells grown in low methionine by the value for cells grown in high methionine.

The 5S rRNA showed no change in abundance in low versus high methionine conditions (Table 4.1). This result was expected as 5S rRNA was chosen as a negative control that should not respond to methionine limitation. In contrast, the putative methionine synthase gene showed a 26-fold increase in transcript abundance after 1 h in low methionine medium (Table 4.1), consistent with the prediction that this gene is

regulated at the level of transcription antitermination in response to methionine limitation. This result serves as a positive control and indicates that the cells were in fact starved for methionine. No change was observed in the abundance of the *metK* transcript even after 3 h of incubation in low methionine medium, conditions under which the SAM pools dropped 8-fold (Fig. 4.1A). This suggests that there is no significant regulation at the level of *metK* transcript abundance in response to depletion of methionine and SAM in *E. faecalis*.

Table 4.1. Transcript levels in *E. faecalis*

	mRNA induction ratio (-met/+met)[a]			
RNA[b]	**0 h**	**1 h**	**2 h**	**3 h**
metK	1.2 ± 0.28	1.5 ± 0.20	1.2 ± 0.60	1.0 ± 0.13
5S rRNA	1.5 ± 0.53	1.1 ± 0.30	1.2 ± 0.31	1.3 ± 0.42
Methionine synthase gene	1.2 ± 0.10	26 ± 5.7	13 ± 1.6	7.4 ± 3.0

[a]Total RNA was extracted at 1 h intervals from *E. faecalis* cells grown in CDM medium containing either low (1.8 μM) or high methionine (230 μM). The abundance of each transcript was determined by reverse transcription followed by quantitative PCR. mRNA induction ratio is the abundance of the transcript under low methionine conditions divided by the transcript abundance under high methionine conditions. [b]Specific primers for reverse transcription were designed for each RNA. The methionine synthase gene is predicted based on sequence similarity; it contains a T box riboswitch motif that is predicted to respond to a reduction in aminoacylated $tRNA^{Met}$ (122).

4.3.3 Stability of the *metK* transcript in *E. faecalis*

After measuring the effect of methionine limitation on transcript abundance, we wanted to determine if methionine limitation affects transcript stability. Addition of rifampicin followed by quantitative PCR was utilized to determine the half-life ($t_{1/2}$) of transcripts in *E. faecalis*. Rifampicin blocks transcription of new RNA in cells by binding and inactivating RNA polymerase. In the absence of new synthesis, the amount

of a given transcript remaining in the cell at any time after rifampicin addition is a function of the stability of the transcript. The $t_{1/2}$ can be determined by measuring the decrease in transcript abundance over time. 5S rRNA was included as a stable RNA control, and no significant decrease in 5S rRNA abundance was observed over the time course (data not shown).

The $t_{1/2}$ of the putative methionine synthase gene transcript, which is predicted to be under the regulatory control of a T box element, was 3.9 +/- 0.78 min in cells grown in the presence of high methionine (Table 4.2). The $t_{1/2}$ increased slightly to 5.4 +/- 0.32 min after growth in low methionine for 1 h. This increase was transient, however, and the $t_{1/2}$ returned to a value similar to the initial value after 2 h. This trend is consistent with the transcript abundance data which showed a 26-fold increase at 1 h (which was attributed to readthrough of the T box regulatory system in the presence of uncharged $tRNA^{Met}$) that tapered off at the 2 and 3 h time points (Table 4.1). Although transcript stability may play a small role in this trend, the T box antitermination mechanism is likely to be the most important factor in the initial 26-fold increase in transcript levels.

Table 4.2. Half-life ($t_{1/2}$) of RNA transcripts *in vivo*

		Growth conditions (time, [methionine])[a]		
		1 h, 230 μM	**1 h, 1.8 μM**	**2 h, 1.8 μM**
RNA[b]	**Riboswitch**	**$t_{1/2}$ (min)**	**$t_{1/2}$ (min)**	**$t_{1/2}$ (min)**
E. fae metK	S_{MK} box	2.8 +/- 0.33	3.5 +/- 0.62	3.7 +/- 0.35
E. fae metK, 3' coding region	S_{MK} box	3.2 +/- 0.15	4.4 +/- 0.36	3.9 +/- 0.61
E. fae putative met gene	T box	3.9 +/- 0.78	5.4 +/- 0.32	3.6 +/- 0.13
B. sub metK	S box	1.3 +/- 0.43	N/A	N/A
B. sub yitJ	S box	1.4 +/- 0.10	N/A	N/A

[a]Cells were grown in CDM medium or Spizizen medium (2, 66) until mid-log phase for *E. faecalis* or *B. subtilis*, respectively. The cells were collected by centrifugation and resuspended in medium containing either low (1.8 μM) or high (230 μM) methionine, as indicated. After incubation at 37ºC for 1 or 2 h, rifampicin was added to 0.5 mg/mL and samples of the culture were taken at 0, 2, 5, and 10 min. Total RNA was extracted and transcript abundance was measured by quantitative PCR. The $t_{1/2}$ was determined from the decrease in transcript abundance over time. [b]Primers for reverse transcription were designed for each transcript and were complementary to a region near the 5' end of the coding region unless otherwise indicated.

A primer complimentary to the 5' end of the *metK* coding region resulted in a transcript $t_{1/2}$ of 2.8 +/- 0.33 minutes under high methionine conditions (Table 4.2). A similar result was obtained when a primer corresponding to the 3' end of the *metK* coding region was used, indicating that the full-length transcript exhibits similar stability. Growth under low methionine conditions resulted in a slight increase in the stability of the transcript. This increase in stability is likely to be due to protection of the RNA from degradation by increased translation, resulting from increased availability of the SD region in the absence of SAM.

The stability of two S box transcripts (*metK* and *yitJ*) in *B. subtilis* was also measured in order to compare S_{MK} box RNA stability with the stability of transcripts

containing SAM-dependent riboswitches that are regulated at the level of premature transcription termination. Both of the S box gene transcripts were determined to have a $t_{1/2}$ of less than 1.5 min (Table 4.2). Although this data set is rather limited relative to the number of known S box sequences, these results suggest that S box gene transcripts have a more rapid rate of turnover than S_{MK} box transcripts, consistent with regulation at the level of transcription attenuation rather than translation initiation.

4.3.4 Expression of *metK-gusA* fusions and SAM synthetase in *E. faecalis*

Although it is apparent from *lacZ* fusion data (see Chapter 2) that the S_{MK} box sequence is a functional gene regulator in *B. subtilis* (a heterologous host), there was still a question of whether the S_{MK} box is a functional gene regulator in a native organism. In order to address this question, a fusion of the *E. faecalis metK* leader to the *E. coli gusA* reporter gene, which encodes β-glucuronidase, was generated and incorporated into the *E. faecalis* chromosome. The fusion consisted of the *metK* sequence encompassing residues -80 to +116 (relative to the predicted transcription start site) fused to the *gusA* coding sequence. *E. faecalis* cells containing the fusion were grown in low (4.6 μM) or high (1.4 mM) methionine conditions and samples were taken at 1 h intervals and assayed for β-glucuronidase activity. At each timepoint, there was less than a 2-fold difference in β-glucuronidase expression between cells grown in low or high methionine (Fig. 4.2A). This is in contrast to *metK-lacZ* fusions in *B. subtilis* where expression is repressed 5-fold under high methionine conditions (Table 2.1).

One possible explanation for the apparent lack of regulation in *E. faecalis* is that a second layer of regulation at the *metK* promoter masks the regulatory effect of the S_{MK}

box. In order to test this possibility, a fusion was made that contained residues 15-116 of the *E. faecalis metK* sequence positioned downstream of a *B. subtilis glyQS* promoter. The expression of this fusion was also not repressed under high methionine conditions (Fig. 4.2B). This result suggests that the lack of S_{MK} box-dependent regulation of gene fusions in *E. faecalis* is not due to an effect of the *metK* promoter.

