CHARACTERIZATION OF AN ORPHAN RIBOSWITCH: IDENTIFICATION OF A METAL-SENSING REGULATORY RNA

APPROVED BY SUPERVISORY COMMITTEE

_____
Wade Winkler, Ph. D.

_____
Vanessa Sperandio, Ph. D.

_____
Marie-Alda Gilles-Gonzalez, Ph. D.

_____
Paul Blount, Ph. D.

DEDICATION

I would like to thank everybody would made this project possible. First and foremost, I would like to thank my mentor Dr. Wade Winkler who is truly one of the greatest scientists that I have ever met. His never-ending enthusiasm for my project and all of the other projects in his lab is truly inspiring and his lively scientific discussions have sparked a love of science in me that I would never have thought imaginable. I would also like to thank the members of my dissertation committee: Drs. Vanessa Sperandio, Paul Blount, and Marie-Alda Gilles-Gonzalez who have all been very supportive of my project and provided me with excellent feedback.

Other key players in this project include Dr. Charles E. Dann III who generated the crystal structure of the M-box RNA, and Dr. Arati Ramesh who has done extensive analysis on the COG categories regulated by the M-box RNA, generated the electrostatic surface potential images presented in this work, and

shared her knowledge of metal ion homeostasis with me. I would also like to thank Dr. Chad A. Brautigam for assistance with AUC, Dr. Alexis Kertsburg for assisting with milligram-scale RNA production, Stephanie C. Baker for help with numerous experiments, and Jennifer A. Collins for the construction of the *mgtE-lacZ* transcriptional fusion. Additionally I would like to acknowledge all past and present members of the Winkler lab who have all given my helpful advice on many occasions and made the lab a great environment in these past years. In particular, I would like to acknowledge Jennifer Stearns whose organizational skills are truly awe-inspiring for making the Winkler lab an excellent work atmosphere. Additionally, I would like to give a special acknowledgement to Irnov, a fellow graduate student in the lab, for many a fascinating scientific discussion.

I would also like to thank my family who has been very supportive of me during my graduate studies (and my whole life in general). My husband Beaux Griffith has always been very understanding when I am working late hours and has always been around to help me out when I feel stressed and cheer me up when I feel down. My parents, John and Sharon Wakeman, give me continuous encouragement and are truly the greatest parents a girl could ever ask for. Additionally, I would like to thank my brother who pushed me to do my best by picking on me and saying that I would never publish in any journal other than

RNA and *Nucleic Acids Research* (although I still believe that there is nothing wrong with those particular journals).

CHARACTERIZATION OF AN ORPHAN RIBOSWITCH: IDENTIFICATION OF A METAL-SENSING REGULATORY RNA

by

CATHERINE ANN WAKEMAN

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

December, 2009

CHARACTERIZATION OF AN ORPHAN RIBOSWITCH: IDENTIFICATION OF A METAL-SENSING REGULATORY RNA

CATHERINE ANN WAKEMAN, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2009

WADE C. WINKLER, Ph.D.

Riboswitches are RNA-based genetic control elements found in untranslated regions of the mRNA transcript that they regulate. These RNA motifs are highly structured and bind metabolites to elicit control of gene expression. Typically, the metabolite sensed by these RNAs is a component of the metabolic pathway in which the regulated gene product resides. The focus of this project has been the identification of the ligand for a riboswitch that was discovered using bioinformatics-based search methods. This riboswitch was designated the *ykoK* RNA element due to its location in the 5′ UTR of the *B. subtilis ykoK* (*mgtE*) gene, which appears to be a Mg^{2+} transporter. Therefore,

the possibility that this RNA senses Mg^{2+} levels was explored. The data revealed that the RNA element imparts Mg^{2+}-responsive regulation to the *ykoK* gene. These data also indicated which portions of the RNA are essential for genetic regulation. The results of a battery of biochemical tests demonstrated that Mg^{2+} triggers a concerted conformational change in the RNA such that it adopts a compacted tertiary structure. Resolution of the three-dimensional structure of the RNA in the Mg^{2+} bound state revealed the basis of this metal-induced tertiary conformation and how this relates to genetic control. Intriguingly, this structure revealed the presence of six Mg^{2+}, making this the first example of multiple ligands binding to a single riboswitch aptamer. When individual metal-sites were eliminated using phosphorothioate substitutions, it became evident that all six of these Mg^{2+} and up to three additional metal sites are required for function of this RNA. Therefore, these data demonstrate that the *ykoK* RNA element, now designated the M-box RNA, directly senses intracellular Mg^{2+} levels for the purposes of genetic control. These findings should have broad implications given that this RNA element is wide spread among Gram-positive bacteria and appears to regulate many additional gene categories such as ABC transporters, cell division proteins, and proteins of unknown function. The exploration of the connection between Mg^{2+} concentration and the expression levels of these proteins might provide insights into previously undefined functional roles.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

Wakeman, C. A., Ramesh, A., and Winkler, W. C. (2009). Multiple metal-binding cores are required for metalloregulation by M-box riboswitch RNAs. *J. Mol. Biol.* **392**, 723-735.

Wakeman, C. A. and Winkler, W. C. (2009). Analysis of the RNA backbone: Structural analysis of riboswitches by in-line probing and selective 2´-hydroxyl acylation and primer extension. *Methods Mol. Biol.* **540**, 173-191.

Brautigam, C. A., **Wakeman, C. A.**, and Winkler, W. C. (2009). Methods for analysis of ligand-induced RNA conformational changes. *Methods Mol. Biol.* **540**, 77-95.

Wakeman, C. A. and Winkler, W. C. (2009). Structural probing techniques on natural aptamers. *Methods Mol. Biol.* **535**, 115-133.

Dann, C. E. 3rd*, **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**, 878-892.

Wakeman, C. A., Winkler, W. C., and Dann, C. E. 3rd (2007). Structural features of metabolite-sensing riboswitches. *Trends Biochem. Sci.* **32**, 415-424.

Gore, H. M., **Wakeman, C. A.**, Hull, R. M., and McKillip, J. L. (2003). Real-time molecular beacon NASBA reveals *hblC* expression from *Bacillus* spp. in milk. *Biochem. Biophys. Res. Commun*. **311**, 386-390.

*These authors contributed equally to this work.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Amp	Ampicillin
ATP	Adenosine triphosphate
ATPαS	Adenosine triphosphate phosphorothioate analog
AUC	Analytical ultracentrifugation
BME	β-mercaptoethanol
BSA	Bovine serum albumin
Cm	Chloramphenicol
CTP	Cytidine triphosphate
CTPαS	Cytidine triphosphate phosphorothioate analog
DTT	Dithiothreitol
EC_{50}	Half maximal effective concentration
EDTA	Ethylenediaminetetraacetic acid
EtBr	Ethidium Bromide
FMN	Flavin mononucleotide
FPLC	Fast protein liquid chromatography
GlcN6P	Glucosamine-6-phosphate
GTP	Guanosine triphosphate
GTPαS	Guanosine triphosphate phosphorothioate analog
HPLC	High pressure liquid chromatography
NMIA	N-methylisatoic anhydride
NTP	Nucleotide triphosphosphate

PAGE	Polyacrylamide gel electrophoresis
PCP	Cytidine 3′,5′-bisphosphate
PEG	Polyethylene glycol
RBS	Ribosomal binding site
RT	Room temperature or Reverse transcriptase
SAH	*S*-adenosylhomocysteine
SAM	*S*-adenosylmethionine
SAXS	Small angle X-ray scattering
SHAPE	Selective 2′-hydroxyl acylation analyzed by primer extension
Spec	Spectinomycin
UTP	Uridine triphosphate
UTPαS	Uridine triphosphate phosphorothioate analog
UTR	Untranslated region
TBAB	Tryptone blood agar base
TPP	Thiamine pyrophosphate
WT	Wild-type
X-gal	Bromo-chloro-indolyl-galactopyranoside

CHAPTER ONE

INTRODUCTION AND LITERATURE REVIEW

Regulatory RNAs:

The 'central dogma' states that information is stored as DNA, transcribed into RNA, and translated into proteins which act as the workhorses of the cell to satisfy both structural and catalytic requirements. Each step of this process is tightly regulated. The majority of this regulation is thought to occur through the action of regulatory proteins. However, in recent years, the importance of noncoding RNAs in these processes has become increasingly appreciated (Hannon et al., 2006; Wakeman et al., 2007; Dambach and Winkler, 2009; Winkler, 2005a; Winkler, 2005b).

Prior to the full understanding of the importance and versatility of the regulatory RNAs found in nature, in vitro selection experiments revealed the impressive potential of these molecules as something greater than a simple intermediate step between DNA and protein. Using in vitro selection, RNA molecules possessing the ability to bind chemical compounds with high affinity and specificity can be evolved from random sequences. These nucleic acid-based molecules capable of forming ligand binding pockets were termed 'aptamers', a term derived from the Greek words 'aptus' meaning 'to fit' and 'meros' meaning 'particle.' The ligands bound by these artificially-evolved aptamers range from simple inorganic materials such as ions to small organic molecules such as nucleotides and amino acids to complex ligands such as peptides and proteins

(Stoltenburg et al., 2007). The available structural models (such as those of the ATP-, GTP-, vitamin B_{12}-, and FMN-binding aptamers) reveal the multiple types of interactions employed by these RNAs in order to achieve specific binding of small organic molecules (Figure 1). These RNA-ligand interactions include base-pairing, aromatic stacking, and metal-ion-dependent stabilization of electronegative moieties (Wakeman et al., 2007). In addition to RNA molecules capable of ligand-binding, RNAs with certain catalytic functions have been produced through in vitro selection. RNA-mediated catalytic activities can include self-cleavage, alkyl transfer, phosphorylation, dephosphorylation, redox chemistry, and carbon-carbon bond synthesis (Ellington et al., 2009).

Today it is well known that various types of naturally occurring regulatory RNAs exist. These RNA molecules are capable of eliciting control of gene expression via modulation of transcription termination (e.g., Mironov et al., 2002; reviewed in Gollnick and Babitzke, 2002), translation initiation efficiency (e.g., Winkler et al., 2002; reviewed in Winkler, 2005a and Winkler, 2005b), mRNA stability (e.g., Winkler et al., 2004; Collins et al., 2007), or alternative splicing (Cheah et al., 2007; Wachter et al., 2007). Regulatory RNAs can function either through intermolecular (*trans*) or intramolecular (*cis*) interactions to regulate the expression of the target mRNA. *Trans*-acting RNAs are well-known in eukaryotic organisms (i.e. miRNAs, etc.) but are also common in eubacteria (reviewed in Gottesman et al., 2006; Storz et al., 2006). *Cis*-acting RNAs are most prevalent in bacteria but can also be found in archeabacteria and even eukaryotes (e.g. Cheah et al., 2007; Wachter et al., 2007). *Cis*-acting regulatory RNAs are found within untranslated regions of the mRNA transcript that they regulate and elicit control

over gene expression either through the sole action of the RNA molecule or in conjunction with recruited protein factors.

A classic example of *cis*-acting regulatory RNA mechanism is the control of transcription attenuation, translation initiation, or mRNA stability found in the tryptophan biosynthesis genes of certain bacterial species (Yanofsky, 2004; Gollnick et al., 2005; Yanofsky, 2007). Within this operon, the regulatory RNAs function through feedback inhibition in which high amounts of the biosynthesis proteins lead to a build-up of tryptophan that is sensed through the action of the RNA and accessory factors thereby decreasing expression of the tryptophan biosynthesis operon. One such regulatory RNA-based genetic mechanism incorporates a combination of a stretch of tryptophan codons and RNA sequences that are capable of forming hairpins. Under tryptophan-replete conditions, an intrinsic terminator helix is able to form, shutting off gene expression. Under tryptophan-depleted conditions, translating ribosomes that are coupled with the nascent transcript will become stalled at the stretch of tryptophan codons. This allows for the formation of an antiterminator helix, thereby preventing the formation of the terminator helix and allowing gene expression to occur. Other RNA-based regulatory mechanisms found within tryptophan biosynthesis operons involve tryptophan-binding proteins that associate with different regulatory RNA sequences to elicit control of either transcription attenuation or translation initiation (Yanofsky, 2007).

A more recently characterized class of *cis*-acting regulatory RNA has been termed riboswitches. This type of RNA-based regulation appears to be prevalent in eubacterial organisms. For example, in *Bacillus subtilis,* more than 2.1% of the genome is thought to be under the control of this class of regulatory RNA (Winkler, 2005b).

Riboswitches accomplish genetic regulation through conformational changes that occur upon the binding of these RNAs to a target small molecule metabolite ligand (Wakeman et al., 2007; Schwalbe et al., 2007; Montange and Batey, 2008; Cochrane and Strobel, 2008; Roth and Breaker, 2009). These RNAs contain a distinct aptamer domain that is responsible for ligand binding. This region of the RNA undergoes a conformational change that modulates downstream genetic regulatory features retained within a domain referred to as the expression platform. This domain usually contains multiple but mutually exclusive base-pairing potentials. Depending on the type of expression platform present, ligand binding can be coupled to control of gene expression through modulation of transcription attenuation (Mironov et al., 2002; Winkler et al., 2002a), translation initiation (Winkler et al., 2002b), mRNA stability (Collins et al., 2007), or alternative splicing (Cheah et al., 2007; Wachter et al., 2007) although regulation at the level of transcription attenuation or translation initiation is most commonly observed. In general, Gram-positive microorganisms tend to prefer mechanisms that exert control over transcription elongation while Gram-negative microorganisms more often employ mechanisms that influence translation efficiency. The ‘on’ and ‘off’ base-pairing schemes of the ligand-bound and unbound states of some of the more extensively characterized riboswitch RNAs can be viewed in Figure 2.

Different classes of riboswitches have been shown to respond to a variety of small organic metabolites including amino acids, nucleobases, nucleosides, enzymatic cofactors, an amino sugar, and a bacterial second messenger molecule (Roth and Breaker, 2009; Dambach and Winkler, 2009). So far, evidence of riboswitch function has been obtained for RNAs that respond to adenine, adenosylcobalamin, arginine, bis-(3′-5′)-

cyclic-di-GMP, FMN, GlcN6P, glycine, guanine, lysine, molybdenum and tungsten cofactors (Moco and Tuco), $preQ_1$, SAH, SAM, and TPP (Serganov, 2009; Dambach and Winkler, 2009). Additionally, two riboswitch classes have been proposed to function as direct sensors of intracellular Mg^{2+} (Cromie et al., 2006; Dann et al., 2007). As with the aforementioned *cis*-acting regulatory RNA mechanisms, these RNA elements typically function through feedback inhibition in that they tend to regulate the expression of the genes responsible for the synthesis and transport of their target metabolites.

One unique aspect of riboswitch RNAs is that while accessory protein factors might influence these genetic mechanisms in vivo, they are not required for binding of metabolites to the RNAs. This capability can be attributed to the fact that these RNA sequences contain all the necessary information in order to fold into intricate and complex architectures that allow for the formation of ligand-binding pockets, utilizing every functional group of the RNA in order to recognize their target metabolite.

Folding and architecture of complex RNAs:

RNA sequences fold in a hierarchical manner to form complex structures. This folding pathway proceeds first with formation of secondary structural elements followed by the compilation of tertiary contacts. This process requires the presence of intracellular metal ions. RNA molecules require cations for folding due to the fact that RNA polymers possess highly negatively charged phosphate backbones and these phosphates are brought into close proximity of each other within the final folded structure. Therefore, metal ions associate with RNA polymers to promote secondary structure by offsetting the

electrostatic penalty incurred in bringing together backbone phosphates. This role is generally accomplished by monovalent ions.

In other cases, cations may also assist in promoting tertiary architecture by bridging distantly located regions of an RNA molecule through formation of specific and direct interactions to distantly located RNA functional groups. While monovalent ions can accomplish this task, divalents are generally more essential for tertiary structure formation (Draper et al., 2005; Woodson, 2005; Sigel and Pyle, 2007; Draper 2008; Chu et al., 2008).

Mg^{2+} and K^{+} are the most common divalent and monovalent ions to associate with RNAs, respectively. These metal ions are either tightly bound within the RNA structure or loosely associated with the negatively charged RNA polymer. Tightly bound or chelated Mg^{2+} ions are coordinated by three major functional groups of the RNA: non-bridging phosphate oxygens, ribose oxygens, and various nucleobase groups (ex. O6 carbonyls of guanines, N7 of purines, and O4 of pyrimidines) (Draper et al., 2005; Sigel and Pyle, 2007; Klein et al., 2004). These tightly associated metal ions are oftentimes essential for the formation of commonly observed RNA folds and motifs, and thereby assist in the formation of complex RNA architectures.

Various well known structural models of complex RNAs reveal the presence of essential metal ions. For example, specific Mg^{2+} sites have been identified for TPP-binding RNAs, group I introns, and the ribosome (DeRose, 2003; Hougland et al., 2005; Klein et al., 2004; Selmer et al., 2006; Serganov et al., 2006; Vicens and Cech, 2006). In these RNAs, Mg^{2+} ions participate in ligand stabilization, chemical catalysis, and structural stabilization, respectively. In each of these cases, the RNA contacts the metal

ion directly rather than recognizing the hydrated form of the ion. A direct contact to the ion is known as inner sphere coordination while an indirect contact mediated by a water molecule is known as outer sphere coordination. Multiple inner sphere coordinations typically suggest that these metal binding pockets are particularly important. However, the majority of RNA structural models do not contain any metal ions with more than one inner sphere contact to the RNA (Draper, 2008). This observation underscores the uniqueness of highly specialized Mg^{2+}-binding sites that involve more than one inner sphere coordination and further highlights the fact that, while cations non-specifically interact with RNA polymers in general, certain RNA molecules have evolved special structural relationships with metal ions.

Metal ion homeostasis:

Metal ions are a requirement for life but become an intracellular threat when present in excess. Studies from multiple laboratories have revealed how homeostasis mechanisms are elegantly controlled for a variety of metals (Moore and Helmann, 2005; Pennella and Giedroc, 2005; Geidroc and Arunkumar, 2007). These studies demonstrate that organisms sense intracellular metal concentrations through the action of metal-binding metalloregulatory proteins. These regulatory proteins act to maintain the correct balance of a variety of intracellular metals by regulating the expression of genes responsible for the import, export, and storage of the target ion (Figure 3). In general, free cytoplasmic ion levels are maintained at a relatively low level.

The mechanisms used to maintain the correct level of intracellular metal are well known for certain ions. One of the most studied metal homeostasis pathways is that of Fe^{2+}. The bulk of Fe^{2+}-responsive gene regulation occurs through a dimeric DNA-binding protein known as Fur (Moore and Helmann, 2005). The Fur regulon consists of ~40 genes primarily involved in acquisition of external Fe^{2+}. These genes encode proteins such as membrane-associated transporters and siderophores which are small, high affinity Fe^{2+}-chelating compounds secreted by microorganisms to scavenge Fe^{2+} from the environment. Fur acts as a repressor of gene expression in its Fe^{2+}-bound state; however, when intracellular Fe^{2+} becomes depleted, Fe^{2+} disperses away from the Fur protein, the Fe^{2+}-free form of Fur dissociates from the DNA, and gene expression of the Fur regulon is upregulated. Thus, when intracellular Fe^{2+} levels are low, the cell can attempt to acquire Fe^{2+} from extracellular sources.

Mechanisms governing Mn^{2+} homeostasis have also been well studied in *Bacillus subtilis*. The metalloregulatory protein MntR is a homolog of a well known Fe^{2+}-sensing protein; however, it has been shown to specifically sense intracellular Mn^{2+} levels (Moore and Helmann, 2005). Just as Fur downregulates Fe^{2+} uptake mechanisms in response to increasing levels of intracellular Fe^{2+}, MntR represses two operons encoding Mn^{2+} permeases: an ABC transporter (*mntABCD*) and an NRAMP family transporter (*mntH*).

A similar regulatory trend can be observed for Zur, a Zn^{2+}-sensing paralog of Fur (Moore and Helmann, 2005). Under Zn^{2+}-replete conditions, Zur represses the expression of two Zn^{2+} uptake systems: the *ycdHI-yceA* encoded ABC transporter and the *yciABC* operon encoding a second, lower affinity Zn^{2+} transport pathway. Therefore, the general mechanism of maintaining homeostasis of biologically important metal ions is to express

genes involved in the uptake of these essential ions until intracellular levels of the metal reach a threshold that signals for the down regulation of metal importers and scavengers.

Despite significant advancements in the knowledge of the factors controlling the intracellular levels of most metal ions, relatively little detail is known regarding the regulation of Mg^{2+} homeostasis even though Mg^{2+} is the most abundant divalent metal within cells and is required for numerous biochemical activities. Studies have shown that there are three major families of Mg^{2+} transporters which function to maintain these relatively high levels of cytoplasmic Mg^{2+}: CorA, MgtE, and MgtA/MgtB P-type ATPase proteins (Gardner, 2003; Hmiel et al., 1986; Kehres and Maguire, 2002; Macdiarmid and Gardner, 1998; Maguire, 2006; Smith and Maguire, 1995, 1998). *Bacillus subtilis*, the model organism for most Gram-positive research, contains as many as five potential Mg^{2+} uptake systems: two CorA homologs (YfjQ and YqxL), an MgtA homolog (YloB), an MgtE homolog (YkoK), and a Mg-citrate transporter (CitM) (Moore and Helmann, 2005; Dann et al., 2007). However, the regulation of these transporters in *B. subtilis* has not been studied.

Most of the previous studies focusing on the regulation of Mg^{2+} levels had been performed in Gram-negative enteric bacteria such as *E. coli* and *S. enterica* (Kato and Groisman, 2008). These studies have focused on the two-component regulatory system PhoP/Q. Typically two component systems consist of a membrane-bound histidine kinase that senses an environmental stimulus and relays the signal via phosphoryl transfer to a protein known as the response regulator, ultimately engendering a change in gene expression. In the case of PhoP/Q, PhoQ acts as the membrane-bound sensor kinase with both extracellular and intracellular domains. The extracellular domain senses Mg^{2+} levels

outside of the cell and, upon Mg^{2+}-deprivation, relays a signal to its intracellular domain in order to phosphorylate PhoP and up-regulate the expression of the Mg^{2+} stimulon (reviewed in Eguchi and Utsumi, 2008). Part of the Mg^{2+} stimulon includes Mg^{2+} uptake genes such as the MgtA Mg^{2+} transporter. Other components of the PhoP/Q regulon include virulence factors such as MgtC. While it is known that bacteria sense the depletion of metals such as Fe^{2+} as signals of entry into the host, the sensing of Mg^{2+} as a signal of host entry is a more controversial subject due to the fact that not all researchers believe that the macrophage is a Mg^{2+} depleted environment (Alix and Blanc-Potard, 2007).

Other than the PhoP/Q two-component system, little else is known about the regulation of these families of Mg^{2+} transporters. As observed in the Fur, MntR, and Zur regulons, additional metal-sensing protein based transcription factors may assist in the regulation of Mg^{2+} homeostasis. However, while proteins have historically been shown to fulfill metalloregulatory roles in organisms, the extent to which these functions should be solely ascribed to proteins has not been determined. As detailed in the introductory section on RNA structure, metal ions (Mg^{2+} in particular) play a critical role in the formation of RNA tertiary structure. Therefore, there are only two criteria required in order to render a structured RNA such as a riboswitch into a metal sensor. First, the compacted structure adopted in the presence of the metal must be capable of governing the accessibility of a short nucleotide tract that results in regulatory function by controlling the formation of either an intrinsic terminator or the exposure of the ribosomal binding site (RBS). Second, the K_D or EC_{50} for this metal-induced structural transition must be tuned to an appropriate physiological intracellular concentration in

order to impart sensory function to the RNA. Given these criteria, it is easy to imagine that an RNA molecule could evolve metalloregulatory function.

As of now, two classes of metalloregulatory riboswitches have been characterized and published (Cromie et al., 2006; Dann et al., 2007). Additionally, another potentially Mg^{2+}-responsive UTR has been identified upstream of *mgtCB* in *S. enterica* (Spinelli et al., 2008) although this element has not yet been definitively classified as a riboswitch. All of these RNAs have been most frequently identified upstream of genes involved in Mg^{2+} transport. The M-box RNA element discussed herein regulates examples of all known families of Mg^{2+} transporters (MtgE, MgtA, and CorA) as well as a number of other metal ion transporters and a few unrelated gene categories (Table 1). This RNA element is highly conserved within many Gram-positive and a few Gram-negative bacteria. The large degree of conservation and the relatively large size (~160 nucleotides) of this RNA are surprising for an RNA element proposed to be binding to and sensing such as simple ligand (Figure 4). However, these features also make the RNA a fascinating molecule to study.

Bioinformatics-added methods have also uncovered several putative classes of riboswitch RNAs that oftentimes appear to be regulating metal ion transporters (Barrick et al., 2004). For example, the *ydaO/yuaA* class of orphan riboswitches has been found upstream of predicted K^+ transporters and the *ykkC/yxkD* class has been associated with putative cation transporters. The *ykoY/yybP* element has also been associated with various putative cation transporters including predicted K^+ and Na^+ transporters and Ca^{2+}/Na^+ antiporter systems. Since riboswitches usually regulate genes via feedback inhibition, it

would not be surprising to find that metal ions are the ligands being sensed by these novel RNA classes.

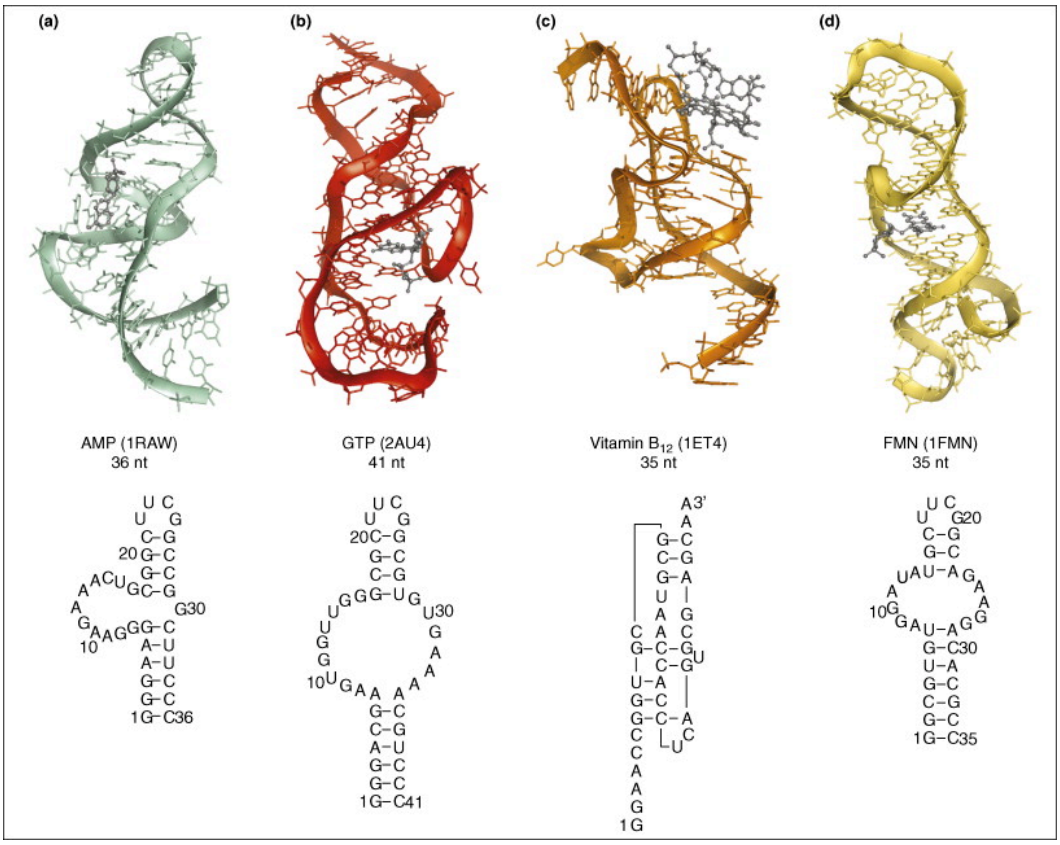


Figure 1: In vitro selection has been used to evolve RNA aptamers capable of binding small organic metabolites with high affinity and specificity. RNA aptamer structures are shown in stick representation with a ribbon drawn through the riboses of the RNA. The bound ligands are indicated, with the PDB accession codes in parentheses and the number of nucleotides. Ligands are shown in gray ball-and-stick representation. The secondary structure and sequence are also indicated. **(a)** An ATP-binding aptamer has been crystallized in complex with AMP (Jiang et al., 1996). This RNA binds ligands with micromolar dissociation constants and binds similar molecules with adenine-like moieties. **(b)** The recently determined GTP aptamer (Carothers et al., 2006) binds its ligand with a K_D of ~75 nM. **(c)** The vitamin B_{12} aptamer (Sussman et al., 2000) binds its ligand peripherally to the surface of the RNA. **(d)** The FMN aptamer (Fan et al., 1996) makes extensive contacts to the isoalloxazine ring of the FMN molecule. Figure from Wakeman, et al., 2007.

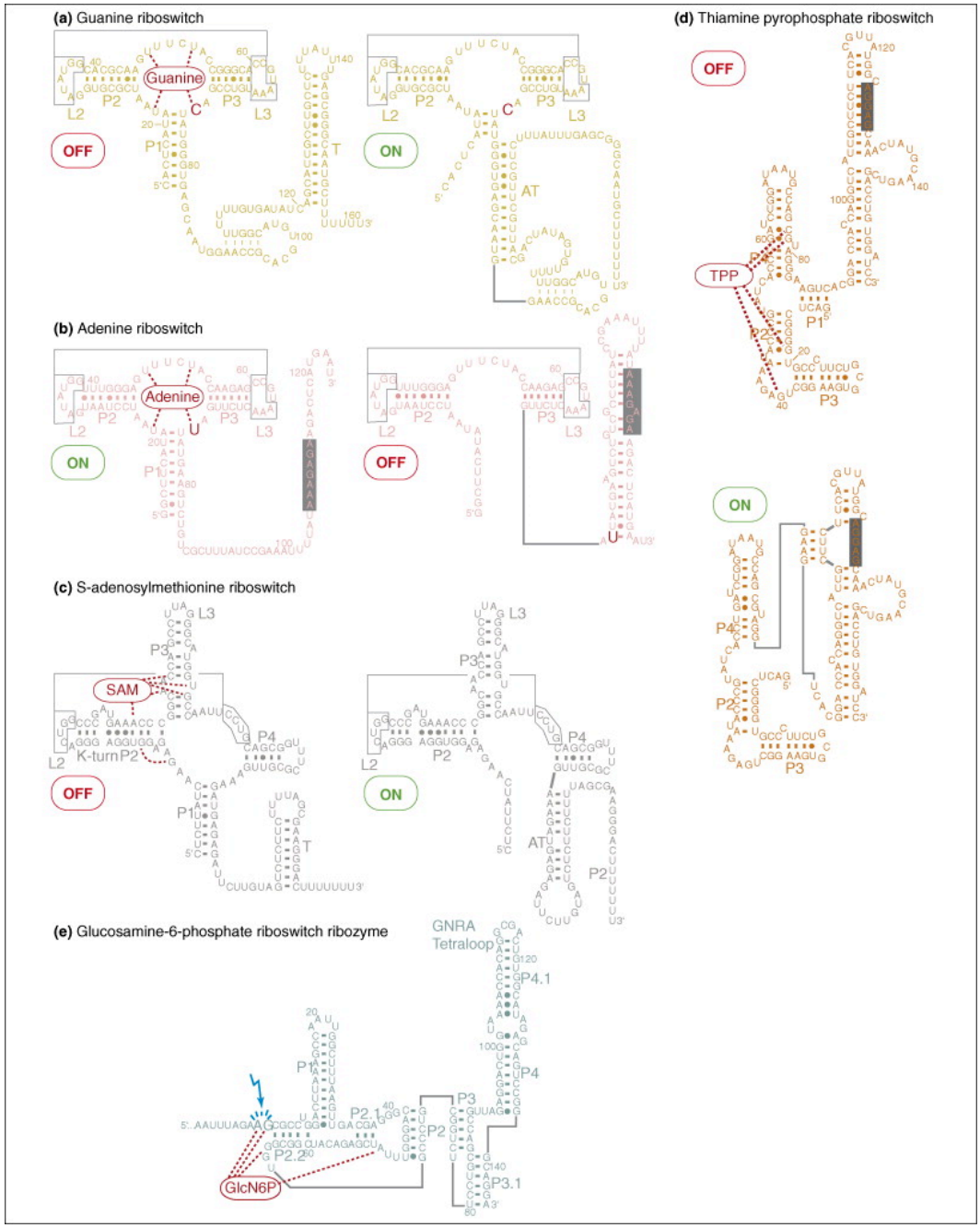

(a) Guanine riboswitch
Guanine
OFF
ON
(b) Adenine riboswitch
Adenine
ON
OFF
(c) S-adenosylmethionine riboswitch
SAM
K-turn
OFF
ON
(d) Thiamine pyrophosphate riboswitch
OFF
TPP
ON
(e) Glucosamine-6-phosphate riboswitch ribozyme
GNRA
Tetraloop
GlcN6P