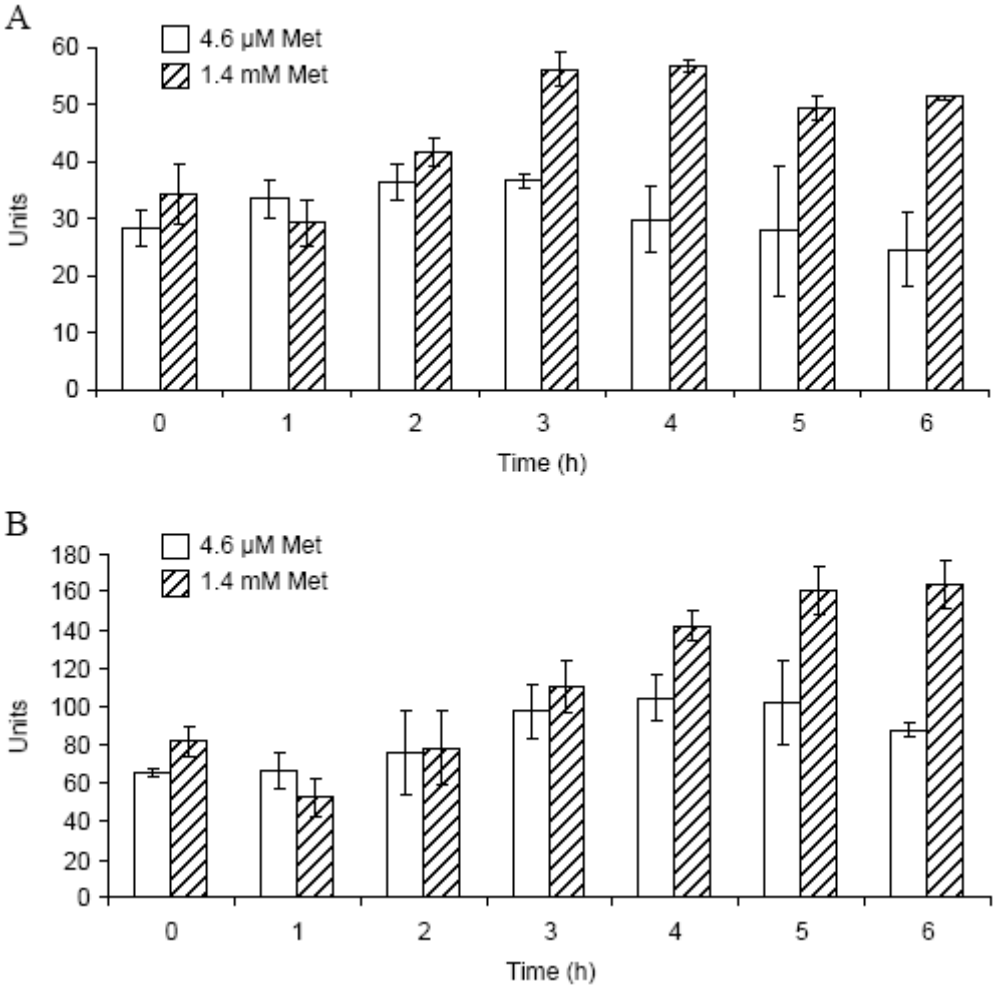

Figure 4.2. Expression of *gusA* fusions in *E. faecalis*. (**A**) Translational fusion of *E. faecalis metK* (containing residues -80 to +116 relative to the predicted transcription start site). (**B**) Translational fusion of *E. faecalis metK* (containing residues 15-116 relative to the predicted transcription start site) under the control of the *B. subtilis glyQS* promoter. Cells were grown in CDM minimal medium until early log phase (66). Cells were then resuspended in fresh CDM medium containing methionine at either 4.6 µM or 1.4 mM. Samples were taken at 1 h intervals and assayed for β-glucuronidase activity. Units of β-glucuronidase activity were calculated in a manner similar to that of Miller Units (104).

Due to potential problems associated with gene fusions, a direct assay for MetK (SAM synthetase) activity in *E. faecalis* cellular extracts was utilized. This assay measures the rate of conversion of radiolabeled methionine into SAM, as described by Ochi and Freese (73). *E. faecalis* OG1 cells were grown to mid-log phase and then resuspended under high (230 μM) or low (1.8 μM) methionine conditions. The cells were incubated at 37ºC for 2.5 h and then were treated with lysozyme and mutanolysin, and vortexed in the presence of zirconia/silica beads. The resulting extracts were added to ATP and [methyl-^{3}H]methionine and the conversion of the methionine to [^{3}H]SAM was monitored over time.

Based on the conclusions drawn in Chapters 2 and 3, it was expected that cells grown under high methionine conditions would have reduced SAM synthetase activity due to S_{MK} box-dependent repression of *metK* expression. Unfortunately, the results of the *E. faecalis* SAM synthetase assays were inconclusive at both the quantitative and qualitative levels. The amount of SAM synthetase activity in each trial was either undetectable or generated SAM at a rate of no more than 0.36 fmol min^{-1} OD_{595}^{-1} (Table 4.3). This is in contrast with the previously reported amount of activity in *B. subtilis* extracts where SAM was generated at a rate of 2.4 to 3.5 fmol min^{-1} OD_{595}^{-1} (60). This suggests that this assay is not suitable for detecting *E. faecalis* SAM synthetase activity, possibly because of difficulties in efficient cell lysis without denaturation of cellular proteins. Thus, the data generated by this assay are not valid.

Table 4.3. SAM synthetase activity of *E. faecalis* cell extracts[a]		
	SAM synthetase activity (fmol min^{-1} OD_{595}^{-1})[a]	
	Low methionine[b]	**High methionine[b]**
Trial 1	0.22	0.0
Trial 2	0.0	0.084
Trial 3	0.36	0.10
Average	0.19 ± 0.18	0.061 ± 0.054

[a]SAM synthetase activity was measured by adding cellular extracts to a solution containing [methyl-^{3}H]methionine and ATP and monitoring the conversion of [methyl-^{3}H]methionine to [methyl-^{3}H]SAM over time. [b]Cells were grown in CDM medium containing 230 μM methionine until mid-log phase. The cells were then collected by centrifugation, resuspended in fresh medium containing either low (1.8 μM) or high (230 μM) methionine, and incubated for 2.5 h at 37ºC before cell extracts were made.

4.4 Discussion

The S_{MK} box is a riboswitch that is predicted to regulate *metK* gene expression at the level of translation initiation through SAM-dependent sequestration of the Shine-Dalgarno sequence. In Chapter 2, *in vivo* studies were presented in which the effect of the *E. faecalis* S_{MK} box on gene expression was determined in the heterologous host *B. subtilis*. *B. subtilis* was chosen because of the ease of genetic manipulation of that organism, and because of the known effect on intracellular SAM pools of growth of a *B. subtilis* methionine auxotroph in the presence or absence of methionine (60, 87, 151). These studies demonstrated that the S_{MK} box regulates expression of reporter gene fusions in *B. subtilis* and that regulation is at the level of translation. Investigation of S_{MK} box-mediated regulation of *metK* in the native host, *E. faecalis*, required establishment of growth conditions that allowed modulation of intracellular SAM pools. In this chapter, it

was determined that *E. faecalis* SAM pools decrease in response to methionine limitation. It was also determined that the abundance and stability of the *metK* transcript do not change significantly when SAM pools are depleted. These results provide additional evidence that regulation of *metK* expression in response to SAM is not at the level of transcription or transcript stability. Unfortunately, we were unable to observe S_{MK} box-dependent regulation of a reporter gene fusion in *E. faecalis*. Thus, the regulatory role of the S_{MK} box in a native organism is unclear.

A major question in the riboswitch field is whether metabolite binding riboswitches can act as reversible switches, rather than as sensors that monitor their cognate signal and make a single gene expression decision. It has been shown that the SAM-responsive S box riboswitch, which controls genes involved in biosynthesis of methionine and SAM in many Firmicutes, including *B. subtilis* (41, 58, 87, 100), is likely to be irreversible. The S box RNA-SAM complex is highly stable, with a $t_{1/2}$ in excess of 4 min in most genes in this family (87). For example, the $t_{1/2}$ of the *B. subtilis yitJ*-SAM complex is 7.2 min; in comparison, we determined that the $t_{1/2}$ of the *yitJ* transcript *in vivo* is 1.4 min (Table 4.2). This provides additional evidence that binding of SAM is an irreversible event and that only one regulatory decision can be made before the transcript is degraded. Most S box riboswitches, including those in *B. subtilis*, regulate gene expression at the level of premature termination of expression. This results in a short window of opportunity for SAM binding (and potentially for dissociation and rebinding) during transcription of the ~200 nt leader RNA. This observation combined with our current transcript stability data suggests that the S box is likely to be an irreversible

riboswitch, and that once SAM binding occurs, the messenger RNA is committed to premature transcription termination.

In this chapter, we show that the *E. faecalis metK* transcript has a $t_{1/2}$ of ~3 min in cells grown under high methionine conditions. The stability of the transcript increased slightly to ~3.5 minutes under conditions of low methionine, which is likely to be due to protection conferred by the ribosome as translation increases. In all conditions tested, the measured $t_{1/2}$ of the mRNA transcript was much longer than the $t_{1/2}$ of the SAM-S_{MK} box complex (7.8 sec) determined *in vitro* (A. Smith, unpublished results). This indicates that the S_{MK} box may be a reversible switch, as SAM dissociation can occur before the mRNA is degraded, leaving the *metK* leader transcript accessible for either translation initiation or interaction with another SAM molecule. The relative rate constants for each of these processes, and the concentrations of reaction components available at a given time, are likely to determine which of these events will occur. A reversible mechanism of regulation at the level of translation initiation would ensure that cells are poised to utilize existing mRNA transcripts to rapidly respond to transient fluctuations in SAM pools. It has been determined that conformational switching of the *metK* leader between SAM-bound and SAM-free forms can be observed in response to fluctuating levels of SAM *in vitro* (A. Smith, unpublished results). Additionally, expression of a fusion of the *metK* leader to *lacZ* in *B. subtilis* has been shown to decrease within minutes after addition of methionine or SAM to methionine starved cells (A. Smith, unpublished results). Together these results suggest that a rapid SAM-dependent change in gene regulation by the S_{MK} box is possible.