Figure 2: The diversity of riboswitch-based genetic control mechanisms. Shown are the secondary structures and mechanisms of genetic control for structurally characterized riboswitch classes. **(a)** Guanine riboswitch from *B. subtilis*. **(b)** Adenine riboswitch from *Vibrio vulnificus*. **(c)** SAM riboswitch from *Thermoanaerobacter tencongensis*. **(d)** TPP riboswitch from *T. tencongensis*. **(e)** *glmS* ribozyme (GlcN6P riboswitch) from *E. coli*. The guanine and SAM riboswitches use transcription attenuation through mutually exclusive terminator–antiterminator pairing elements for genetic regulation. The adenine and TPP riboswitches exert regulatory control over translation inhibition. The *glmS* RNA is a metabolite-sensing ribozyme that uses GlcN6P to activate *cis* cleavage of transcripts to control mRNA stability. Comparison of the guanine **(a)** and adenine **(b)** riboswitches highlights how *cis*-acting regulatory RNAs can repress or activate gene expression. Broken red lines denote nucleotide positions that form hydrogen bonds to the metabolite ligands. 'Off' and 'on' indicate conditions where expression of downstream genes is reduced or increased, respectively. Pseudoknot interactions in **(a–c)** are depicted by linked boxes. Gray boxes in **(b)** and **(d)** denote the ribosome-binding site for the downstream gene. Blue arrow in **(e)** indicates the site of ribozyme self-cleavage that occurs in response to GlcN6P binding. Pairing elements (helices) in aptamer domains are assigned the letter 'P' and numbered with respect to their proximity to the 5′ end of the RNA transcript. Similarly, terminal loops are designated by 'L'. Figure from Wakeman et al., 2007.

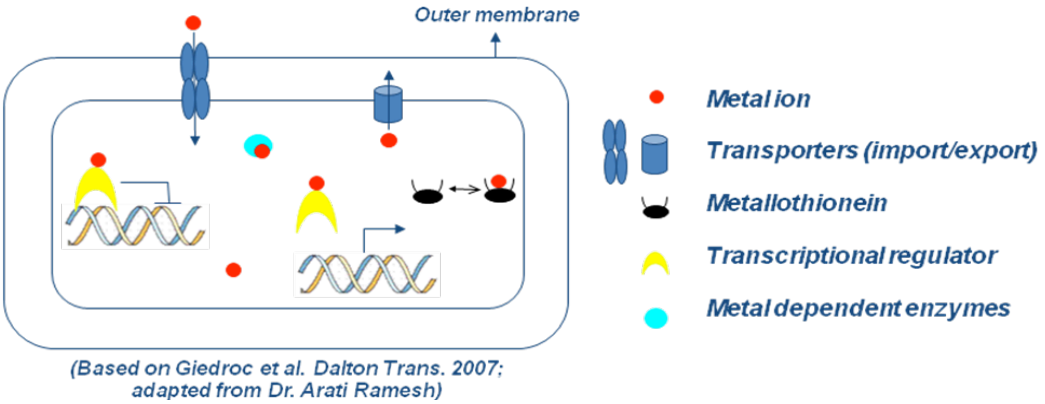


Figure 3: Overview of metal ion homeostasis. Set concentrations of intracellular metal ions are required for the function of metal-dependent enzymes. However, in excess these ions become toxic. Therefore, metal sensing transcriptional regulators regulate the levels of import and export proteins as well as metal sequestering metallothioneins in order to maintain relatively low levels of free intracellular divalent ions.

	Mg(II) Transporter	Cation Transporter	Other Transporter	Other
Aeromonas hydrophila	MgtE (2)			
Alkaliphilus metalliredigens	MgtE			
Bacillus anthracis	MgtA			
Bacillus cereus	MgtA (2)	COG1285S		
Bacillus halodurans	MgtE (4)	COG1285S (2)		COG3464L
Bacillus subtilis	MgtE			
Bacillus thuringiensis	MgtA			
Caldicellulisiruptor saccharolyticus	MgtE	COG1914P		
Carboxydothermus hydrogenoformans	MgtE	COG1914P		
Chromobacterium violaceum	MgtE			
Clostridium acetobutylicum	MgtE	COG1914P (2)		Hypothetical protein
Clostridium thermocellum	CorA			
Desulfotomaculum reducens	CorA			
Enterococcus faecalis	MgtA (2)		COG1116P (2), COG4986P (2)	COG1316K
Enterococcus faecium				COG0772D
Geobacter uraniumreducens	MgtE	COG1914P		
Halorhodospira halophila	MgtE			
Klebsiella pneumoniae	MgtE			
Lactobacillus acidophilus	MgtA			
Lactobacillus brevis			COG4986P	
Lactobacillus johnsonii	MgtA (2)			
Listeria innocua	MgtA			
Listeria monocytogenes	MgtA			COG0772D
Listeria welshimeri	MgtA			COG0772D (2)
Moorella thermoacetica			COG1116P, COG4986P	
Mycobacterium avium		COG1914P		Hypothetical protein (2), COG5651N (2), PE family protein
Mycobacterium bovis				Hypothetical protein, COG5651N (2), PE family protein
Mycobacterium leprae		COG1914P		
Mycobacterium tuberculosis				Hypothetical protein (2), PE protein
Pelobacter propinoicus	MgtE	COG1914P		
Pelotomaculum thermopropionicum	CorA			
Photorhabdus luminescens	MgtE			
Roseiflexus sp.	CorA (2)		COG4986P	
Serratia proteamaculans	MgtE			
Sodalis glossinidius	MgtE			
Streptococcus sanguinis	MgtA			
Syntrophobacter fumaroxidans	CorA	COG1914P		
Thermoanaerobacter tengcongensis	MgtA			
Vibrio cholerae	MgtE			

Table 1: A partial summary of M-box regulated gene categories. Examples of all three types of Mg^{2+} transporters are regulated by the M-box as shown in the 2nd column. Additional M-box reguluated cation transporters include MgtC, a virulence factor also involved in ion transport (COG1285S) and NRAMP family Mn^{2+}/Fe^{2+} transporters (COG1914P). Other transporters include ABC-type transporters (COG1116P) and protein-dependent transporters (COG4986P). Non-transport proteins include transposase (COG3464L), transcriptional regulators (COG1316K), cell division proteins (COG0772D), and PPE family proteins (COG5651N).

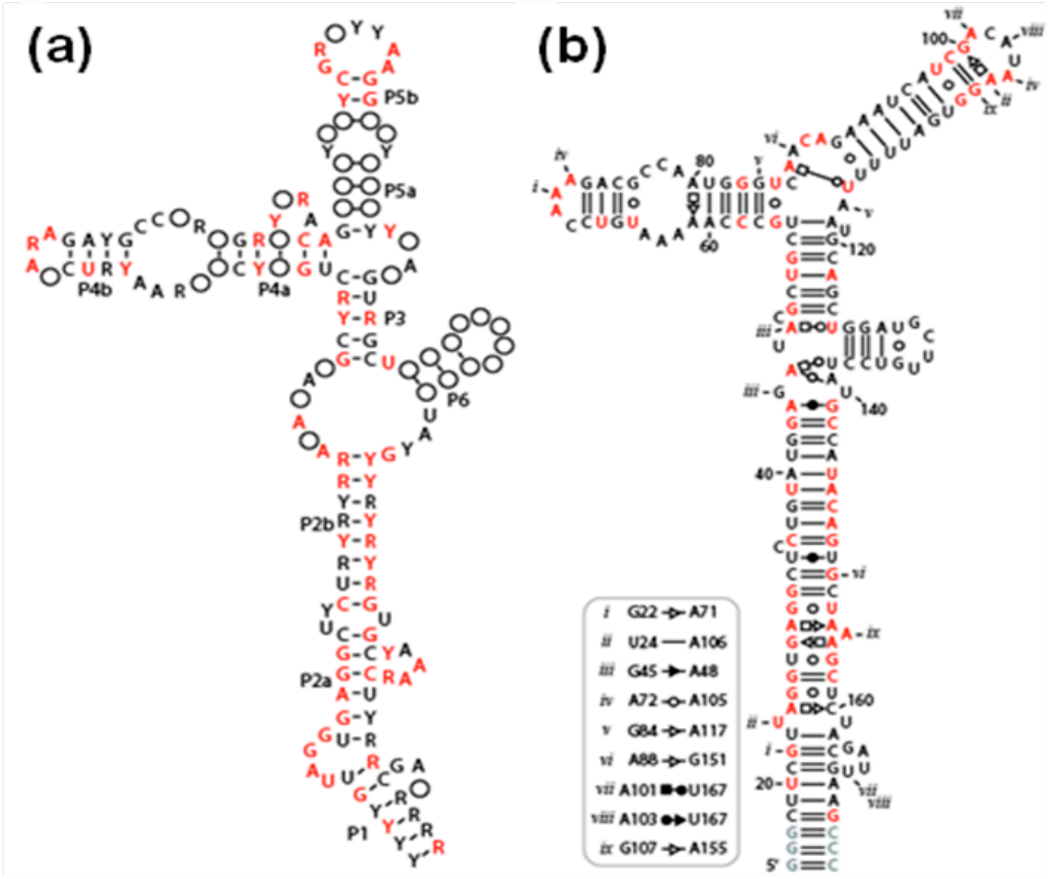


Figure 4: The conserved secondary structure of the M-box RNA element. **(a)** The consensus sequence and secondary structure pattern of the M-box RNA was adapted from Rfam (Griffiths-Jones et. al. 2005) and predicts pairing elements 1-6 (P1-P6). Black and red nucleotides are conserved at greater than 80% and 95% frequency, respectively. Circles represent non-conserved positions. Y = pyrimidine, R = purine. **(b)** Secondary structure of the Mg^{2+}-bound state of the *B. subtilis mgtE* RNA aptamer as determined by X-ray crystallography. The secondary structure is depicted using a geometric nomenclature that offers a useful method for designating noncanonical pairings (Leontis and Westhof, 2001). Circles, triangles, and squares represent bonding interactions utilizing Watson-Crick, sugar, and Hoogsten edges, respectively. Open and shaded symbols specify *cis* and *trans* interactions between nucleotides, respectively. Watson-Crick base pairing is denoted by a double line for GC base pairs and a single line for AU base pairs. GU wobble base pairs are represented by open circles. Gray nucleotides depict nucleotides that are present in the crystallized construct but have been altered from the wild type sequence. Red nucleotides are conserved at greater than 95%. Figure adapted from Dann, Wakeman et al., 2007.

CHAPTER TWO

MATERIALS AND METHODS

DNA Oligonucleotides, Chemicals, Strains, and Plasmids

DNA oligonucleotides were purchased from Integrated DNA Technologies. The listing of all oligonucleotide sequences used in this study can be found in Table 2. Chemicals were purchased from Sigma. All strains were derived from IA40 (Bacillus Genetic Stock Center, Ohio). Fusions to *lacZ* were accomplished via pDG1661, which ectopically inserts into the nonessential *amyE* locus. Correct transformants were checked for Cm resistance (5 μg ml^{-1}) and Spec sensitivity (100 μg ml^{-1}). In order to construct the *mgtE-lacZ* fusion, nucleotides -486 to + 25 relative to the *mgtE* translational start were cloned into pDG1661 via restriction sites added to oligonucleotide primers. Nucleotides -487 to -143 and -487 to -235 were cloned into pDG1661 for construction of promoter-*lacZ* and aptamer-*lacZ* fusions, respectively. Mutations were introduced using the QuikChange mutagenesis protocol (Strategene) and were verified by DNA sequencing. DNA was transformed as described (Jarmer et al., 2002).

Growth Conditions

In general, cells were incubated at 37 °C in glucose minimal medium [0.5% glucose, 0.5 mM $CaCl_2$, 5 μM $MnCl_2$, 15 mM $(NH_4)_2SO_4$, 80 mM K_2HPO_4, 44 mM KH_2PO_4, 3.9 mM sodium citrate, and 50 μg ml^{-1} amino acids (tryptophan, methionine,

lysine)]. Cultures containing 2.5 mM $MgCl_2$ were grown at 37 °C overnight, without shaking. The following day, cells were incubated while shaking, cultured to mid-exponential growth phase, pelleted (t_0), washed twice in medium lacking Mg^{2+}, and resuspended to a final OD_{600} of 0.1. $MgCl_2$ was added to 5 μM or 2.5 mM for ‘‘low’’ or ‘‘high’’ conditions, respectively, prior to incubation again at 37 °C. At $t_{5.5(hours)}$, cells were harvested for analysis. Growth of manganese-limited cells was achieved similarly except that 2.5 mM $MgCl_2$ was maintained throughout the experimentation and at OD_{600} = 0.1 $MnCl_2$ was either eliminated or added at 5 μM for ‘‘low’’ and ‘‘high’’ conditions, respectively. For control of iron levels, either 5 μM $FeCl_2$ or 100 μM of the iron chelator 2,2′-dipyridyl (Biachoo et al., 2002) was added to cultures at OD_{600} = 0.1 to generate ‘‘high’’ or ‘‘low’’ iron conditions, respectively. For assays involving limitation of all divalents, cell pellets were resuspended in glucose minimal medium lacking $MgCl_2$, $CaCl_2$, and $MnCl_2$. Each divalent was individually added to cultures at a concentration of 50 μM.

β-galactosidase Assays

For measurements of β-galactosidase activity, cells were harvested by centrifugation and assays were performed in triplicate using a protocol similar to that described previously (Miller, 1972). Briefly, cell pellets were resuspended in 1.0 ml Z buffer (Miller, 1972). 10 μl of toluene was added and the cell suspension was vortexed for 30 seconds whereupon the toluene was evaporated under a hood for 15 minutes. 0.15 ml (0.6 mg) orthonitrophenyl-β-*D*-galactoside was added to 0.75 ml permeabilized cell

suspension to initiate the enzymatic assay (T_i). Reactions were terminated (T_f) by addition of 0.375 ml 1M Na_2CO_3, centrifuged briefly, and analyzed for absorbance at A_{420}. Data was analyzed relative to control reactions containing Z buffer alone. Miller Units were calculated using the equation $[(1000 \times A_{420})/((T_f - T_i) \times 0.75 \text{ ml} \times A_{600})]$. All A_{420} readings were taken within 20 minutes of reaction termination.

Quantitative RT-PCR

Quantitative RT-PCR was performed in as previously described by Bookout et al., 2006 using the Applied Biosystems 7900HT 384-well format system. Total RNA was isolated from WT *B. subtilis* IA40 grown for 5 hours under 'high' (2.5 mM) Mg^{2+} or 'low' (5 μM) Mg^{2+} conditions using a combination of bead-beating and RNAwiz (Ambion) and DNase treated. cDNA was generated using the *i*script reaction mix from Bio-Rad. qRT-RCR was performed using the *i*Taq reaction mix (Bio-Rad) and quantified based on SYBR green fluorescence. A ΔΔCt analysis was performed (Bookout et al., 2006) and data was normalized to 5S rRNA, a highly expressed transcript whose levels remain unchanged in fluctuating Mg^{2+} conditions.