The fact that the S_{MK} box riboswitch did not regulate expression of a *gusA* reporter gene in *E. faecalis* is surprising since it has been shown to regulate a *lacZ* reporter gene in a non-native host (see Chapter 2). A possible explanation for the lack of regulation in *E. faecalis* is that the *gusA* portion of the fusion transcript might interfere with the *metK* leader sequence and disrupt S_{MK} box function. This potential issue could be investigated by introducing a *metK-gusA* fusion into *B. subtilis* (to determine if regulation is similar to that of the *lacZ* fusions) or by introducing a *metK-lacZ* fusion into *E. faecalis* (to see if there is regulation when a different reporter gene is used). It is also possible that the *gusA* reporter does not accurately reflect *metK* expression levels under these growth conditions, for unknown reasons. Unfortunately, direct measurement of MetK enzymatic activity was unsuccessful, probably due to technical difficulties with cell lysis. Additionally, it is possible that the effect of the S_{MK} box and its ability to bind SAM could be hindered or enhanced by other regulatory systems that act upon the *metK* transcript *in vivo*.

In addition to understanding the utility of individual riboswitch classes, it is also important to explore the role of these regulatory RNAs within a contextual framework. The S box system from *B. subtilis* provides a good example of a riboswitch RNA that resides upstream of multiple transcriptional units (~11 in *B. subtilis*), and demonstrates a differential response to SAM in terms of the level of regulation for each of the associated genes (87). This suggests that fine tuning of the cellular response to intracellular levels of SAM varies depending on the gene that is being regulated. Differences in the regulatory response may be due to subtle differences in riboswitch sequence and structure,

or to overlapping regulatory mechanisms that operate in concert to modulate the overall level of expression.

While most genes regulated by the S box undergo tight repression in response to SAM, there are exceptions. The *metK* gene resides downstream of an S box element in *B. subtilis*, but it does not undergo the same regulatory response as other S box regulated genes (e.g., *yitJ*). Expression levels during repressive conditions are higher than for most S box transcriptional units (87). This is not surprising since the gene product, SAM synthetase, is required by the cell at all times. What is surprising is that although transcript levels increase transiently in response to methionine starvation, no persistent induction of gene expression has been observed in response to SAM depletion in a variety of *in vitro* and *in vivo* analyses (87). This suggests that *B. subtilis metK* is subject to regulatory mechanisms that act in concert with the S box element.

Despite the presence of an S_{MK} riboswitch element in the 5'-UTR, differential expression of *metK* has not been observed in *E. faecalis*. Thus, *E. faecalis* and *B. subtilis* both have a SAM-responsive riboswitch in the leader region of *metK*, but the role of that riboswitch *in vivo* is unclear. One possibility is that the cell monitors not only SAM levels but also methionine levels to determine the amount *metK* expression (e.g., inducing gene expression when methionine is high and repressing gene expression when methionine is low). It may make sense for the cell to do so because both SAM and methionine are vital for cellular survival. Our *in vivo* experiments rely on indirectly manipulating SAM levels by adjusting the levels of methionine. Thus, if the cells sense both SAM and methionine to determine *metK* expression, then the effects of the changing SAM levels on *metK* expression may be masked by the concurrent change in methionine

levels. It appears that the importance of the essential *metK* gene has caused *E. faecalis* and *B. subtilis* to develop a more complicated mechanism to maintain a proper level of expression. Investigation of the *metK* regulation scheme, and determining the roles of the S_{MK} and S box riboswitches within that scheme, will be a challenge that could yield fascinating new information about bacterial gene regulation.

CHAPTER 5

SUMMARY AND FUTURE DIRECTIONS

5.1 Summary of the S_{MK} box riboswitch

Riboswitches are conserved RNA sequences that undergo a structural change when they sense a signal (or effector molecule) and act in *cis* to control gene expression. Recent studies have shown that gene regulation by riboswitches is a commonly used mechanism in many bacteria and that riboswitches are present in all three domains of life. Bacterial riboswitches are typically in the 5'-UTR of the gene or operon that they regulate and generally consist of an aptamer (signal binding) domain and an expression platform, which is the structure responsible for the regulatory effect. Expression platforms are typically either an intrinsic transcriptional terminator, so that regulation is at the level of transcription attenuation, or a structure that sequesters the SD sequence which results in regulation at the level of translation initiation. The structure and/or formation of the expression platform changes as a result of signal recognition by the aptamer domain. The main goal of this work was to characterize a newly discovered riboswitch motif in the Lactobacillales called the S_{MK} box. Specifically, we wanted to investigate the mechanism of action of this riboswitch, identify the effector molecule it

responds to, and identify crucial structural requirements for effector molecule binding and riboswitch function.

It was known that in many bacterial species the *metK* gene (which encodes SAM synthetase) contains a SAM-responsive S box riboswitch motif in the 5'-UTR. Most species of the Lactobacillales order do not contain an S box motif in *metK*, but the 5'-UTR contains regions of conservation. It also contained regions of covariation, indicating the potential for RNA secondary structure, that included the possibility for pairing between part of the SD sequence with a conserved ASD sequence. We named this element the S_{MK} box and predicted that it may be a SAM-responsive riboswitch element that regulates at the level of translation initiation.

We demonstrated that S_{MK} box RNA binds SAM *in vitro* and discriminates against the closely related analog SAH. SAM binding causes structural changes in the RNA and these changes result in increased structure in the SD region. Mutational analysis showed that pairing between the ASD and SD is required for binding of SAM. It has been shown that the addition of SAM is sufficient to inhibit binding of *E. coli* 30S ribosomal subunits to the RNA *in vitro*. Additionally, primer extension inhibition assays confirmed that 30S subunits bind to the SD region of the RNA and also indicated that ASD-SD pairing and inhibition of 30S subunit binding is promoted by SAM. X-ray crystallography data for the SAM-bound form of the RNA provided additional evidence for ASD-SD pairing, showed that SAM is intercalated in a three-way helical junction in the RNA, and indicated that SAM is recognized by interactions between five residues in the RNA, including one residue in the SD sequence, that are 100% conserved. It has also been demonstrated that expression of a translational fusion of the *metK* 5'-UTR sequence

to a *lacZ* reporter gene is repressed in *B. subtilis* when cells are grown under conditions in which SAM pools are high. All of these results suggest that the S_{MK} box is a riboswitch that regulates gene expression through inhibition of translation by SAM-promoted sequestration of the SD region.

All of the initial work with the S_{MK} box riboswitch was performed either *in vitro*, with *E. coli* components, or in *B. subtilis*, and these organisms do not encode the S_{MK} box. We decided to investigate the properties of the S_{MK} box and the *metK* transcript in *E. faecalis* in order to elucidate the role of the S_{MK} box in a native organism. Our results show that *metK* transcript levels did not increase when cells are grown under conditions under which SAM pools are low. These results indicate that there is no regulation of *metK* at the level of transcript abundance in response to SAM. This leads to the conclusion that, as expected, the S_{MK} box riboswitch in *metK* does not regulate at the level of transcription. Additionally, the stability of the *metK* transcript did not significantly change under conditions under which SAM pools are low. This indicates that there is no SAM-dependent regulation at the level of transcript stability for *metK*. Also, the stability ($t_{1/2}$) of the transcript *in vivo* was measured to be much higher than the $t_{1/2}$ of the SAM-S_{MK} box complex *in vitro* (A. Smith, unpublished results) and SAM binding has been shown to be reversible *in vitro* (A. Smith, unpublished results). Although comparing *in vitro* results with *in vivo* results can be problematic, together these results suggest that SAM binding to the riboswitch may be reversible *in vivo*.

After we concluded that the S_{MK} box riboswitch does not regulate gene expression at the level of transcription or transcript stability in *E. faecalis*, experiments were performed to determine if regulation is at the level of translation as expected.

Unfortunately, translational *metK-gusA* fusions showed constant β-glucuronidase expression under all conditions tested. Additionally, attempts to detect differences in MetK activity in cellular extracts were inconclusive. Thus, the role of the S_{MK} box riboswitch in *E. faecalis* remains unclear.

5.2 Future directions

There is still much to be learned about the S_{MK} box riboswitch. Although the structure of the SAM-bound form of the RNA has been established, little is known about the SAM-free structure of the RNA. For example, it is not known whether or not the SAM binding pocket is preformed in the RNA or if formation of the pocket is induced by the presence of SAM. Nuclear Magnetic Resonance (NMR) experiments are being performed in order to help determine the SAM-free structure and address this question (A. Smith, R.Wilson). Additionally, the stages of folding and structure formation during S_{MK} box RNA transcription are not known. The dynamics of RNA folding could be investigated using single molecule techniques. Specifically, a technique called force-measuring optical tweezers (FMOT) allows the folding states of an RNA to be determined as it emerges from RNAP *in situ.* This technique has been used to determine the folding intermediates of an adenine riboswitch (123). Determination of the folding states of the S_{MK} box riboswitch in the presence and absence of SAM at different stages of transcription may help to elucidate the mechanism (or stages) of SAM binding.

Another question is that of the persistence of the *metK* transcript in the cell. The RNA stability data indicate that the *metK* transcript is much more stable *in vivo* than the S_{MK} box-SAM complex. This suggests that SAM binding may be reversible and that

each transcript may persist in the cell for a sufficient amount of time to make multiple SAM-dependent regulatory decisions. The stability data, however, represent an average of the entire population of *metK* transcripts. It is possible that there is variation in the stability of each transcript and that the variability is dependent upon whether it is bound to SAM, free of SAM, or being actively translated by ribosomes. In order to test this, the stability of a mutant *metK* sequence with a pyrimidine substitution in the SD sequence could be measured. A pyrimidine substitution in the SD would inhibit binding of ribosomal subunits, but would also probably disrupt pairing in the RNA and SAM binding. It would be beneficial if the mutated sequence included a compensatory mutation to restore pairing so that SAM binding is maintained and the overall RNA structure is likely to be similar to wild-type. A good candidate sequence to test would be the C24G/G91C mutant that maintains SAM binding (Fig. 2.5), but showed no detectable expression in *B. subtilis* when fused to *lacZ* (unpublished results), presumably because the SD sequence was disrupted (the G91C substitution creates a GGGCG SD sequence) and ribosomes could not initiate translation. If the stability of this mutant transcript is significantly reduced under both high and low SAM conditions, then that would indicate that the persistence of the *metK* transcript in the cell is largely dependent upon active translation. This would suggest that the S_{MK} box riboswitch mechanism is not in fact reversible as transcripts lacking bound ribosomes would experience rapid turnover. In contrast, if the stability of the transcript remained the same, then the persistence of the *metK* transcript in the cell is likely to be independent of translation. This result would support our hypothesis that the S_{MK} box riboswitch can make multiple SAM-dependent regulatory decisions prior to *metK* transcript degradation.

Perhaps the most important question to investigate in the future is that of the role of the S_{MK} box *in vivo* in *E. faecalis*. Differential expression of MetK and *metK-gusA* fusions has yet to be demonstrated in *E. faecalis*. One problem is that the MetK activity assays of cell extracts were inconclusive. This assay suffers from the fact cells must be lysed under non-denaturing conditions in order to preserve protein function and *E. faecalis* cells are inherently difficult to lyse under these conditions. Development of a method to improve lysis may improve the results of this assay. Additionally, simply increasing the amount of cells used to make extracts may yield valid results. Furthermore, refinement of the reaction conditions may produce more significant results. It is possible that *E. faecalis* MetK is not very active and/or stable under the reaction conditions, which were based on those used for *B. subtilis* MetK. Although it is not clear what conditions may improve yields for *E. faecalis*, it has been shown that addition of *p*-toluenesulfonate, β-mercaptoethanol, acetonitrile, or urea have increased yields of SAM in large-scale reactions utilizing *E. coli* MetK (171).

Another potential problem that might have caused the lack of observed S_{MK} box riboswitch-mediated regulation is that the level of SAM in the cells was indirectly manipulated by controlling the availability of methionine in the medium. It is possible that *E. faecalis* regulates *metK* expression not only in response to SAM but also in a different manner in response to methionine. Thus, the effects of the change in SAM levels may be masked by the corresponding change in methionine levels. It would be beneficial to develop a method to reduce the SAM pool in *E. faecalis* without changing methionine availability in order to test the direct effects of SAM on regulation. One possible way that this may be accomplished is by incorporation of a SAM hydrolase gene

into cells to reduce SAM pools. Bacteriophage T3 encodes a SAM hydrolase (SAMase) enzyme that has been shown to be active in *E. coli* and derepress the *met* regulon (172, 173). Derepression of the regulon is attributed to the hydrolysis of SAM, which acts as a corepressor with the MetJ repressor protein. The activity of the SAMase enzyme in *E. faecalis* would need to be assessed in order to determine the validity of this approach for analysis of S_{MK} box-dependent regulation.

Although a significant amount of data have already been generated regarding the S_{MK} box RNA; the remaining open questions combined with the mystery of its regulatory role in native organisms suggests that there is much more information about this interesting riboswitch that is awaiting discovery.

LIST OF REFERENCES

1. Altman, S., Wesolowski, D., Guerrier-Takada, C., and Li, Y. 2005. RNase P cleaves transient structures in some riboswitches. *Proc. Natl. Acad. Sci. USA* **102**: 11284-9.

2. Anagnostopoulos, C. and Spizizen, J. 1961. Requirements for transformation in *Bacillus subtilis.* *J. Bacteriol.* **81**: 741-6.

3. Anupam, R., Nayek, A., Green, N.J., Grundy, F.J., Henkin, T.M., Means, J.A., Bergmeier, S.C., and Hines, J.V. 2008. 4,5-Disubstituted oxazolidinones: High affinity molecular effectors of RNA function. *Bioorg. Med. Chem. Lett.* **18**: 3541-4.

4. Bae, T., Clerc-Bardin, S., and Dunny, G.M. 2000. Analysis of expression of *prgX*, a key negative regulator of the transfer of the *Enterococcus faecalis* pheromone-inducible plasmid pCF10. *J. Mol. Biol.* **297**: 861-75.

5. Bae, T., Kozlowicz, B., and Dunny, G.M. 2002. Two targets in pCF10 DNA for *PrgX* binding: their role in production of Qa and *prgX* mRNA and in regulation of pheromone-inducible conjugation. *J. Mol. Biol.* **315**: 995-1007.

6. Baker, A.M. and Draper, D.E. 1995. Messenger RNA recognition by fragments of ribosomal protein S4. *J. Biol. Chem.* **270**: 22939-45.

7. Barrick, J.E. and Breaker, R.R. 2007. The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol.* **8**: R239.

8. Batey, R.T., Gilbert, S.D., and Montange, R.K. 2004. Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. *Nature* **432**: 411-5.

9. Bernstein, J.A., Khodursky, A.B., Lin, P.H., Lin-Chao, S., and Cohen, S.N. 2002. Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. *Proc. Natl. Acad. Sci. USA* **99**: 9697-702.

10. Blount, K.F. and Breaker, R.R. 2006. Riboswitches as antibacterial drug targets. *Nat. Biotechnol.* **24**: 1558-64.

11. Boiziau, C., Larrouy, B., Sproat, B.S., and Toulme, J.J. 1995. Antisense 2'-O-alkyl oligoribonucleotides are efficient inhibitors of reverse transcription. *Nucleic Acids Res.* **23**: 64-71.

12. Brantl, S. 2006. Bacterial gene regulation: metal ion sensing by proteins or RNA. *Trends Biotechnol.* **24**: 383-6.

13. Burgin, A.B. and Pace, N.R. 1990. Mapping the active site of ribonuclease P RNA using a substrate containing a photoaffinity agent. *EMBO J.* **9**: 4111-8.

14. Burke, D.H. and Gold, L. 1997. RNA aptamers to the adenosine moiety of *S*-adenosyl methionine: structural inferences from variations on a theme and the reproducibility of SELEX. *Nucleic Acids Res.* **25**: 2020-4.

15. Cheah, M.T., Wachter, A., Sudarsan, N., and Breaker, R.R. 2007. Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. *Nature* **447**: 497-500.

16. Chiba, Y., Sakurai, R., Yoshino, M., Ominato, K., Ishikawa, M., Onouchi, H., and Naito, S. 2003. *S*-adenosyl-L-methionine is an effector in the posttranscriptional autoregulation of the cystathionine gamma-synthase gene in Arabidopsis. *Proc. Natl. Acad. Sci. USA* **100**: 10225-30.

17. Condon, C. 2003. RNA processing and degradation in *Bacillus subtilis*. *Microbiol. Mol. Biol. Rev.* **67**: 157-74.

18. Condon, C. 2007. Maturation and degradation of RNA in bacteria. *Curr. Opin. Microbiol.* **10**: 271-8.

19. Corbino, K.A., Barrick, J.E., Lim, J., Welz, R., Tucker, B.J., Puskarz, I., Mandal, M., Rudnick, N.D., and Breaker, R.R. 2005. Evidence for a second class of *S*-adenosylmethionine riboswitches and other regulatory RNA motifs in alpha-proteobacteria. *Genome Biol.* **6**: R70.