Microarrays

Total RNA for microarray analyses was isolated from WT *B. subtilis* grown for 5 hours under 'high' (2.5 mM) Mg^{2+} or 'low' (5 μM) Mg^{2+} conditions using a combination of bead-beating, LETS buffer (0.1 M LiCl, 10 mM EDTA, 10 mM Tris HCl pH 7.4, 1%

SDS), and TRIzol (Invitrogen). These RNAs were subsequently DNase treated. cDNA was generated using Superscript reverse transcriptase III and random primers (Invitrogen) and RNA was digested with alkaline treatment that was subsequently neutralized through the addition of HCl. cDNA was fragmented with 0.2 U DNase / 1 µg cDNA. These cDNAs were then biotinylated using terminal deoxynucleotide transferase. These fragmented, biotinylated cDNA samples were then sent to the UT Southwestern Microarray Core Facility for further manipulation and analysis.

Northern Blotting / mRNA half-lives

Total RNA for Northern blot analysis was isolated from WT *B. subtilis* grown for 5.5 hours under 'high' (2.5 mM) Mg^{2+} or 'low' (5 µM) Mg^{2+} conditions. With the intention of determining mRNA half-lives, each culture was grown in a 60 mL volume. At 5.5 hours, 14 mL of sample was removed, pelleted, and immediately frozen in liquid nitrogen to serve as the t_0 time point. Transcription was then ceased with the addition of 9 mL of 1 mg/mL rifampin. At 1, 3, 10, 20, and 30 minutes, 10 mL aliquots were removed, pelleted and frozen in liquid nitrogen. Total RNA was extracted using a combination of bead-beating and PureZOL (Bio-Rad). RNAs were resolved on formaldehyde-containing agarose gels. A sample of RNA and a small amount of an RNA size marker were mixed with EtBr and resolved on the gel adjacent to the samples that were to be analyzed by Northern blot. The ribosomal bands from the sample and the RNA size markers allow a standard curve to be generated so that RNA bands within the experimental samples can later be mapped. The resolved RNA samples were then transferred onto a nylon

membrane using an OWL transfer apparatus set at 77 mA / 8.5X4.5 cm gel for 1 hour. Samples were crosslinked to the membrane with UV using 120 mJ for 30 seconds. The membrane was dried by baking at 85 °C for 15 minutes and probed using body-labeled RNA probes complimentary to the *ykoK* UTR. The membrane was washed with 2X RT low stringency washes (1X SSC, 0.1% SDS, 1 mM EDTA) for 5 minutes, 1X 42 °C high stringency wash (0.2X SSC, 0.1% SDS, 1 mM EDTA) and 1X 65 °C high stringency wash for 15 minutes. The Northern blot was then wrapped in plastic wrap and imaged using a phosphorimaging screen.

RNA preparation

DNA templates for in vitro studies of different RNAs were generated by PCR using appropriate oligonucleotide primer set which incorporated the T7 promoter sequence. PCR products were then prepared using the QIAGEN PCR clean-up kit or ethanol precipitated with two volumes of cold 100% ethanol. For wild-type M-box RNAs, the DNA template was amplified from *B. subtilis* IA40 chromosomal DNA (Bacillus Genetic Stock Center, Columbus, Ohio). Aptamer domain RNAs included positions +14 to +172 relative to the endogenous transcriptional start site of the *B. subtilis mgtE* gene. Full-length RNAs for in-line probing and SHAPE included positions +1 – 220 and +1 – 265, respectively. An extra G was added at the 5'-terminus to facilitate T7 transcription. For M3 mutant RNAs, the DNA template was amplified from a plasmid containing a version of the M-box in which nucleotides 107-109 had been mutated from GGU to CCA. TPP-, SAM-, and FMN-binding riboswitches were also amplified from the

B. subtilis genome. Additional M-box RNA constructs were amplified from the *B. cereus, V. cholerae* and *C.acetobutilicum* genomes. RNAs were synthesized in vitro at 37 ˚C in 25 – 50 µl reaction mixtures that included 10 – 30 pmol DNA templates, 30 mM Tris-HCl (pH 8.0), 10 mM DTT, 0.1% Triton X-100, 0.1 mM spermidine-HCl, 2.5-5.0 mM each NTP (Roche), 40 mM $MgCl_2$, and ~50 µg mL^{-1} T7 RNA polymerase. Reactions were terminated after 2.5 hours with 2 volumes of 8 M urea. Products were resolved by denaturing 6% PAGE and RNA-containing bands were excised, cut into ~ 1 mm cubes, and equilibrated in 200 mM NaCl, 10 mM Tris-HCl (pH 7.5), and 10 mM EDTA (pH 8.0) for < 2 hours at 23 ºC. The passively eluted RNAs were ethanol precipitated, resuspended in water, quantified by A_{260} and utilized as markers during nondenaturing gel electrophoresis.

In-line Probing

RNA substrates for in-line probing were dephosphorylated using calf intestinal alkaline phosphatase (New England Biolabs) and 5'-radiolabeled using T4 polynucleotide kinase (New England Biolabs) and γ-^{32}P ATP (Amersham). Reactions contained ~1 nM RNA, 50 mM Tris-HCl (pH 8.3), 100 mM KCl, and specified concentrations of $MgCl_2$ or cobalt hexammine. Certain reactions (described in text) also included 2 M NaCl. RNAs were incubated at 23 ºC for ~40 hours and products were resolved by 10% PAGE adjacent to control lanes containing: partial digestion by RNase T_1 (cleavage after G), a hydroxyl cleavage ladder (cleavage at every position), and an aliquot of unreacted RNA.

See Wakeman and Winkler, 2009a and Wakeman and Winkler, 2009b for further details on the in-line probing technique.

Hydroxyl Radical Footprinting

Hydroxyl radical probing procedures were modestly adapted from a published protocol (Kieft et al., 1999). These reactions consisted of 30 mM HEPES pH 7.4, 1 µg yeast tRNA, 1 µl of a 10X Mg^{2+} solution, and ~100 kcpm of ^{32}P 5'-labeled RNA in a final volume of 7 µl. This mixture was incubated at 37 °C for 5 min. A Fe-EDTA solution (20 mM EDTA pH 8.0; 10 mM ferrous ammonium sulfate) was made fresh prior to each experiment. This solution was mixed 1:1 with 50 mM sodium ascorbate and 2 µl of this mixture was spotted in the lid of each reaction. A 0.03% H_2O_2 solution was also prepared and 1 µl of this solution was spotted on a separate area of each lid. The reactions were initiated by briefly centrifuging the tubes. The reactions were then incubated at 37 °C for 2 min. At this point, the reactions were stopped with the addition of a mixture containing 169 µl of H_2O, 20 µl of 3 M sodium acetate pH 5.2, and 1 µl glycogen, followed by the addition of 500 µl of 100% ethanol. The reactions were precipitated, resuspended in a 2:1 ratio of H_2O and formamide loading buffer (95% formamide; 20 mM EDTA pH 8.0; bromophenol blue and xylene cyanol), and resolved by 10% denaturing PAGE alongside non-reacted RNA, partial RNase T_1, and partial alkaline digestion ladders. See Wakeman and Winkler, 2009b for further details on the hydroxyl radical footprinting technique.

Selective 2´-Hydroxyl Acylation and Primer Extension (SHAPE) Probing

SHAPE assays were performed essentially as described (Merino et al., 2005) with minor alterations. 1 pmol RNA was added to a 9 µl solution containing 5 µl of 2X SHAPE buffer (100 mM Tris-HCl (pH 8), 200mM KCl), and 1 µl of a 10X $MgCl_2$ or other divalent ion solution. In experiments using high monovalent ions, 2X SHAPE buffer was supplemented with 4 M NaCl. These solutions were equilibrated for 2 hours at 23°C, whereupon 1 µl of freshly prepared 65 mM N-methylisatoic anhydride (NMIA) solution was added, followed by incubation for 45 minutes at 37 °C. Reactions were ethanol precipitated and resuspended in 11 µl water with the appropriate 5'-^{32}P-labeled oligonucleotide [CAW18 (5' ccgcacctcctttgattgg 3') for full length constructs and SCB4 (5' ccttcaatcgtagagc 3') for aptamer RNAs]. Extension by reverse transcriptase (Roche) was performed as described elsewhere (Merino et al., 2005). See Wakeman and Winkler, 2009a for further details on the SHAPE probing technique.

3′ ^{32}P end labeling of RNAs

For some of the in-line probing gels shown in this work, the RNA molecules were labeled at the 3´ terminus. This can be advantageous if one wants to better observe structural transitions occurring at this end of the molecule. To accomplish this type of labeling, the radioactive nucleotide cytidine 3´,5´-bisphosphate (PCP) with 5´ ^{32}P is incorporated into the RNA polymer using T4 RNA ligase. For ideal labeling, combine T4 RNA ligase, 10 pmol RNA, 5 µL PCP, 50 µM, 10% DMSO, 10 mM DTT, and 27.5%

PEG in a 40 µL final volume. RNA should be phenol:chloroform extracted and EtOH precipitated prior to being resolved by 6% PAGE.

Analysis of Structural Probing

All polyacrylamide (National Diagnostics) gels were visualized and quantified using a PhosphorImager (Amersham) with ImageQuant software (Molecular Dynamics). Regions of increased, decreased, or unchanging reactivity were designated based upon the direct comparison of volume reports for line traces that had passed through all bands for reactions containing low or high Mg^{2+} (ImageQuant). A difference map between line traces was graphed; peaks and valleys denoted positions that were modified in their reactivity during conditions of high or low Mg^{2+}. EC_{50} values were estimated by plotting the normalized fraction of RNA cleaved (in-line probing) or individual RT stops (SHAPE) versus the logarithm of the concentration of Mg^{2+} in molar units. These data were also analyzed with curve-fitting functions (SigmaPlot) in order to determine the concentration needed to induce half-maximal modulation (EC_{50}).

Reverse Transcriptase (RT) Stop Assay

A primer complimentary to the *ykoK* UTR approximately 40 nucleotides downstream of the terminator element was end-labeled at the 5′ terminus. Reaction mixes contained approximately 10 pmol of WT or mutant RNA constructs encompassing the full length aptamer and expression platform domains plus about 45 nucleotides. These

RNAs were combined with 0.02 mM dNTP, ~1,000,000 cpm primer, commercial buffer, and M-MulV reverse transcriptase, an enzyme that is highly sensitive to RNA secondary structure, in a 20 µL volume and incubated at 37 °C for 1 hour. The reverse transcriptase reactions were resolved by 10% PAGE adjacent to sequencing ladders generated using the Thermosequenase DNA sequencing kit (USB).

5′ (Promoter) Mapping

WT *B. subtilis* IA40 was grown to mid-log phase in glucose minimal media containing 5 µM Mg^{2+}. Total RNA was extracted using a combination of bead-beating and RNAwiz (Ambion) and DNase treated. A DNA oligo complimentary to a portion of the M-box aptamer domain (caw36 – ggcgtctttggacattttaccgcagcagcgtatc) was radioactively labeled at the 5′ terminus with ^{32}P. Reaction mixes were set up with Transcriptor RT, a reverse transcriptase known to be relatively insensitive to RNA structure. These reactions also contained ~2.5 µg/µL of RNA, 0.4 mM dNTPs, commercial buffer, and ~500,000 cpm radioactively labeled primer. The reaction mix was incubated at 42 °C for 2 hours. The reverse transcriptase reactions were resolved by 10% PAGE adjacent to sequencing ladders generated with the Thermosequenase DNA sequencing kit (USB) using the WT M-box:*lacZ* fusion plasmid as template.

Transcription Termination Assays

Templates for transcription termination assays were generated by PCR. All templates contained a deletion of the first three transcribed nucleotides. Wild-type templates included sequence from -455 to -143 relative to the downstream AUG translational start site. M3 templates were the same as wild-type except they included the M3 site-directed mutations. A template was also prepared that was identical to wild-type except that it contained a deletion of the position immediately following the terminator helix, C221, as described in text. For templates lacking the aptamer domain but retaining the terminator element, the promoter region (positions -455 to -399 relative to the translational start site) was directly fused to nucleotides -212 to -143 (which encompassed the terminator helix). The template for the *B. cereus* M-box RNA was comprised of positions -563 to -85 relative to the translational start site of the *bc4140* gene. The template for the SAM-binding riboswitch encompassed positions -273 to +57 relative to the translational start site of the *B. subtilis yjcI* gene. 50 ng DNA templates containing the endogenous *B. subtilis mgtE* promoter sequence were incubated at 37°C for 30 minutes with 25 µl 2X transcription buffer (140 mM Tris-HCl (pH 8.0); 140 mM NaCl; 0.2 mM EDTA; 28 mM BME; 70 µg ml-1 BSA), 5 µl 10X $MgCl_2$ solution, 2 µCi γ-^{32}P-UTP, and 0.5 U *E. coli* RNA polymerase σ^{70} holoenzyme (Epicentre), mixed at a 35 µl final volume. Transcription was initiated with the addition of 15 µl of a freshly prepared NTP mix (0.32 mM ATP, CTP, and GTP; 0.1 mM UTP; 0.2 mg ml^{-1} heparin) and the reactions were incubated at 37°C for 10 minutes whereupon they were terminated by the addition of 8 M urea and resolved by denaturing 6% PAGE. Unless noted, all

reactions included *E. coli* RNA polymerase. The transcription termination assays using *B. subtilis* RNA polymerase were carried out in a 25 µl volume containing 3 µCi γ-^{32}P-UTP, the appropriate concentration of $MgCl_2$, ~2 U RNA polymerase, 500 ng DNA template, buffering conditions identical to the *E. coli* RNA polymerase reactions, 0.5 mM ATP, CTP, and GTP, and 0.05 mM UTP. These reactions were incubated for 30 minutes at 37 °C.

Velocity Sedimentation

A Beckman Analytical Ultracentrifuge XL-I with absorbance optics was used for all velocity sedimentation experiments. Samples and reference buffer (390 µl each) were loaded into dual-sector charcoal-filled epon centerpieces, which were equilibrated to 20 °C for 1 hour. A_{260} readings were measured while the samples were centrifuged in an An60Ti rotor at 42,000 rpm for at least 4 hours. All samples contained RNA concentrations yielding at least 0.2 absorbance units at A_{260}. SEDNTERP was used to calculate the density and viscosity of each buffer (Laue et al., 1992). The Mg^{2+} titration experiments were performed in a buffer consisting of 10 mM Tris pH 7.5, 100 mM KCl, and the specified Mg^{2+} concentration. Experiments analyzing the role of K^+ in the Mg^{2+}-induced conformational change were performed in a buffer consisting of 20 mM Tris pH 8.0, either 0 or 100 mM KCl, and either 30 µM or 10 mM $MgCl_2$. All data was analyzed through use of SedFit v9.4 (Schuck, 2000). Analyses assume a partial specific volume (vbar) of 0.53 cm^3 gm^{-1} and a hydration value of 0.59 g/g (Takamoto et al., 2002). These analyses yielded values for the sedimentation coefficient, R_H, a/b, and D. Sedimentation

coefficients were converted into $s_{20,w}$ by multiplying by the conversion factor obtained with the following equation:

$$(\eta_{T,b}/\eta_{20,w}) \times ((1-\rho_{20,w}*vbar)/(1-\rho_{T,b}*vbar))$$

where $\eta_{T,b}$ is the viscosity of the buffer, $\eta_{20,w}$ is the viscosity of water at 20 °C, $\rho_{20,w}$ is the density of water at 20 °C, vbar is the vbar of the RNA at 20 °C, and $\rho_{T,b}$ is the density of the buffer. See Brautigam et al., 2009 for further details on probing global conformational changes in RNAs using velocity sedimentation.

RNA Size Exclusion Chromatography

M-box RNAs were synthesized in vitro, ethanol precipitated, and resuspended in 10 mM Tris-HCl pH 7.5 and 20 mM EDTA or H_2O. All chromatography experiments were performed at room temperature using an 18-1900-26 ÄKTA^{TM} FPLC and a Superdex 200 10/300 GL column with a flow rate of 0.5 ml/min. Fractions were collected in 0.5 ml increments. The running buffer contained 10 mM Tris-HCl (pH 7.5), 100 mM KCl, and either 10 mM, 1 mM or 0.5 mM $MgCl_2$. See Brautigam et al., 2009 for further details on separating RNA molecules using size exclusion chromatography.

For RNAs containing random phosphorothioate substitutions fractions corresponding to the extended and compacted RNA species, were collected, concentrated, and radioactively labeled at the 5′ terminus prior to separation by nondenaturing polyacrylamide gel electrophoresis. For experiments demonstrating Mg^{2+}-induced changes in hydrodynamic radius, 100 µL of RNA yielding 0.5-2.0 OD_{254} was loaded onto the column that was equilibrated with 10 mM Tris HCl pH 7.5 and either 30

μM or 10 mM $MgCl_2$. For experiments analyzing the ability of other divalents to substitute for Mg^{2+}, the column was equilibrated with 10 mM of the specified metals. The elution volume of the wild type construct with each divalent was compared to that of the M3 mutant construct. The latter mutant has been shown under high Mg^{2+} conditions to possess essentially the same structure as the wild type construct during low Mg^{2+} conditions. The elution volume of the RNA under each condition was tracked through UV absorbance at A_{254}. KCl-free RNA was purified via the a similar method except the column was equilibrated with a buffer containing 20 mM Tris pH 8.0 and 30 μM $MgCl_2$.

Crystallization of the M-box RNA

Structural coordinates were obtained through a collaboration with Dr. Charles E. Dann III using methods detailed in Dann et al., 2007. These coordinates were deposited in the RCSB Protein Data Bank under accession code 2QBZ.

Non-denaturing PAGE

All RNAs were dephosphorylated using calf intestinal alkaline phosphatase (New England Biolabs), 5'-radiolabeled using T4 polynucleotide kinase (New England Biolabs) and γ-^{32}P ATP (Amersham), and resolved by 6% denaturing gel electrophoresis. Radiolabeled RNAs were then excised, passively eluted into 200 mM NaCl, 10 mM Tris-HCl (pH 7.5), and 10 mM EDTA (pH 8.0), ethanol precipitated and resuspended in RNase-free water. 8% nondenaturing polyacrylamide gels (29:1 acrylamide:

bisacrylamide) were prepared with either 1 mM $MgCl_2$, 0.5 mM $MgCl_2$, or with a mixture of 0.5 mM $MgCl_2$ and 0.5 mM $CdCl_2$ and Tris-borate buffer containing the same metal ion concentrations. Samples were electrophoresed at constant 20 watts for 4 hours at 4 °C. RNA-containing bands were visualized by autoradiography, excised and cut into ~ 1 mm cubes, and equilibrated in 200 mM NaCl, 10 mM Tris-HCl (pH 7.5), and 10 mM EDTA (pH 8.0) for < 2 hours at 23 ºC. These passively eluted RNAs were then ethanol precipitated and resuspended in water to approximately match 100,000 cpm/µL, as verified by scintillation counting. See Brautigam et al., 2009 for further details on separating different conformations of the same RNA molecule using non-denaturing/native PAGE.

Random Phosphorothioate Substitution

M-box RNAs with randomly incorporated phosphorothioate modifications were transcribed in 1 mL reaction mixtures that included 0.5 – 1 nmol DNA template, 30 mM Tris-HCl, (pH 8.0), 10 mM DTT, 0.1% Triton X-100, 0.1 mM spermidine-HCl, 1 mM each NTP (Roche), 40 mM $MgCl_2$, ~50 µg mL^{-1} T7 RNA polymerase, 1 U yeast inorganic pyrophosphatase (Sigma), and either 50 µM ATPαS, 100 µM CTPαS, 50 µM GTPαS, or 50 µM UTPαS (Glen Research). The levels of nucleotide analogs added to the transcription reactions were chosen for ~5% incorporation overall, allowing for single phosphorothioate substitutions per RNA molecule (Christian and Yarus, 1992; Christian and Yarus, 1993). Reactions were terminated after 2.5 hours with the addition of 1 mL buffered phenol, followed by extraction with phenol:chloroform. The RNA transcripts

were ethanol precipitated and resuspended in 125 µL water. 100 µL were removed for size exclusion chromatography (conducted at room temperature) and subsequent nondenaturing gel electrophoresis (resolved at approximately 4 ºC). The remaining RNA was retained for use as the unselected, parent RNA pool to which the extended and compacted RNA molecules were compared. After initial separation by size exclusion chromatography, the reactions were radioactively labeled at the 5′ end and resolved by denaturing PAGE prior to separation by nondenaturing PAGE.

Visualization and quantitation of phosphorothioate interferences

The four individual RNA pools, resulting from incorporation of ATPαS, CTPαS, GTPαS, or UTPαS and radiolabeled with P^{32} at the 5' terminus, were matched by scintillation counting. Specifically, two aliquots of 500,000 cpm RNA were each resuspended in 8 µL water and 1 µL yeast total RNA (Ambion). Cleavage at phosphorothioate substitutions was induced for one of the aliquots via addition of 1 µL of 100 µM iodine dissolved in ethanol. 1 µL water was added to the second aliquot as a control population for background cleavages. 5 µL of 3x formamide loading buffer (95% formamide, 20 mM EDTA, 0.01% bromophenol blue, and 0.01% xylene cyanol) was added to each tube and the reactions were heated to 95 ˚C for 2 minutes prior to resolution by 10% denaturing polyacrylamide gel electrophoresis. The resulting gels were dried and analyzed by phosphorimaging instrumentation (Molecular Dynamics Typhoon) and phosphor storage screens (Amersham).

Individual band intensities were quantified using Semi-Automated Footprinting Analysis (SAFA) (Das et al., 2005b) and ImageQuant software (Molecular Dynamics) via standard methodology. Individual bands corresponding to sites of random phosphorothioate substitutions were quantified by SAFA both for untreated RNA pools and those that were subjected to iodine-mediated oxidation. Background intensities (resulting from analysis of untreated RNAs) were subtracted from band intensities for iodine-treated RNAs. Preliminary interference values were obtained by dividing the values measured for the parental (unselected) pool by values measured for individual selected pools. These data were normalized to account for minor variations in loading. This was accomplished by calculating a mean value from all positions within a particular data set, excluding a small subset of positions whose interference values were greater than 5 or less than 0.2. All values within a given data set were divided by the average of values falling within one standard deviation of this mean. The resulting interference values were arbitrarily deemed strong if they exceeded 10, moderate if they exceeded 5, and weak if they exceeded 2. Similar analyses of phosphorothioate incorporation experimentation are reported elsewhere (Ryder and Strobel, 1999; Jansen et al., 2006). All data shown in this paper are quantified from the average of at least two independent experiments.

Nucleotides 128-170 could not easily be analyzed using SAFA alone due to the fact that these regions were not well resolved; therefore, these regions were compared using line trace software in ImageQuant and any positions showing potential interferences were analyzed by individual band boxing quantification. Also, increased levels of background cleavages at positions A25, A48, A70, A79, C86-A90, and C102-

A106 obscured analyses of these positions for the extended RNAs to varying degrees, but analyses of compacted RNAs remained relatively free of background and were therefore analyzed at all positions.

Structural Images

Structural figures were generated using UCSF Chimera (Petersen et al., 2004) or PyMol (Delano, 2002).

Electrostatic Surface Potential

Calculations were performed by Arati Ramesh as described in Stahley et al. 2007 and Wakeman et al., 2009.

Table 2: List of oligonucleotides used in this study

Name	Sequence	Description
caw18	5' ccgcacctcctttgattgg 3'	Rev. primer for ykoK UTR + Term. + 45bp
caw19	5' taatacgactcactataggcagcatctcgttaggcgaggc 3'	Fwd primer for vch vc1655 ykoK + T7
caw20	5' gccgcaacaacagagtgg 3'	Rev. for vc1655 ykoK + Term.
caw23	5' taatacgactcactatagggtctctgttaggtgaggc 3'	fwd for b. cereus bc4140 ykoK + T7
caw24	5' gtaaggtgcatatagacg 3'	rev. for b. cereus bc4140 ykoK
caw31	5' taatacgactcactataggtaattgtagtcatcgactggg 3'	fwd for spiR UTR of Salmonella + T7
caw32	5' gcagcgtattcttgag 3'	rev for spiR UTR
caw33	Gtgtgtgaattcaattaggaac	ykok no promoter + EcoR1
caw34	taatacgactcactataggaacttcgttaggtgagggacctg	mutate bulge in ykok P2 + T7
caw35	gatacgctgctgcggtaaaatgtccaaagacgcc	mutation of ykoK P3
caw36	ggcgtctttggacattttaccgcagcagcgtatc	mutation of ykoK P3
caw37	catcgacataaccagatttttaatgcagctggatg	mutation of ykoK P4 "M3 mutant"
caw38	catccagctgcattaaaaatctggttatgtcgatg	mutation of ykoK P4 "M3 mutant"

caw41	cttcgttaggtgaggcaggtgtatggagatacgc	mutation of bulge in ykoK P2 "M4 mutant"
caw42	gcgtatctccatacacctgcctcacctaacgaag	mutation of bulge in ykoK P2 "M4 mutant"
caw46	cttgtttggtaccgtaattgtgatgtaagcgc	mutation of ykoK P2
caw47	ctttttagaattctaatgagtgaatgaatgcagtggc	mutation of ykok P2
caw48	ttttctgtaatctcagtgtaattcaattctttcaccttcaatccc	ykok terminator alone-adds 1/2 ykoK promoter
caw49	actttttttgagttgacatagcttgtttttttctgtaatctcagtg	adds the rest of the ykok promoter to caw48 PCR
caw50	ctttcaccttcaatcccgaaggtttttttatgcc	mutation to eliminate C221
caw51	ggcataaaaaaaaccttcgggattgaaggtgaaag	rev. primer to go with caw50
caw55	caataccgaattcttttgctgtattttcttttag	C. a ykoK CA_C3329 UTR Fwd + EcoRI
caw56	Acctagggatcccaacagcttgttc	C. a ykoK CA_C3329 UTR Rev + BamHI
caw57	agaaaggaattcaaaagtagtatctgcg	C. a. ykoK CA_C0683 UTR Fwd + EcoRI
caw58	catatacggatccataaagtacttcactc	C. a. ykoK CA_C0683 UTR Rev + BamHI
caw59	taatacgactcactatagggataatgtattgttaggtgaggc	C. a. CA_C3329 aptamer Fwd + T7 promoter
caw60	Atacactatctttcgttgagc	C. a. CA_C3329 aptamer Rev
caw61	taatacgactcactatagggaaaagttttagttaggtgaggc	C. a. CA_C0683 aptamer Fwd + T7
caw62	Tacgaaattagacagtagagc	C. a. CA_C0683 aptamer Rev
caw78	cagaaatcatcgacaAaaggtgatttttaatgcagc	B.s ykoK U104 to A fwd
caw79	gctgcattaaaaatcaccttTtgtcgatgatttctg	B.s ykoK U104 to A rev
caw84	caattaggaacttcgtAaggtgaggctcctgtatgg	B.s. ykoK U24 to A fwd
caw85	ccatacaggagcctcacctTacgaagttcctaattg	B.s. ykoK U24 to A rev
caw90	cctatgccatacagtgctaCagctctacgattgaag	B.s. ykoK A155 to C fwd
caw91	cttcaatcgtagagctGtagcactgtatggcatagg	B.s. ykoK A155 to C rev
caw92	Gatcttaagaaaacattggaagg	B.s. tenA UTR fwd
caw93	Ggatggacaaagctcccctccc	B.s tenA UTR rev
caw94	Aacggttcggcacacgaag	B.s. yjcl UTR fwd
caw95	Gacttcatcgctacggttccc	B.s. yjcl UTR rev
caw96	Tttatatatgttagtttgccatacc	B.c. bc1581 fwd
caw97	Tatgggtaacggaaccgagtc	B.c. bc1581 rev
caw98	Aaatccgtttaatattttaaaaacacag	B.c. bc4140 fwd
caw99	Aacgtatgggtatctgggccg	B.c. bc4140 rev
caw102	taatacgactcactataggtCtCtGttAGGtG	fwd. for b. cereus bc4140 ykoK aptamer + T7
caw103	Acctctccctcgttgag	rev. for b. cereus bc4140 ykoK aptamer
caw104	gctgctgcccaaaaatgtcGaaagacgccaatg	fwd. mutation of B.s. M-box tetraloop
caw105	cattggcgtctttCgacatttttgggcagcagc	rev. mutation of B.s. M-box tetraloop
caw125	gacgccaatgggtccacagaaatcatcgac	forward oligo for B.s. M-box A87C mutation
caw126	gtcgatgatttctgtggacccattggcgtc	reverse oligo for B.s. M-box A87C mutation
caw127	cgacataaggtgatttttcatgcagctggatgc	forward oligo for B.s. M-box A117C mutation
caw128	gcatccagctgcatgaaaaatcaccttatgtcg	reverse oligo for B.s. M-box A117C mutation
caw129	cgctgctgcccaaaaatgtccacagacgccaatg	forward oligo for B.s. M-box A71C mutation
caw130	cattggcgtctgtggacatttttgggcagcagcg	reverse oligo for B.s. M-box A71C mutation
caw152	gagatacgctgctgcccaTTGGtgtccaaagacgc	fwd mutation of Bs Mbox to eliminate P4 bulge
caw153	gcgtctttggacaCCAAtgggcagcagcgtatctc	rev mutation of Bs Mbox to eliminate P4 bulge
caw154	gacgccaatgggtcaCcagaaatcatcgac	forward oligo for B.s. M-box A88C mutation

caw155	gtcgatgatttctgGtgacccattggcgtc	reverse oligo for B.s. M-box A88C mutation
caw156	cagaaatcatcgacatCaggtgatttttaatgcagc	forward oligo for B.s. M-box A105C mutation
caw157	gctgcattaaaaatcacctgatgtcgatgatttctg	reverse oligo for B.s. M-box A105C mutation
caw158	gctaaagctctacgatAgaaggcgcccgcacgc	forward oligo for B.s. M-box U167A mutation
caw159	gcgtgcgggcgccttcTatcgtagagctttagc	reverse oligo for B.s. M-box U167A mutation
jc1	Ggccctgaattcgttccgtaattgtgatg	fwd ykoK UTR + EcoRI
jc2	gtaaaggatcctcctgataggtcatgttttg	rev ykoK UTR + BamHI
scb3	taatacgactcactatagggaacttcgttaggtg	fwd ykoK aptamer + T7
scb4	Ccttcaatcgtagagc	rev ykoK aptamer
wcw34	Catccttctcccatccagac	rev oligo for ribD aptamer
wcw405	taatacgactcactatagggtatccttcggggcaggg	fwd oligo for ribD aptamer + T7
wcw89	taatacgactcactataggacttcctgacacgaa	fwd oligo for yitJ +T7
wcw453	Ctcttcttatcttccaagc	rev oligo for yitJ aptamer
wcw77	taatacgactcactataggcagaacaattcaatatg	fwd oligo for tenA + T7
wcw452	Gctccacttccctacgc	rev oligo for tenA

CHAPTER THREE

GENETIC ANALYSIS OF A METALLOREGULATORY RNA

INTRODUCTION

In addition to the known, well-characterized classes of riboswitch RNAs, bioinformatics-aided efforts have uncovered several RNA elements that share traits with established riboswitches but whose metabolic signals remain to be identified (Barrick et al., 2004; Corbino et al., 2005). These computer-based search methods identify candidate regulatory RNAs using several characteristic features: these RNAs are typically found in relatively long UTRs of mRNAs, the ligand binding domain (aptamer) of a riboswitch is evolutionarily conserved within all organisms containing a particular class of riboswitch, and they have a domain known as the expression platform which has a low level of evolutionary conservation at the primary sequence level but does have the potential to participate in genetic control processes either through transcription attenuation or translation initiation by forming an intrinsic terminator helix or an RBS-sequestering helix. These putative riboswitch classes that have been predicted using bioinformatics are known as orphan riboswitches due to the fact that the ligand sensed by these RNA molecules remains unknown.