20. Deikus, G. and Bechhofer, D.H. 2007. Initiation of decay of *Bacillus subtilis* trp leader RNA. *J. Biol. Chem.* **282**: 20238-44.

21. Dubey, A.K., Baker, C.S., Suzuki, K., Jones, A.D., Pandit, P., Romeo, T., and Babitzke, P. 2003. CsrA regulates translation of the *Escherichia coli* carbon starvation gene, *cstA*, by blocking ribosome access to the *cstA* transcript. *J. Bacteriol.* **185**: 4450-60.

22. Dunny, G.M., Lee, L.N., and LeBlanc, D.J. 1991. Improved electroporation and cloning vector system for gram-positive bacteria. *Appl. Environ. Microbiol.* **57**: 1194-201.

23. Feng, L. and Niu, D.K. 2007. Relationship between mRNA stability and length: an old question with a new twist. *Biochem. Genet.* **45**: 131-7.

24. Fernandez-Luna, M.T. and Miranda-Rios, J. 2008. Riboswitch folding: one at a time and step by step. *RNA Biol.* **5**: 20-3.

25. Fuchs, R.T., Grundy, F.J., and Henkin, T.M. 2006. The S_{MK} box is a new SAM-binding RNA for translational regulation of SAM synthetase. *Nat. Struct. Mol. Biol.* **13**: 226-33.

26. Fuchs, R.T., Grundy, F.J., and Henkin, T.M. 2007. *S*-adenosylmethionine directly inhibits binding of 30S ribosomal subunits to the S_{MK} box translational riboswitch RNA. *Proc. Natl. Acad. Sci. USA* **104**: 4876-80.

27. Giard, J.C., Laplace, J.M., Rince, A., Pichereau, V., Benachour, A., Leboeuf, C., Flahaut, S., Auffray, Y., and Hartke, A. 2001. The stress proteome of *Enterococcus faecalis*. *Electrophoresis* **22**: 2947-54.

28. Gilbert, S.D., Rambo, R.P., Van Tyne, D., and Batey, R.T. 2008. Structure of the SAM-II riboswitch bound to *S*-adenosylmethionine. *Nat. Struct. Mol. Biol.* **15**: 177-82.

29. Gollnick, P. and Babitzke, P. 2002. Transcription attenuation. *Biochim. Biophys. Acta.* **1577**: 240-50.

30. Grundy, F.J. and Henkin, T.M. 1998. The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in gram-positive bacteria. *Mol. Microbiol.* **30**: 737-49.

31. Grundy, F.J., Winkler, W.C., and Henkin, T.M. 2002. tRNA-mediated transcription antitermination in vitro: codon-anticodon pairing independent of the ribosome. *Proc. Natl. Acad. Sci. USA* **99**: 11121-6.

32. Grundy, F.J., Lehman, S.C., and Henkin, T.M. 2003. The L box regulon: lysine sensing by leader RNAs of bacterial lysine biosynthesis genes. *Proc. Natl. Acad. Sci. USA* **100**: 12057-62.

33. Grundy, F.J., Yousef, M.R., and Henkin, T.M. 2005. Monitoring uncharged tRNA during transcription of the *Bacillus subtilis glyQS* gene. *J. Mol. Biol.* **346**: 73-81.

34. Grundy, F.J. and Henkin, T.M. 2006. From ribosome to riboswitch: control of gene expression in bacteria by RNA structural rearrangements. *Crit. Rev. Biochem. Mol. Biol.* **41**: 329-38.

35. Hambraeus, G., von Wachenfeldt, C., and Hederstedt, L. 2003. Genome-wide survey of mRNA half-lives in *Bacillus subtilis* identifies extremely stable mRNAs. *Mol. Genet. Genomics* **269**: 706-14.

36. Hampel, K.J. and Tinsley, M.M. 2006. Evidence for preorganization of the *glmS* ribozyme ligand binding pocket. *Biochemistry* **45**: 7861-71.

37. Hanson, S., Bauer, G., Fink, B., and Suess, B. 2005. Molecular analysis of a synthetic tetracycline-binding riboswitch. *RNA* **11**: 503-11.

38. Harvey, I., Garneau, P., and Pelletier, J. 2002. Inhibition of translation by RNA-small molecule interactions. *RNA* **8**: 452-63.

39. Helm, M. 2006. Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucleic Acids Res.* **34**: 721-33.

40. Henkin, T.M. and Yanofsky, C. 2002. Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions. *Bioessays* **24**: 700-7.

41. Henkin, T.M. 2008. Riboswitch RNAs: using RNA to sense cellular metabolism. *Genes Dev.* **22**: 3383-90.

42. Heppell, B. and Lafontaine, D.A. 2008. Folding of the SAM aptamer is determined by the formation of a K-turn-dependent pseudoknot. *Biochemistry* **47**: 1490-9.

43. Jin, H., Zhao, Q., Gonzalez de Valdivia, E.I., Ardell, D.H., Stenstrom, M., and Isaksson, L.A. 2006. Influences on gene expression *in vivo* by a Shine-Dalgarno sequence. *Mol. Microbiol.* **60**: 480-92.

44. Karlson, A.G. and Ulrich, J.A. 1969. Stability of rifampin in dimethylsulfoxide. *Appl. Microbiol.* **18**: 692-3.

45. Kim, J.N., Roth, A., and Breaker, R.R. 2007. Guanine riboswitch variants from *Mesoplasma florum* selectively recognize 2'-deoxyguanosine. *Proc. Natl. Acad. Sci. USA* **104**: 16092-7.

46. Klein, D.J. and Ferre-D'Amare, A.R. 2006. Structural basis of *glmS* ribozyme activation by glucosamine-6-phosphate. *Science* **313**: 1752-6.

47. Kristich, C.J., Manias, D.A., and Dunny, G.M. 2005. Development of a method for markerless genetic exchange in *Enterococcus faecalis* and its use in construction of a *srtA* mutant. *Appl. Environ. Microbiol.* **71**: 5837-49.

48. Kristich, C.J., Chandler, J.R., and Dunny, G.M. 2007. Development of a host-genotype-independent counterselectable marker and a high-frequency conjugative delivery system and their use in genetic analysis of *Enterococcus faecalis*. *Plasmid* **57**: 131-44.

49. Leclerc, G.J., Leclerc, G.M., and Barredo, J.C. 2002. Real-time RT-PCR analysis of mRNA decay: half-life of Beta-actin mRNA in human leukemia CCRF-CEM and Nalm-6 cell lines. *Cancer Cell Int.* **2**: 1.

50. Leenhouts, K., Buist, G., Bolhuis, A., ten Berge, A., Kiel, J., Mierau, I., Dabrowska, M., Venema, G., and Kok, J. 1996. A general system for generating unlabelled gene replacements in bacterial chromosomes. *Mol. Gen. Genet.* **253**: 217-24.

51. Lemay, J.F., Penedo, J.C., Tremblay, R., Lilley, D.M., and Lafontaine, D.A. 2006. Folding of the adenine riboswitch. *Chem. Biol.* **13**: 857-68.

52. Lieber, C.S. and Packer, L. 2002. *S*-Adenosylmethionine: molecular, biological, and clinical aspects--an introduction. *Am. J. Clin. Nutr.* **76**: 1148S-50S.

53. Lockard, R.E. and Kumar, A. 1981. Mapping tRNA structure in solution using double-strand-specific ribonuclease V1 from cobra venom. *Nucleic Acids Res.* **9**: 5125-40.

54. Lowman, H.B. and Draper, D.E. 1986. On the recognition of helical RNA by cobra venom V1 nuclease. *J. Biol. Chem.* **261**: 5396-403.

55. Lu, C., Smith, A.M., Fuchs, R.T., Ding, F., Rajashankar, K., Henkin, T.M., and Ke, A. 2008. Crystal structures of the SAM-III/S_{MK} riboswitch reveal the SAM-dependent translation inhibition mechanism. *Nat. Struct. Mol. Biol.* **15**: 1076-83.

56. Mandal, M., Lee, M., Barrick, J.E., Weinberg, Z., Emilsson, G.M., Ruzzo, W.L., and Breaker, R.R. 2004. A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science* **306**: 275-9.

57. McCarthy, T.J., Plog, M.A., Floy, S.A., Jansen, J.A., Soukup, J.K., and Soukup, G.A. 2005. Ligand requirements for *glmS* ribozyme self-cleavage. *Chem. Biol.* **12**: 1221-6.

58. McDaniel, B.A., Grundy, F.J., Artsimovitch, I., and Henkin, T.M. 2003. Transcription termination control of the S box system: direct measurement of *S*-adenosylmethionine by the leader RNA. *Proc. Natl. Acad. Sci. USA* **100**: 3083-8.

59. McDaniel, B.A., Grundy, F.J., and Henkin, T.M. 2005. A tertiary structural element in S box leader RNAs is required for *S*-adenosylmethionine-directed transcription termination. *Mol. Microbiol.* **57**: 1008-21.