One of the first orphan riboswitch classes to be identified was originally termed the *ykoK* element due to its location upstream of the *ykoK* (*mtgE*) gene of *B. subtilis* (Barrick et al., 2004). This element has now been renamed the M-box RNA and will be

referred to as such throughout this paper. This structural motif is generally located upstream of Mg^{2+} transport genes in several Gram-positive and a few Gram-negative bacteria (Table 1) (Barrick et al., 2004; Griffiths-Jones et al., 2005). Mg^{2+} transport in bacteria is achieved primarily through the use of three protein families: CorA, MgtE, and MgtA/MgtB P-type ATPase proteins (Gardner, 2003; Hmiel et al., 1986; Kehres and Maguire, 2002; Macdiarmid and Gardner, 1998; Maguire, 2006; Smith and Maguire, 1995, 1998). Depending on the organism, M-box RNAs can be found in UTRs of all three classes of transport genes, although they appear to primarily function as regulatory elements for members of the *mgtA* and *mgtE* families. While M-box RNAs are usually coupled with transport proteins, the RNA element is also localized with genes of unknown function (Table 1) (Barrick et al., 2004; Griffiths-Jones et al., 2005; Ramesh and Winkler, 2009).

The major focus of my project upon joining the Winkler lab was to determine the ligand for the M-box RNA class. Generally, with a few exceptions, riboswitches tend to regulate gene expression through feedback inhibition meaning that expression of the regulated genes leads to higher levels of a molecule that is, in turn, sensed by the riboswitch in order to shut down gene expression. Therefore, since this element was oftentimes found upstream of Mg^{2+} transporters, it appeared likely that this RNA would be responsive to intracellular Mg^{2+} levels either by directly associating with Mg^{2+} or by binding to a metabolite that fluctuates in response to changing Mg^{2+} levels.

RESULTS

Reporter gene fusions

B. subtilis contains at least one homolog for all three of the major Mg^{2+} transporter classes, although only the *mgtE* homolog is preceded by the M-box RNA. Manual inspection of the *mgtE* locus revealed a candidate promoter element that was confirmed by 5′ mapping, indicating that the *mgtE* mRNA contains a 409 nucleotide 5′ UTR that encompasses the M-box RNA element (Figure 5a). This putative promoter region was fused to a *lacZ* reporter and shown to be functional. Visual inspection of this UTR also revealed the presence of a candidate intrinsic transcription terminator, an observation that was supported by reverse transcriptase stop experiments (Figure 5b) and structural probing data discussed in subsequent chapters.

To investigate the basis of genetic regulation via the M-box RNA element, the entire 5′ UTR of *mgtE* was fused to *lacZ* and ectopically integrated into the genome. Additionally, constructs (designated M1 and M2) containing mutations in relatively unconserved portions of the M-box aptamer were made to test whether structural integrity of the aptamer domain was required for genetic regulation. The M1 mutation altered a nonconserved portion of the P5 helix and the M2 mutation was designed to restore base-pairing to this helix despite the alteration to the primary sequence. The initial tests on these constructs were performed using Petri dishes containing glucose minimal media with either normal amounts of Mg^{2+} (2.5 mM) referred to as ‘high’ or no Mg^{2+} (~0 μM) referred to as ‘low’. After growth overnight at 37 °C, all of these constructs appeared as

light blue colonies and none stood out as exhibiting significant differences in the level of gene expression. This initial observation was highly disappointing; however, upon incubation at room temperature over the course of several days, these colonies gradually became bluer and the differences were more evident. Each of the constructs appeared brilliantly blue on the low Mg^{2+} media. However, on the high Mg^{2+} media, the WT M-box and the M2 restoration mutant appeared to be capable of repressing reporter gene expression to some extent, appearing as lighter blue colonies as compared to the colonies on the low Mg^{2+} media (Figure 6). In contrast, the M1 disruption mutant was still very blue even in Mg^{2+}-replete media. These observations supported the hypothesis that the M-box RNA is responsive to Mg^{2+}-depletion. In addition, these data appeared to support the P5 pairing as predicted by sequence analyses. The fact that the *lacZ* reporter gene fusion constructs required several days of incubation to exhibit any difference between the high/normal and low Mg^{2+} conditions suggested that this method of Mg^{2+} limitation could be improved. Therefore the liquid culture method that is described in detail in Chapter 2 was utilized for the remainder of the reporter gene fusion experiments. Under solid media conditions, the bacteria seemed capable of acquiring trace amounts of Mg^{2+} from the media due to the fact that the bacteria on the 'no' Mg^{2+} media were still able to grow at approximately the same rate as the bacteria on the Mg^{2+}-rich media. However, upon growth in liquid culture, Mg^{2+}-depletion seemed more effective. Under ~0 mM Mg^{2+} conditions in liquid culture, the bacteria would not grow at all. Therefore, 5 μM Mg^{2+} was supplemented into the low Mg^{2+} liquid media which allowed for approximately two cell doublings to occur over the course of 5 hours of growth at 37 °C.

In addition to providing accurate and reproducible Mg^{2+} starvation results, the liquid culture method facilitated the observation of a Mg^{2+}-depletion-induced growth phenotype in WT *B. subtilis* cells that could be interesting and may warrant subsequent studies. Cells grown in low Mg^{2+} conditions for ~5 hours become elongated and filamentous as demonstrated by both light microscopy and electron microscopy (Figure 7). This observation becomes rather interesting when the additional classes of M-box-regulated genes that are not involved in Mg^{2+} transport are taken into account. In at least three microorganisms, *Enterococcus faecium*, *Listeria monocytogenes*, and *Listeria welshmeri* the M-box RNA potentially regulates the expression of FtsW-like cell division proteins (Table 1). While in *B. subtilis*, no known Mg^{2+} responsive system appears to regulate cell division genes, the ability to divide is clearly tied to Mg^{2+} levels. Therefore, a general trend of connections between cell division and Mg^{2+} homeostasis begins to emerge. This observation also raises the question of what other undiscovered Mg^{2+}-responsive mechanisms exist in *B. subtilis* that lead to cell division phenotype.

To determine whether or not general metal depletion leads to changes in gene expression for M-box-regulated genes, cells were cultured under 'normal' defined glucose minimal medium conditions (i.e. adequate metal concentrations) except that a different divalent ion was selectively depleted during each experiment (Figure 8b-d). Expression of *lacZ* was substantially increased in response to Mg^{2+} deprivation, but not upon depletion of Mn^{2+} or Fe^{2+}. As a correlative experiment, cells were depleted for all divalent ions and *lacZ* expression levels were measured upon addition of individual metals (Figure 8f). Only the addition of Mg^{2+} repressed *lacZ* expression under these conditions. Together, these data demonstrated that *mgtE-lacZ* expression is selectively

repressed by Mg^{2+} in vivo. This regulatory response did not occur at the level of transcription initiation as expression of the promoter:*lacZ* construct remained unchanged with varying Mg^{2+} (Figure 8e). To rule out the possibility that Mg^{2+}-responsive regulation was due to RNA structural defects in general, *lacZ* fusions to unrelated RNAs were also tested. In contrast to *mgtE*, Mg^{2+} limitation did not increase expression of genes regulated by either SAM- or GlcN-6-P-sensing RNAs (Figure 8e).

In order to demonstrate that this Mg^{2+} responsive regulation was due to the M-box riboswitch, a mutational analysis on different regions of the M-box aptamer domain was performed. As discussed previously, reporter gene fusions to the M1 and M2 mutations had already suggested that Mg^{2+} responsiveness appeared to rely upon an intact aptamer domain via the solid media culture assays (Figure 6). However, to be more quantitative, these assays were performed in liquid culture so that Miller units, the classical measurement of *lacZ* reporter gene expression, could be recorded. M1 reduced Mg^{2+}-induced repression to 2.2- from 7.5-fold (Figure 8e). The compensatory mutations (M2) designed to restore base-pairing within P5 helix partially restored Mg^{2+} repression, similar to what was observed in solid culture. While the M1 and M2 mutations were introduced in non-conserved regions of the M-box, other mutations (M3 and M4) were designed to disrupt more conserved portions of the RNA. The M3 mutation altered a highly conserved part of P5 and exhibited almost complete elimination of Mg^{2+}-induced repression (Figure 8a,e). Because this particular mutant appeared to be the most disruptive to M-box function, it has been used as a control to represent the non-Mg^{2+}-associated state of the M-box in many subsequent experiments. A similar loss in Mg^{2+}-responsiveness resulted from alteration of a conserved side-bulge within P2 (M4).

Together, these data are consistent with expectations that the aptamer region functions as a sensory domain for an intracellular metabolic signal.

In order to function as a riboswitch, the M-box RNA must couple ligand binding in the aptamer domain to a functional readout in the expression platform. As discussed earlier, a putative terminator helix can be found downstream of the M-box aptamer domain but upstream of the *ykoK* (*mgtE*) coding region. The 5′ portion of the terminator helix appeared to be capable of pairing with the 3′ portion of a helix (P1) within the M-box domain, thereby forming an antiterminator helix, a mutually exclusive base pairing scheme whose formation would prevent the formation of the terminator helix (Figure 8a). Most other members of this regulatory RNA class were also found to contain candidate terminator and antiterminator elements (Figure 9) (Barrick et al., 2004; Griffiths-Jones et al., 2005). Together, these observations suggest that M-box RNAs control expression of downstream genes via regulated terminator formation, a mechanism commonly referred to as transcription attenuation (Landick and Yanofsky, 1987). As a preliminary test of this mechanism, the aptamer domain was fused to *lacZ*, thereby deleting the downstream terminator region. Expression was increased 6.5-fold relative to the full-length sequence, reflecting the loss of the termination signal (Figure 8e). This construct also exhibited a loss of Mg^{2+}-induced repression, demonstrating that the terminator helix is required for regulation. Site-specific mutation of the terminator (M5) also led to a similar increase (5-fold) in Mg^{2+}-independent expression. In contrast, compensatory mutations designed to restore the terminator helix but disrupt the antiterminator (M6) decreased overall expression and diminished responsiveness to Mg^{2+}. The M6 mutant's partial responsiveness to Mg^{2+} is likely due to the presence of some base pairing potential

remaining in the 5′ portion of antiterminator helix. Together, these mutational analyses demonstrate that interruption of the interplay between terminator and antiterminator elements deleteriously impacts Mg^{2+}-responsive regulation of *mgtE* in vivo. Therefore, both the aptamer and expression platform domains of M-box riboswitch are important for its function in vivo.

Effect of Mg^{2+}-depletion on transcript abundance

All of the previous experimentation using reporter gene fusions focused on detection of gene expression at the protein level. However, if the M-box RNA is regulating gene expression via transcription attenuation, Mg^{2+}-responsiveness could be demonstrated at the level of transcript abundance. Transcript abundance under Mg^{2+}-depleted growth conditions was first analyzed using Northern blotting. The Northern blots were initially performed in an attempt to measure the half-life of M-box regulated genes to determine whether or not Mg^{2+} levels affect mRNA stability/turnover. Originally these experiments were inspired in part by the observation that in *E. faecalis* an M-box RNA can be found within a putative 3′ UTR of a transcript. This observation suggested the possibility that a given riboswitch could function in the 3′ UTR if it controlled mRNA stability. An example of a riboswitch capable of controlling mRNA stability has already been found in the 5′ UTR of the *glmS* gene in certain Gram-positive bacteria (Winkler et al., 2004; Collins et al., 2007). However, riboswitches controlling mRNA stability could theoretically also function at the 3′ end of transcripts. Additional examples of riboswitch RNAs located within 3′ UTRs have been recently proposed (Toledo-Arana,

et al., 2009). Therefore, it would be very interesting to determine how riboswitch RNAs could potentially regulate gene expression from the 3′ UTR.

Unfortunately, mRNA half-lives of the full length *mgtE* transcript under Mg^{2+}-replete conditions were never determined due to very low signal as compared to bands representative of ribosomal RNA (Figure 10b). These rRNA bands were particularly problematic due to the fact that 30S subunit RNA is approximately the same size as the full-length *mgtE* transcript. Interpretation of these Northern blots was further complicated by the presence of a number of intermediate RNA species appearing in high abundance under the 'low' Mg^{2+} conditions (Figure 10a). While the origin of these intermediates is unknown, they might be degradation intermediates.

Despite the issues complicating data analysis, these Northern blots were still informative. The mRNA species resolving at the predicted sizes for the full-length and terminated transcripts supported overall the transcription attenuation mechanism as predicted for the M-box regulatory RNA. Under Mg^{2+}-replete conditions, the dominant species appeared to be a short product corresponding to termination within the 5′ UTR. Upon Mg^{2+} deprivation, less of this terminated species was present and a band representative of the full-length transcript could be observed. While the M-box RNA found in the 3′ UTR of *E. faecalis* might be controlling mRNA stability, the transcript analysis of the *B. subtilis mgtE* suggested that the *B. subtilis* M-box RNA does not control stability and instead regulates transcription termination.

In addition to Northern blotting, quantitative real-time PCR (qRT-PCR) was performed in order to determine the effect of Mg^{2+} deprivation on transcript abundance. Interestingly, the *B. subtilis* genome contains two candidate *corA* genes (*yfjQ, yqxL*), a

candidate *mgtA* gene (*yloB*), one *mgtE* homolog (*ykoK*), as well as a Mg^{2+}-citrate transporter (*citM*) (Moore and Helmann, 2005; Dann et al., 2007). qRT-PCR reveals that under Mg^{2+}-limited conditions, transcript abundance of *mgtE* was significantly increased while *corA* and *mgtA* transcript levels were relatively unchanged (Figure 11). Additionally, as would be predicted, the levels of the *mntA* transcript previously shown to fluctuate in response to Mn^{2+} deprivation but not the depletion of many other divalent ions (Que and Helmann, 2000) remain constant upon Mg^{2+} starvation. According to some reports, the *corA* homolog *yqxL* appears to be constitutively expressed while the other Mg^{2+} transporters in *B. subtilis* are regulated via unknown mechanisms in response to unidentified signals (Moore and Helmann, 2005 and references therein). However, these data indicate that the baseline expression of *yqxL* is lower than that of *yloB*. These expression data together suggest that Mg^{2+} transporters other than MgtE do not appear to be significantly regulated in response to Mg^{2+} starvation in *B. subtilis*. Therefore, while other Mg^{2+}-responsive systems may exist in *B. subtilis*, they do not appear to be involved in the regulation of Mg^{2+} homeostasis via control of Mg^{2+} transport.

In order to identify additional genes regulated in response to Mg^{2+}-depletion, microarray analyses were performed in which *B. subtilis* cells grown under 'normal' and 'low' Mg^{2+} conditions were compared. Because RNA was isolated from the 'low' Mg^{2+} cells after they were grown for 5.5 hours after Mg^{2+}-depletion, many pleiotropic effects were noted that would likely be observed under any conditions of general stress. Approximately 1/3 of the genomic transcripts exhibited increased abundance, 1/3 of the genomic transcripts exhibited decreased abundance, and 1/3 of the genome remained unchanged. The global changes observed could simply be a result of whatever factor is

causing the build-up of intermediate species under Mg^{2+}-depletion on the Northern blot (Figure 10a). The origin of these intermediates is unknown but they are likely either degradation products or the remnants of aborted transcripts. Most metallonucleases utilize Mg^{2+} as a cofactor although some require Ca^{2+} or Zn^{2+} instead (Dupureur, 2008). Hence, prolonged Mg^{2+} deprivation would likely lead to a build-up of incompletely processed degradation products. RNA polymerase is also a Mg^{2+}-dependent enzyme. Ergo, insufficient intracellular Mg^{2+} might also lead to the accumulation of aborted transcripts. However, a Mg^{2+} sensor like the M-box should be able to sense depleted Mg^{2+} levels prior to the point that global cellular processes are affected. Therefore, it would be much more appropriate to repeat this experiment with analysis at much shorter intervals of Mg^{2+} starvation. Despite the fact that this particular data set is largely uninterruptible, it was noted that some of the unchanged genes did include other riboswitch regulated genes such as *glmS* and *yitJ*.

DISCUSSION

In total, the findings from the in vivo experimentation on the M-box RNA element confirm that this regulatory RNA senses Mg^{2+}-depletion in order to elicit control of downstream gene expression. However, these data do not distinguish whether or not the M-box RNA is directly associating with Mg^{2+} in order to elicit control of gene expression or perhaps sensing the presence of an indirect metabolite whose levels fluctuate in response to intracellular Mg^{2+} concentration. In either case, this regulatory RNA is designed as an ‘off’ switch in which ligand binding stabilizes the formation of

the P1 anti-antiterminator helix, which prevents the occurrence of the antiterminator helix and allows the formation of the intrinsic terminator helix, thereby shutting off gene expression. In short, the ligand bound state is the 'off' state. These data indicate that this 'off' state occurs under high Mg^{2+} conditions. This observation fits well with the hypothesis that the RNA is directly sensing Mg^{2+} although it does not rule out the possibility that it is sensing another ligand that is only present when Mg^{2+} is abundant or perhaps sensing a small organic metabolite in complex with Mg^{2+}. Thus, at the end of the in vivo experimentation, the model for M-box regulatory function can be summarized in Figure 12.