60. McDaniel, B.A., Grundy, F.J., Kurlekar, V.P., Tomsic, J., and Henkin, T.M. 2006. Identification of a mutation in the *Bacillus subtilis S*-adenosylmethionine synthetase gene that results in derepression of S-box gene expression. *J. Bacteriol.* **188**: 3674-81.

61. Miranda-Rios, J., Navarro, M., and Soberon, M. 2001. A conserved RNA structure (thi box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. *Proc. Natl. Acad. Sci. USA* **98**: 9736-41.

62. Montange, R.K. and Batey, R.T. 2006. Structure of the *S*-adenosylmethionine riboswitch regulatory mRNA element. *Nature* **441**: 1172-5.

63. Morana, A., Di Lernia, I., Carteni, M., De Rosa, R., and De Rosa, M. 2000. Synthesis and characterisation of a new class of stable *S*-adenosyl-L-methionine salts. *Int. J. Pharm.* **194**: 61-8.

64. Muller, M., Weigand, J.E., Weichenrieder, O., and Suess, B. 2006. Thermodynamic characterization of an engineered tetracycline-binding riboswitch. *Nucleic Acids Res.* **34**: 2607-17.

65. Murphy, B.A., Grundy, F.J., and Henkin, T.M. 2002. Prediction of gene function in methylthioadenosine recycling from regulatory signals. *J. Bacteriol.* **184**: 2314-8.

66. Murray, B.E., Singh, K.V., Ross, R.P., Heath, J.D., Dunny, G.M., and Weinstock, G.M. 1993. Generation of restriction map of *Enterococcus faecalis* OG1 and investigation of growth requirements and regions encoding biosynthetic function. *J. Bacteriol.* **175**: 5216-23.

67. Nahvi, A., Sudarsan, N., Ebert, M.S., Zou, X., Brown, K.L., and Breaker, R.R. 2002. Genetic control by a metabolite binding mRNA. *Chem. Biol.* **9**: 1043.

68. Nelson, A.R., Henkin, T.M., and Agris, P.F. 2006. tRNA regulation of gene expression: interactions of an mRNA 5'-UTR with a regulatory tRNA. *RNA* **12**: 1254-61.

69. Nomura, Y. and Yokobayashi, Y. 2007. Reengineering a natural riboswitch by dual genetic selection. *J. Am. Chem. Soc.* **129**: 13814-5.

70. Nou, X. and Kadner, R.J. 2000. Adenosylcobalamin inhibits ribosome binding to *btuB* RNA. *Proc. Natl. Acad. Sci. USA* **97**: 7190-5.

71. Nudler, E. and Mironov, A.S. 2004. The riboswitch control of bacterial metabolism. *Trends Biochem. Sci.* **29**: 11-7.

72. Ochi, K., Kandala, J.C., and Freese, E. 1981. Initiation of *Bacillus subtilis* sporulation by the stringent response to partial amino acid deprivation. *J. Biol. Chem.* **256**: 6866-75.

73. Ochi, K. and Freese, E. 1982. A decrease in *S*-adenosylmethionine synthetase activity increases the probability of spontaneous sporulation. *J. Bacteriol.* **152**: 400-10.

74. Ontiveros-Palacios, N., Smith, A.M., Grundy, F.J., Soberon, M., Henkin, T.M., and Miranda-Rios, J. 2008. Molecular basis of gene regulation by the THI-box riboswitch. *Mol. Microbiol.* **67**: 793-803.

75. Philippe, C., Eyermann, F., Benard, L., Portier, C., Ehresmann, B., and Ehresmann, C. 1993. Ribosomal protein S15 from *Escherichia coli* modulates its own translation by trapping the ribosome on the mRNA initiation loading site. *Proc. Natl. Acad. Sci. USA* **90**: 4394-8.

76. Posnick, L.M. and Samson, L.D. 1999. Influence of *S*-adenosylmethionine pool size on spontaneous mutation, dam methylation, and cell growth of *Escherichia coli*. *J. Bacteriol.* **181**: 6756-62.

77. Rentmeister, A., Mayer, G., Kuhn, N., and Famulok, M. 2007. Conformational changes in the expression domain of the *Escherichia coli thiM* riboswitch. *Nucleic Acids Res.* **35**: 3713-22.

78. Rodionov, D.A., Vitreschak, A.G., Mironov, A.A., and Gelfand, M.S. 2004. Comparative genomics of the methionine metabolism in Gram-positive bacteria: a variety of regulatory systems. *Nucleic Acids Res.* **32**: 3340-53.

79. Roth, A., Winkler, W.C., Regulski, E.E., Lee, B.W., Lim, J., Jona, I., Barrick, J.E., Ritwik, A., Kim, J.N., Welz, R., Iwata-Reuyl, D., and Breaker, R.R. 2007. A riboswitch selective for the queuosine precursor preQ1 contains an unusually small aptamer domain. *Nat. Struct. Mol. Biol.* **14**: 308-17.

80. Schlax, P.J., Xavier, K.A., Gluick, T.C., and Draper, D.E. 2001. Translational repression of the *Escherichia coli* alpha operon mRNA: importance of an mRNA conformational switch and a ternary entrapment complex. *J. Biol. Chem.* **276**: 38494-501.

81. Serganov, A., Yuan, Y.R., Pikovskaya, O., Polonskaia, A., Malinina, L., Phan, A.T., Hobartner, C., Micura, R., Breaker, R.R., and Patel, D.J. 2004. Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. *Chem. Biol.* **11**: 1729-41.

82. Serganov, A., A. Polonskaia, A.T. Phan, R.R. Breaker, and Patel, D.J. 2006. Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature* **441**: 1167-71.

83. Sharp, J.S. and Bechhofer, D.H. 2003. Effect of translational signals on mRNA decay in *Bacillus subtilis*. *J. Bacteriol.* **185**: 5372-9.

84. Sperandio, B., Polard, P., Ehrlich, D.S., Renault, P., and Guedon, E. 2005. Sulfur amino acid metabolism and its control in *Lactococcus lactis* IL1403. *J. Bacteriol.* **187**: 3762-78.

85. Sudarsan, N., Hammond, M.C., Block, K.F., Welz, R., Barrick, J.E., Roth, A., and Breaker, R.R. 2006. Tandem riboswitch architectures exhibit complex gene control functions. *Science* **314**: 300-4.

86. Suess, B., Fink, B., Berens, C., Stentz, R., and Hillen, W. 2004. A theophylline responsive riboswitch based on helix slipping controls gene expression *in vivo*. *Nucleic Acids Res.* **32**: 1610-4.

87. Tomsic, J., McDaniel, B.A., Grundy, F.J., and Henkin, T.M. 2008. Natural variability in *S*-adenosylmethionine (SAM)-dependent riboswitches: S-box elements in *Bacillus subtilis* exhibit differential sensitivity to SAM *in vivo* and *in vitro*. *J. Bacteriol.* **190**: 823-33.

88. Uhlenbeck, O.C. 1995. Keeping RNA happy. *RNA* **1**: 4-6.

89. Uhlenbeck, O.C. 2003. Less isn't always more. *RNA* **9**: 1415-7.

90. Vitreschak, A.G., Rodionov, D.A., Mironov, A.A., and Gelfand, M.S. 2004. Riboswitches: the oldest mechanism for the regulation of gene expression? *Trends Genet.* **20**: 44-50.

91. Wang, J.X. and Breaker, R.R. 2008. Riboswitches that sense *S*-adenosylmethionine and *S*-adenosylhomocysteine. *Biochem. Cell Biol.* **86**: 157-68.

92. Wang, J.X., Lee, E.R., Morales, D.R., Lim, J., and Breaker, R.R. 2008. Riboswitches that sense *S*-adenosylhomocysteine and activate genes involved in coenzyme recycling. *Mol. Cell* **29**: 691-702.

93. Weaver, K.E., Ehli, E.A., Nelson, J.S., and Patel, S. 2004. Antisense RNA regulation by stable complex formation in the *Enterococcus faecalis* plasmid pAD1 par addiction system. *J. Bacteriol.* **186**: 6400-8.

94. Weigand, J.E. and Suess, B. 2007. Tetracycline aptamer-controlled regulation of pre-mRNA splicing in yeast. *Nucleic Acids Res.* **35**: 4179-85.

95. Weinberg, Z., Regulski, E.E., Hammond, M.C., Barrick, J.E., Yao, Z., Ruzzo, W.L., and Breaker, R.R. 2008. The aptamer core of SAM-IV riboswitches mimics the ligand-binding site of SAM-I riboswitches. *RNA* **14**: 822-8.

96. Wickiser, J.K., Winkler, W.C., Breaker, R.R., and Crothers, D.M. 2005. The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. *Mol. Cell* **18**: 49-60.

97. Winkler, W.C., Grundy, F.J., Murphy, B.A., and Henkin, T.M. 2001. The GA motif: an RNA element common to bacterial antitermination systems, rRNA, and eukaryotic RNAs. *RNA* **7**: 1165-72.

98. Winkler, W.C., Cohen-Chalamish, S., and Breaker, R.R. 2002. An mRNA structure that controls gene expression by binding FMN. *Proc. Natl. Acad. Sci. USA* **99**: 15908-13.

99. Winkler, W., Nahvi, A., and Breaker, R.R. 2002. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**: 952-6.

100. Winkler, W.C., Nahvi, A., Sudarsan, N., Barrick, J.E., and Breaker, R.R. 2003. An mRNA structure that controls gene expression by binding *S*-adenosylmethionine. *Nat. Struct. Biol.* **10**: 701-7.

101. Winkler, W.C., Nahvi, A., Roth, A., Collins, J.A., and Breaker, R.R. 2004. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428**: 281-6.

102. Yocum, R.R., Perkins, J.B., Howitt, C.L., and Pero, J. 1996. Cloning and characterization of the *metE* gene encoding *S*-adenosylmethionine synthetase from *Bacillus subtilis*. *J. Bacteriol.* **178**: 4604-10.

103. Yousef, M.R., Grundy, F.J., and Henkin, T.M. 2005. Structural transitions induced by the interaction between $tRNA^{Gly}$ and the *Bacillus subtilis glyQS* T box leader RNA. *J. Mol. Biol.* **349**: 273-87.

104. Miller, J. 1972. *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

105. Sudarsan, N., Barrick, J.E., and Breaker, R.R. 2003. Metabolite-binding RNA domains are present in the genes of eukaryotes. *RNA* **9**: 644-7.

106. Fox, K.A., Ramesh, A., Stearns, J.E., Bourgogne, A., Reyes-Jara, A., Winkler, W.C., and Garsin, D.A. 2009. Multiple posttranscriptional regulatory mechanisms partner to control ethanolamine utilization in *Enterococcus faecalis*. *Proc. Natl. Acad. Sci. USA* **106**: 4435-40.

107. Johansson, J., Mandin, P., Renzoni, A., Chiaruttini, C., Springer, M., and Cossart, P. 2002. An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*. *Cell* **110**: 551-61.

108. Morita, M.T., Tanaka, Y., Kodama, T.S., Kyogoku, Y., Yanagi, H., and Yura, T. 1999. Translational induction of heat shock transcription factor σ^{32}: evidence for a built-in RNA thermosensor. *Genes Dev.* **13**: 655-65.

109. Narberhaus, F., Waldminghaus, T., and Chowdhury, S. 2006. RNA thermometers. *FEMS Microbiol. Rev.* **30**: 3-16.

110. Mandal, M. and Breaker, R.R. 2004. Adenine riboswitches and gene activation by disruption of a transcription terminator. *Nat. Struct. Mol. Biol.* **11**: 29-35.

111. Sudarsan, N., Lee, E.R., Weinberg, Z., Moy, R.H., Kim, J.N., Link, K.H., and Breaker, R.R. 2008. Riboswitches in eubacteria sense the second messenger cyclic di-GMP. *Science* **321**: 411-3.

112. Dann, C.E., Wakeman, C.A., Sieling, C.L, Baker, S.C., Irnov, I., and Winkler, W.C. 2007. Structure and mechanism of a metal-sensing regulatory RNA. *Cell* **130**: 878-92.

113. Cromie, M.J., Shi, Y., Latifi, T., and Groisman, E.A. 2006. An RNA sensor for intracellular Mg^{2+}. *Cell* **125**: 71-84.

114. Yura, T., Nagai, H., and Mori, H. 1993. Regulation of the heat-shock response in bacteria. *Annu. Rev. Microbiol.* **47**: 321-50.

115. Nocker, A., Hausherr, T., Balsiger, S., Krstulovic, N.P., Hennecke, H., and Naberhaus, F. 2001. mRNA-based thermosensor controls expression of rhizobial heat shock genes. *Nucleic Acids Res.* **29**: 4800-7.

116. Ataide, S.F., Wilson, S.N., Dang, S., Rogers, T.E., Roy, B., Banerjee, R., Henkin, T.M., and Ibba, M. 2007. Mechanisms of resistance to an amino acid antibiotic that targets translation. *ACS Chem. Biol.* **2**: 819-27.

117. Garst, A.D., Heroux, A., Rambo, R.P., and Batey, R.T. 2008. Crystal structure of the lysine riboswitch regulatory mRNA element. *J. Biol. Chem.* **283**: 22347-51.

118. Blount, K.F., Wang, J.X., Lim, J., Sudarsan, N., and Breaker, R.R. 2007. Antibacterial lysine analogs that target lysine riboswitches. *Nat. Chem. Biol.* **3**: 44-9.

119. Rodionov, D.A., Vitreschak, A.G., Mironov, A.A., and Gelfand, M.S. 2003. Regulation of lysine biosynthesis and transport genes in bacteria: yet another RNA riboswitch? *Nucleic Acids Res.* **31**: 6748-57.

120. Mandal, M., Boese, B., Barrick, J.E., Winkler, W.C., and Breaker, R.R. 2003. Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria. *Cell* **113**: 577-86.

121. Wickiser, J.K., Cheah, M.T., Breaker, R.R., and Crothers, D.M. 2005. The kinetics of ligand binding by an adenine-sensing riboswitch. *Biochemistry* **44**: 13404-14.

122. Gutierrez-Preciado, A., Henkin, T.M., Grundy, F.J., Yanofsky, C., and Merino, E. 2009. Biochemical features and functional implications of the RNA-based T-box regulatory mechanism. *Microbiol. Mol. Biol. Rev.* **73**: 36-61.

123. Greenleaf, W.J., Frieda, K.L., Foster, D.A.N., Woodside, M.T., and Block, S.M. 2008. Direct observation of hierarchical folding in single riboswitch aptamers. *Science* **319**: 630-3.

124. Serganov, A., Huang, L., and Patel, D.J. 2009. Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch. *Nature* **458**: 233-7.

125. Vitreschak, A.G., Rodionov, D.A., Mironov, A.A., and Gelfand, M.S. 2002. Regulation of riboflavin biosynthesis and transport genes in bacteria by transcriptional and translational attenuation. *Nucleic Acids Res.* **30**: 3141-51.

126. Thore, S., Leibundgut, M., and Ban, N. 2006. Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand. *Science* **312**: 1208-11.

127. Edwards, T.E., and Ferre-D'Amare, A.R. 2006. Crystal structures of the Thi-box riboswitch bound to thiamine pyrophosphate analogs reveal adaptive RNA-small molecule recognition. *Structure* **14**: 1459-68.

128. Sudarsan, N., Barrick, J.E., and Breaker, R.R. 2003. Metabolite-binding RNA domains are present in the genes of eukaryotes. *RNA* **9**: 644-7.

129. Sudarsan, N., Cohen-Chalamish, S., Nakamura, S., Mitchell Emilsson, G., and Breaker, R.R. 2005. Thiamine pyrophosphate riboswitches are targets for the antimicrobial compound pyrithiamine. *Chem. Biol.* **12**: 1325-35.

130. Nahvi, A., Barrick, J.E., and Breaker, R.R. 2004. Coenzyme B_{12} riboswitches are widespread genetic control elements in prokaryotes. *Nucleic Acids Res.* **32**: 143-50.

131. Mironov, A.S., Gusarov, I., Rafikov, R., Lopez, L.E., Shatalin, K., Kreneva, R.A., Perumov, D.A., and Nudler, E. 2002. Sensing small molecules by nascent RNA: A mechanism to control transcription in bacteria. *Cell* **111**: 747-56.

132. Nou, X., and Kadner, R.J. 1998. Coupled changes in translation and transcription during cobalamin-dependent regulation of *btuB* expression in *Escherichia coli*. *J. Bacteriol.* **180**: 6719-28.

133. Lipfert, J., Das, R., Chu, V.B., Kudaravalli, M., Boyd, N., Herschlag, D., and Doniach, S. 2007. Structural transitions and thermodynamics of a glycine-dependent riboswitch from *Vibrio cholerae*. *J. Mol. Biol.* **365**: 1393-1406.

134. Kwon, M., and Strobel, S.A. 2008. Chemical basis of glycine riboswitch cooperativity. *RNA* **14**: 25-34.

135. Tripp, H.J., Schwalbach, M.S., Meyer, M.M., Kitner, J.B., Breaker, R.R., and Giovannoni, S.J. 2009. Unique glycine-activated riboswitch linked to glycine-serine auxotrophy in SAR11. *Environ. Microbiol.* **11**: 230-8.

136. Meyer, M.M., Roth, A., Chervin, S.M., Garcia, G.A., and Breaker, R.R. 2008. Confirmation of a second natural preQ$_1$ aptamer class in Streptococcaceae bacteria. *RNA* **14**: 685-95.