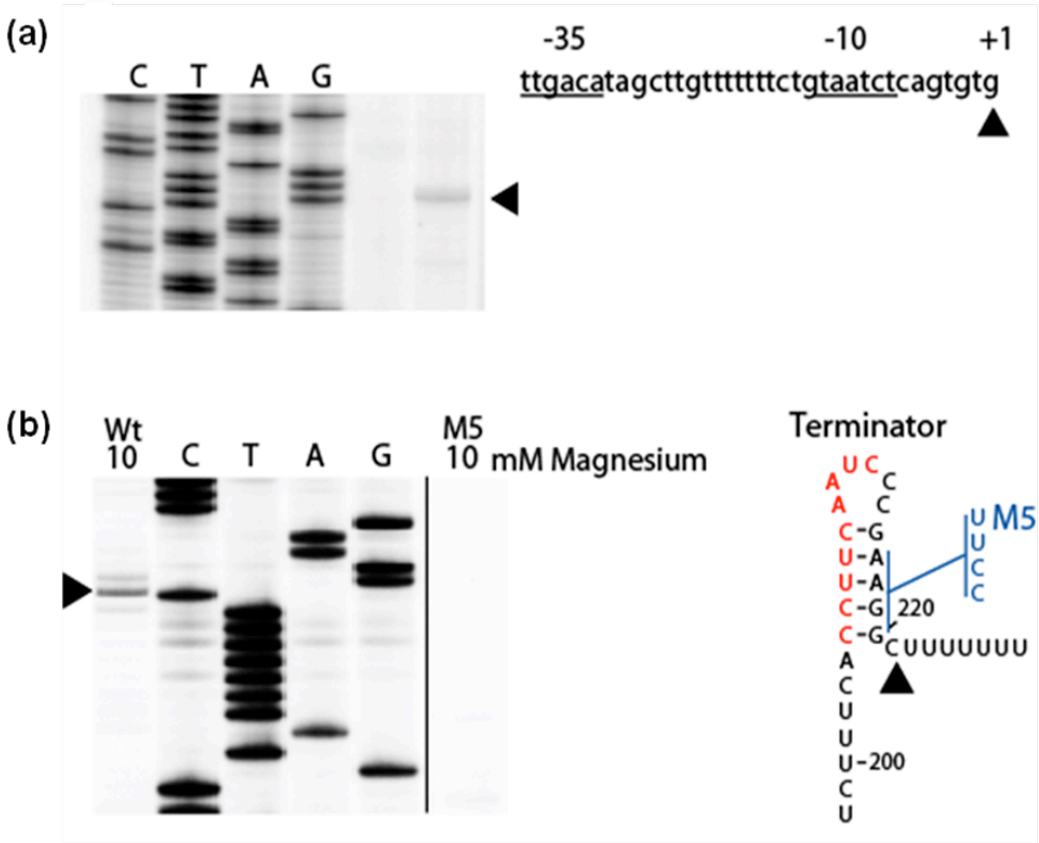


Figure 5: Mapping of the promoter and terminator elements of the *B. subtilis* M-box RNA **(a)** Primer extension indicated the likely 5' transcriptional start site of *B. subtilis mgtE* RNA. The promoter sequence is representative of a typical *B. subtilis* vegetative promoter. **(b)** The presence of the terminator helix can be observed in vitro when reverse transcriptase (RT) is halted by the secondary structure of the terminator. The RT stop observed for the wild-type construct can be compared to the M5 terminator mutant construct which should not display any secondary structure in this region. The terminator sequence is displayed on the right. Red nucleotides represent portions of the terminator that are involved in the antiterminator helix. Mutations introduced in the M5 construct are denoted in blue. Lanes labeled C, T, A, and G correspond to DNA sequencing ladders. Black triangles highlight the key RNA elements identified by these assays. Figure from Dann, Wakeman et al., 2007.

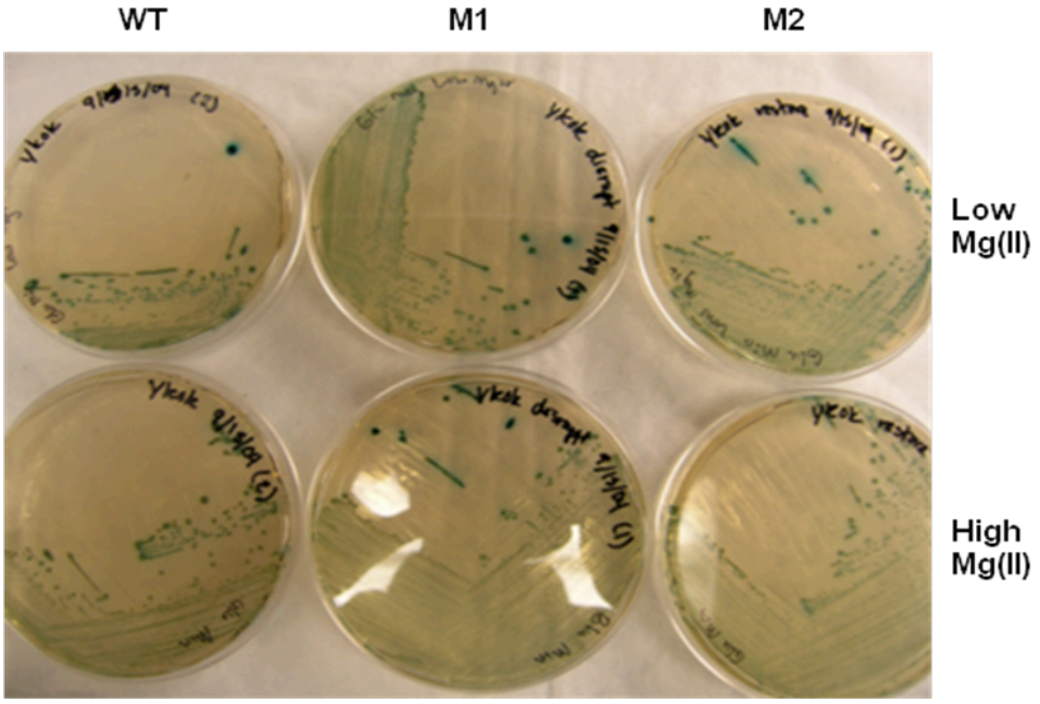


Figure 6: Preliminary analysis of M-box:*lacZ* fusion constructs. Focusing on the individual colonies in the final streak of these glucose minimal media plates, a trend can be observed. While all of the colonies appear blue, WT M-box fusions appear bluer on the low Mg^{2+} media as compared to the high Mg^{2+} media. The M1 disruption mutant exhibits high *lacZ* expression under both conditions, and the M2 restoration mutant appears to follow the same trend as the WT M-box:*lacZ* fusion.

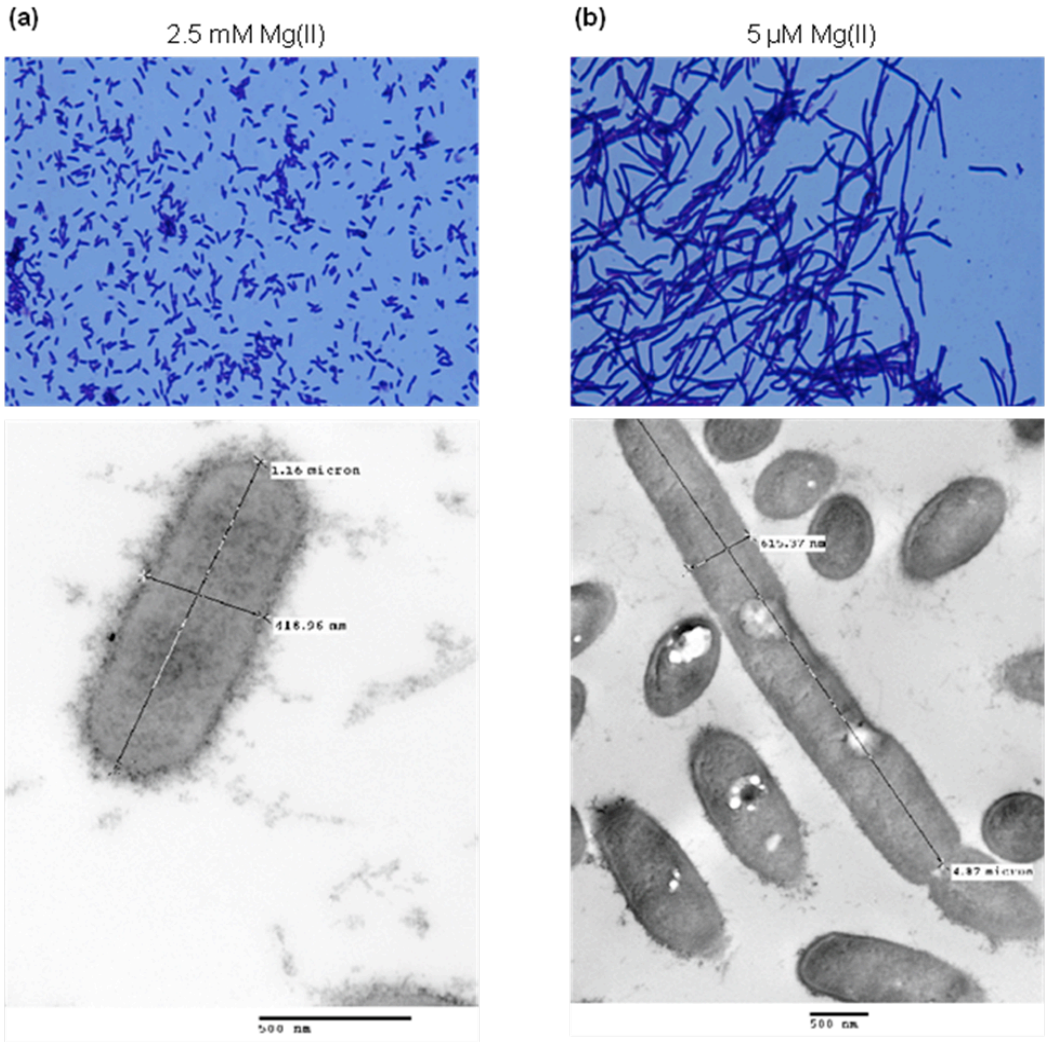
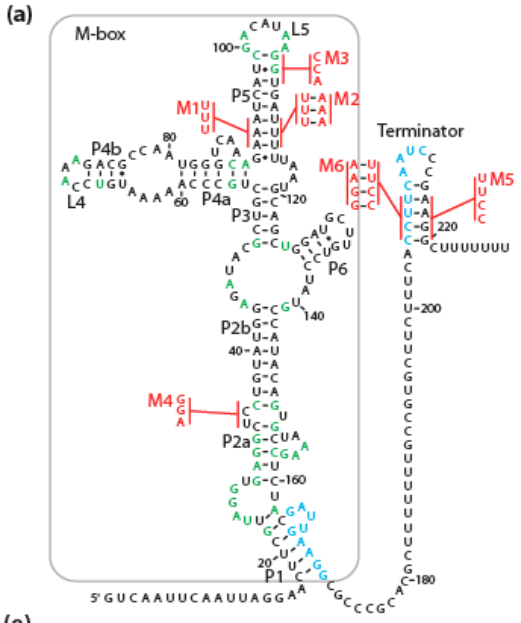
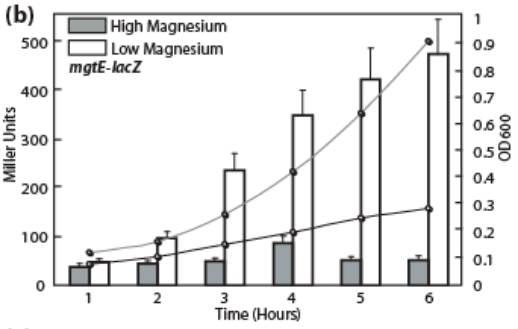
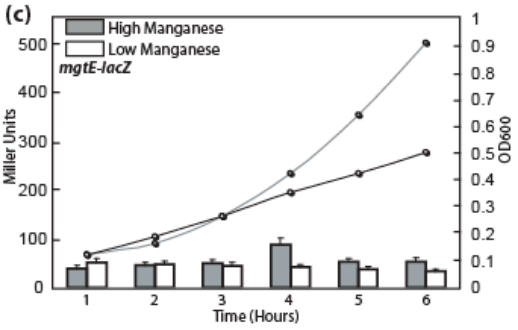
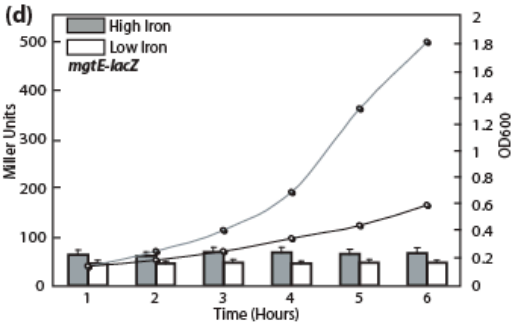


Figure 7: Light and electron microscopy reveal that the relatively short, Gram-positive rods of *B. subtilis* observed under normal growth conditions containing 2.5 mM Mg^{2+} **(a)** become significantly elongated upon growth in media containing 5 µM Mg^{2+} **(b)**. The bar at the bottom of the images indicates the scale in the electron micrographs.









(e)

lacZ Fusion Construct	Low Mg^{2+}	High Mg^{2+}	Fold Repression
Full-Length mgtE 5'UTR	360 +/-34	48 +/-8.9	7.5
mgtE Promoter-lacZ	180 +/-21	240 +/-9.5	0.75
mgtE Aptamer-lacZ	490 +/-82	310 +/-42	1.6
M1	430 +/-74	200 +/-38	2.2
M2	300 +/-35	82 +/-8.8	3.7
M3	350 +/-18	250 +/-7.7	1.4
M4	330 +/-29	170 +/-9.4	1.9
M5	340 +/-7.6	240 +/-8.7	1.4
M6	47 +/-0.65	14 +/-1.1	3.4
SAM Riboswitch-lacZ	34 +/-1.1	130 +/-9.5	0.26
GlcN6P Riboswitch-lacZ	10 +/-0.44	30 +/-2.9	0.33

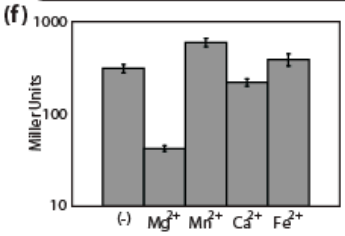


Figure 8: (**a**) Secondary structure of the 5′ portion of *B. subtilis mgtE* highlighting the locations of site-directed mutations denoted in red. Green nucleotides match the consensus pattern for positions that are conserved at 95% or greater. Cyan indicates the putative antiterminator pairing. The boxed region denotes the M-box, the aptamer domain, of the regulatory RNA. (**b–d**) The promoter and entire UTR of the *ykoK* (*mgtE*) gene were fused to *lacZ*. Line plots reflect OD_{600} measurements and correspond to the right-side y axis (shaded and open circles indicate conditions of 'high' and 'low' metals). Bar graphs reflect expression of the *lacZ* reporter fusion (Miller Units) (filled and open bars indicate conditions of ''high'' and ''low'' metals). (**e**) Variants of the *mgtE-lacZ* fusion were tested for expression with 2.5 mM or 0.005 mM Mg^{2+}. All experiments were repeated at least in triplicate. Construction of the *B. subtilis yitJ-lacZ* (SAM riboswitch) and *glmS-lacZ* (glucosamine 6-phosphate riboswitch) fusions are described elsewhere (Winkler et al., 2003, 2004). (**f**) *B. subtilis* strains containing the *mgtE-lacZ* fusion were grown in minimal media containing either no divalent ions or the indicated ion at 50 mM. Under these conditions, only the addition of Mg^{2+} caused specific reduction in *b*-galactosidase activity. Figure from Dann, Wakeman et al., 2007.