137. Klein, D.J., Edwards, T.E., and Ferre-D'Amare, A.R. 2009. Cocrystal structure of a class I preQ$_1$ riboswitch reveals a pseudoknot recognizing an essential hypermodified nucleobase. *Nat. Struct. Mol. Biol.* **16**: 343-4.

138. Kang, M., Peterson, R., and Feigon, J. 2009. Structural insights into riboswitch control of the biosynthesis of queuosine, a modified nucleotide found in the anticodon of tRNA. *Mol. Cell* **33**: 784-90.

139. Chamnongpol, S., and Groisman, E.A. 2000. Acetyl-phosphate-dependent activation of a mutant PhoP response regulator that functions independently of its cognate sensor kinase. *J. Mol. Biol.* **300**: 291-305.

140. Ueland, P.M. 1982. Pharmacological and biochemical aspects of *S*-adenosylhomocysteine and *S*-adenosylhomocysteine hydrolase. *Pharmacol. Rev.* **34**: 223-85.

141. Schubert, H.L., Blumenthal, R.M., and Cheng, X. 2003. Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem. Sci.* **28**: 329-35.

142. Klein, D.J., Schmeing, T.M., Moore, P.B., and Steitz, T.A. 2001. The kink-turn: a new RNA secondary structure motif. *EMBO J.* **20**: 4214-21.

143. Grundy, F.J., and Henkin, T.M. 2002. Synthesis of serine, glycine, cysteine and methionine. In *Bacillus subtilis and Its Closest Relatives: from Genes to Cells* (eds. Sonenshein, A.L., Hoch, J.A., and Losick, R.) 245-54 (American Society for Microbiology, Washington, DC, USA).

144. Grundy, F.J. and Henkin, T.M. 1993. tRNA as a positive regulator of transcription antitermination in *B. subtilis*. *Cell* **74**: 475-82.

145. Grundy, F.J., Moir, T.R., Haldeman, M.T., and Henkin, T.M. 2002. Sequence requirements for terminators and antiterminators in the T box transcription antitermination system: disparity between conservation and functional requirements. *Nucleic Acids Res.* **30**: 1646-55.

146. Grundy, F.J., Waters, D.A., Allen, S.H., and Henkin, T.M. 1993. Regulation of the *Bacillus subtilis* acetate kinase gene by CcpA. *J. Bacteriol.* **175**: 7348-55.

147. Donnelly, C.E., and Sonenshein, A.L. 1984. Promoter-probe plasmid for *Bacillus subtilis*. *J. Bacteriol.* **157**: 965-7.

148. Powers, T. and Noller, H.F. 1991. A functional pseudoknot in 16S ribosomal RNA. *EMBO J.* **10**: 2203-14.

149. Hartz, D., McPheeters, D.S., Traut, R., and Gold, L. 1988. Extension inhibition analysis of translation initiation complexes. *Methods Enzymol.* **164**: 419-25.

150. Fredrick, K., and Noller, H.F. 2002. Accurate translocation of mRNA by the ribosome requires a peptidyl group or its analog on the tRNA moving into the 30S P site. *Mol. Cell* **9**: 1125-31.

151. Wabiko, H., Ochi, K., Nguyen, D.M., Allen, E.R., and Freese, E. 1988. Genetic mapping and physiological consequences of *metE* mutations of *Bacillus subtilis*. *J. Bacteriol.* **170**: 2705-10.

152. Dubey, A.K., Baker, C.S., Suzuki, K., Jones, A.D., Pandit, P., Romeo, T., and Babitzke, P. 2003. CsrA post-transcriptionally represses *pgaABCD*, responsible for synthesis of a biofilm polysaccharide adhesin of *Escherichia coli*. *J. Bacteriol.* **185**: 4450-60.

153. Argaman, L., and Altuvia, S. 2000. *fhlA* repression by OxyS RNA: kissing complex formation at two sites results in a stable antisense-target RNA complex. *J. Mol. Biol.* **300**: 1101-12.

154. Altuvia, S., Zhang, A., Argaman, L., Tiwari, A, and Storz, G. 1998. The *Escherichia coli* OxyS regulatory RNA represses *fhlA* translation by blocking ribosome binding. *EMBO J.* **17**: 6069-75.

155. Schlax, P.J., Xavier, K.A., Gluick, T.C, and Draper, D.E. 2001. Translational repression of the *Escherichia coli* alpha operon mRNA: importance of an mRNA

conformational switch and a ternary entrapment complex. *J. Biol. Chem.* **276**: 38494-501.

156. Grundy, F.J., and Henkin, T.M. 2004. Regulation of gene expression by effectors that bind to RNA. *Curr. Opin. Microbiol.* **7**: 126-31.

157. Winkler, W.C., and Breaker, R.R. 2005. Regulation of gene expression by riboswitches. *Annu. Rev. Microbiol.* **59**: 487-517.

158. Henkin, T.M., and Grundy, F.J. 2006. Sensing metabolic signals with nascent RNA transcripts: the T box and S box riboswitches as paradigms. *Cold Spring Harb. Symp. Quant. Biol.* **71**: 231-7.

159. Grundy, F.J., Rollins, S.M., and Henkin, T.M. 1994. Interaction between the acceptor end of tRNA and the T box stimulates antitermination in the *Bacillus subtilis tyrS* gene: a new role for the discriminator base. *J. Bacteriol.* **176**: 4518-26.

160. Gold, O.G., Jordan, H.V., and van Houte, J. 1975. The prevalence of enterococci in the human mouth and their pathogenicity in animal models. *Arch. Oral. Biol.* **20**: 473-7.

161. Dunny, G.M., Brown, B.L., and Clewell, D.B. 1978. Induced cell aggregation and mating in *Streptococcus faecalis*: evidence for a bacterial sex pheromone. *Proc. Natl. Acad. Sci. USA* **75**: 3479-83.

162. Grundy, F.J., and Henkin, T.M. 2003. The T box and S box transcription termination control systems. *Front. Biosci.* **8**: d20-31.

163. Gagnon, Y., Breton, R., Putzer, H., Pelchat, M., Grunberg-Manago, M., and Lapointe, J. 1994. Clustering and co-transcription of the *Bacillus subtilis* genes encoding the aminoacyl-tRNA synthetase genes specific for glutamate and for cysteine and the first gene for cysteine biosynthesis. *J. Biol. Chem.* **269**: 7473-82.

164. Mansilla, M.C., Albanesi, D., and deMendoza, D. 2000. Transcriptional control of the sulfur-regulated *cysH* operon encoding genes involved in L-cysteine biosynthesis in *Bacillus subtilis*. *J. Bacteriol.* **182**: 5885-92.

165. Guillouard, I., Auger, S., Hullo, M.F., Chetouani, F., Danchin, A., and Martin-Verstraete, I. 2002. Identification of *Bacillus subtilis* CysL, a regulator of the *cysJI* operon, which encodes sulfite reductase. *J. Bacteriol.* **184**: 4681-9.

166. Fernandez, M., Kleerebezem, M., Kuipers, O.P., Siezen, R.J., and van Kranenburg, R. 2002. Regulation of the *metC-cysK* operon, involved in sulfur metabolism in *Lactococcus lactis*. *J. Bacteriol.* **184**: 82-90.

167. Shelver, D., Rajagopal, L., Harris, T.O., and Rubens, C.E. 2003. MtaR, a regulator of methionine transport, is critical for survival of group B streptococcus *in vivo*. *J. Bacteriol.* **185**: 6592-9.

168. Greene, R.C. 1996. Biosynthesis of methionine. In *Escherichia coli and Salmonella: Cellular and Molecular Biology* (eds. Neidhardt, F.C. *et al.*) 542-60 (American Society for Microbiology, Washington, DC, USA).

169. Collins, J.A., Irnov, I., Baker, S., and Winkler, W.C. 2007. Mechanism of mRNA destabilization by the *glmS* ribozyme. *Genes Dev.* **21**: 3356-68.

170. Rodionov, D.A., Vitreschak, A.G., Mironov, A.A., and Gelfand, M.S. 2002. Comparative genomics of thiamin biosynthesis in prokaryotes: new genes and regulatory mechanisms. *J. Biol. Chem.* **277**: 48949-59.

171. Park, J., Tai, J., Roessner, C.A., and Scott, A.I. 1996. Enzymatic synthesis of *S*-adenosyl-L-methionine on the preparative scale. *Bioorg. Med. Chem.* **4**: 2179-85.

172. Spoerel, N., and Herrlich, P. 1978. Colivirus-T3-coded *S*-adenosylmethionine hydrolase. *Eur. J. Biochem.* **95**: 227-33.

173. LaMonte, B.L, and Hughes, J.A. 2006. In vivo hydrolysis of *S*-adenosylmethionine induces the *met* regulon of *Escherichia coli*. *Microbiology* **152**: 1451-9.