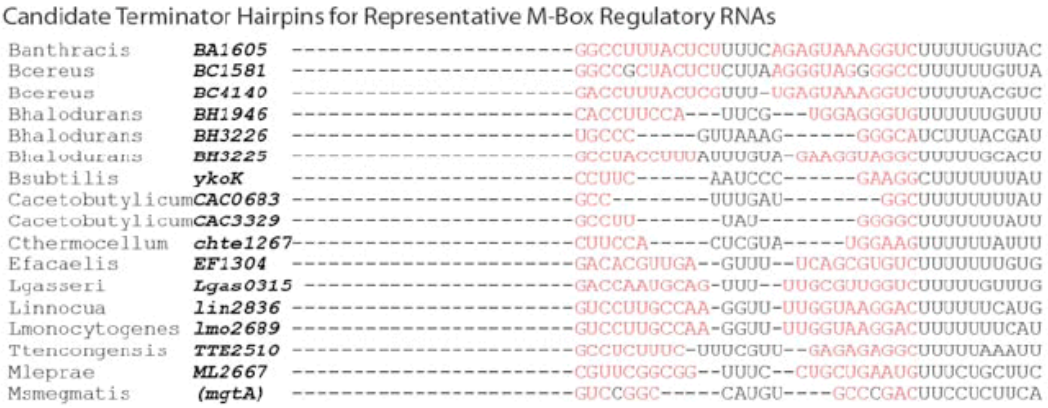


Figure 9: M-box terminator elements. Alignment of M-box expression platforms reveal that, while sequences of these regions vary, many of these RNAs contain G:C rich helices (shown in red) followed by a poly U tract. Figure from Dann, Wakeman et al., 2007.

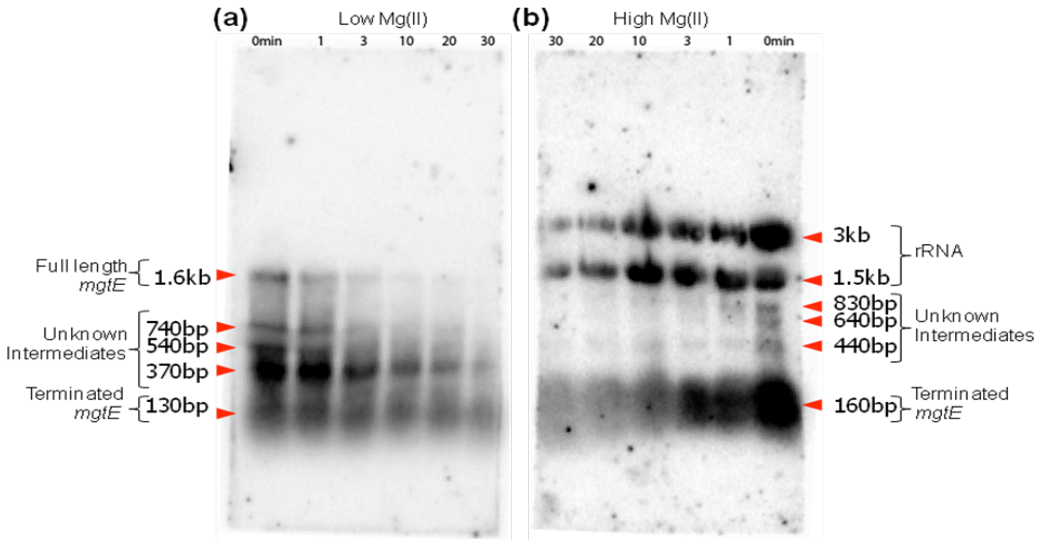


Figure 10: Northern blot analysis of total RNA from cells growth under 5 µM **(a)** and 2.5 mM **(b)** Mg^{2+} conditions. Since this Northern blot analysis was originally an attempt to determine the half-life of the M-box transcript, the numbers shown at the top indicate the amount of time elapsed subsequent to stopping total cellular transcription with rifampin at the time of total mRNA extraction. Under low Mg^{2+} conditions, both full length and terminated *mgtE* gene products are observed at similar levels. Under high Mg^{2+} conditions, a strong band representative of the terminated *mgtE* is present but, due to what appears to be non-specific probing of rRNA, we cannot confirm the presence or absence of the full length *mgtE* transcript. Under both conditions, a large number of unknown species can be observed. These unknown species do appear to become more abundant under low Mg^{2+} conditions.

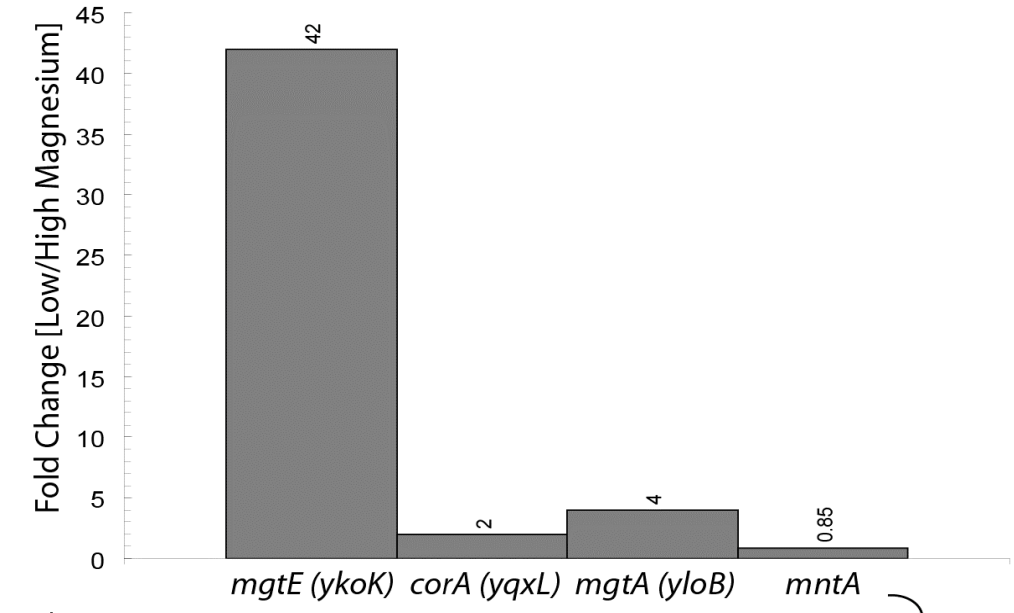


Figure 11: qPCR data reveal that the *mgtE* family Mg^{2+} transport system gene expression is more responsive to Mg^{2+} depletion than other families of metal ion transporters in *B. subtilis*. The graph indicates the fold change in gene expression in low versus high Mg^{2+} growth conditions. The numbers below the graph represents the cycle threshold (Ct) data normalized to the level of 5S rRNA gene expression, a highly expressed RNA product whose levels are unaffected by intracellular Mg^{2+} fluctuations.

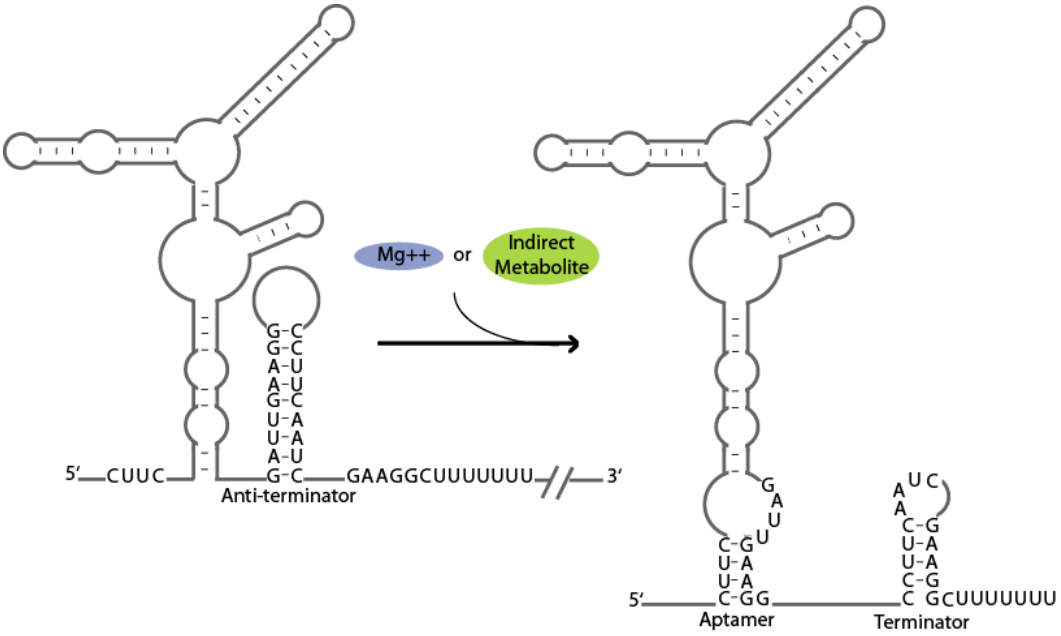


Figure 12: Model of M-box function based on in vivo studies. Through mutational analysis, we have demonstrated that sequences producing the alternate base pairing schemes for the terminator and anti-terminator elements are required for M-box-mediated regulation of gene expression. We have also demonstrated that the terminator base-paring scheme that downregulates gene expression is formed under high Mg^{2+} conditions. However, we do not know whether the M-box is directly or indirectly sensing intracellular Mg^{2+} concentrations.

CHAPTER FOUR

BIOCHEMICAL AND BIOPHYSICAL EVIDENCE FOR DIRECT METAL-SENSING BY THE M-BOX RNA

INTRODUCTION

While the in vivo gene expression data indicate that the M-box element senses intracellular Mg^{2+} levels, these data did not distinguish whether or not the M-box RNA was directly associating with Mg^{2+} in order to elicit control of gene expression or perhaps sensing the presence of an indirect metabolite whose levels fluctuate in response to intracellular Mg^{2+} concentration. All previously studied riboswitch RNAs were shown to sense small organic molecules, making the latter choice seem more likely. However, direct sensing of Mg^{2+} by riboswitch RNAs should be possible. Therefore, one of the aims of this project was to determine the effects of varied levels of Mg^{2+} on the RNA both structurally and functionally and demonstrate that these effects are unique to the M-box as compared to other well characterized examples of unrelated riboswitch RNAs.

RESULTS

As a preliminary measure to demonstrate that M-box RNAs are directly responsive to fluctuations in Mg^{2+} and not simply sensing a more traditional small organic ligand that changes in response to intracellular Mg^{2+} levels, attempts were made

in order to determine the effect of varying Mg^{2+} levels on the structure of these unique RNAs. A diverse assortment of biochemical and biophysical techniques are regularly used to investigate the structural arrangements of biological RNAs. Among these different experimental techniques are structural probing methods, which are often times employed to determine which nucleotides for a given RNA polymer are paired or unpaired and are therefore useful for the assessment of secondary structure. Additionally other gel-based structural probing methods assess whether certain RNA structures undergo dynamical structure changes more indicative of tertiary structure formation. RNA sequences fold upon themselves to form complex structures. Functional analysis of most biological RNAs requires knowledge of secondary structure arrangements and tertiary base interactions. Therefore, rapid and comprehensive methods for assessing RNA structure are highly desirable. Computational tools are oftentimes employed for prediction of secondary structure. However, a greater degree of accuracy is achieved when these methods are combined alongside structure probing experimentation. Multiple probing techniques have been developed to assist identification of base-paired regions. However, most of these techniques investigate only a subset of RNA nucleotides at a time. A combination of structure probing approaches is thus required for analysis of all nucleotides within a given RNA molecule. Therefore, methods that investigate local structure for all positions in a sequence-independent manner can be particularly useful in characterizing secondary structure and RNA conformational changes.

Structural probing reveals significant structural changes upon Mg^{2+} titration

One of the structural probing techniques I used to interrogate the *mgtE* M-box RNA was a recently developed technique known as selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) (Merino et al., 2005). This assay measures the relative rates of reactivity of N-methylisatoic anhydride (NMIA) to 2′-hydroxyls, a reaction influenced by internucleotide flexibility. In general, the internucleotide linkages of bases involved in the formation of helices and certain tertiary contacts will be inflexible while the internucleotide linkages of looped and bulged nucleotides will be largely susceptible to NMIA modification. SHAPE has proved to be a powerful technique over the past few years; however, when it was first employed in the Winkler lab, the SHAPE method had not been examined over a wide range of divalent concentrations. Therefore, *mgtE* RNAs were also analyzed by in-line probing, which measures the relative rate of spontaneous scission at internucleotide linkages. Previous experiments demonstrate that proper in-line configuration between the 2′-hydroxyl and the 5′-oxyanion leaving group is the predominant feature dictating these cleavage rates (Li and Breaker, 1999; Soukup and Breaker, 1999). However, it is likely that nucleophilicity of the 2′-hydroxyl is also influenced at high metal ion concentrations. It was therefore important to apply both techniques to verify that data obtained by either method were consistent and not the result of a heretofore unrecognized Mg^{2+}- specific artifact introduced by each respective technique. These reactions included the M-box domain alone (nucleotides 14-172, relative to the transcriptional start site) or ‘full-length’ RNAs,

which completely encompass the aptamer and terminator regions (nucleotides 1-265 and 1-220 for SHAPE and in-line probing, respectively).

Initial predictions of secondary structural features in the M-box RNA were based on comparative sequence analyses (Figure 4). Comparative sequence analysis is simply a method of predicting secondary structure in RNAs based upon a comparison of multiple known sequences of a certain class of RNA found within a variety of different organisms. If certain nucleotide stretches within these RNA sequences display base pairing potential in all known examples, they are likely to base pair. An even more convincing testament to potential base-pairing is nucleotides within these stretches that exhibit co-variation. Co-variation is when a nucleotide on one side of a potential helix appears to have changed identity over the course of evolution to restore base-pairing potential with another evolutionarily mutated base on the other side of the putative helix. Regions of NMIA reactivity (SHAPE) and spontaneous cleavage (in-line probing) largely agreed with the secondary structure as predicted by comparative sequence analyses, with bands primarily corresponding to nucleotides that were predicted to be unpaired. SHAPE probing reveals that the addition of Mg^{2+} leads to lowered reactivity, indicating greater structural constraint, for over 40 internucleotide linkages (Figure 13). Most of the nucleotide linkages that appeared to become more structured upon addition of Mg^{2+} localized to portions of RNA predicted to form loops or bulges rather than helices, suggesting that the Mg^{2+} was promoting the formation of tertiary contacts between unpaired bases. Each individual change in reactivity occurred at similar Mg^{2+} concentrations, suggesting that the change in RNA conformation was concerted with an EC_{50} of ~0.6 mM and ~2.7 mM Mg^{2+} for aptamer and full-length RNAs, respectively

(Figure 13b). Furthermore, Mg^{2+}-specific changes in the aptamer acted in concert with altered reactivity of terminator and antiterminator positions (Figure 13c-d). Importantly, these latter probing changes agreed well with the predicted switching between antiterminator and terminator base-pairing schemes (Figure 13 d). For example, nucleotides C21 and G217 appear to be unstructured in the low Mg^{2+} state but become structured upon addition of Mg^{2+}. The model predicts that these nucleotides are unstructured/flexible in the unbound form of the M-box RNA but will be sequestered within the anti-antiterminator (P1) and the terminator helices, respectively upon ligand binding; therefore, the SHAPE probing supports this model. Additionally, unlike most internucleotide linkages in the M-box RNA, positions 172 and 211-213 exhibited increased NMIA reactivity upon exposure to high Mg^{2+}, indicating that these nucleotides were becoming less structured with increased Mg^{2+}. This unique response to association with Mg^{2+} is consistent with their relocation from the antiterminator helix to the unstructured region linking the anti-antiterminator to the terminator helix and terminator terminal loop, respectively. The data therefore provide biochemical evidence that Mg^{2+}-induced aptamer changes are responsible for downstream terminator formation. To further bolster these findings, in-line probing assays were performed with both the full-length element (Figure 14) and aptamer domain (Figure 15) of the M-box RNA. These gels also revealed a concerted conformational change as indicated by several decreasing bands indicative of regions of RNA becoming more highly structured. These structural transitions were located in similar positions to those observed for SHAPE probing and occurred with approximately the same EC_{50}.

As previously discussed in Chapter 3, a number of point mutants presumably disrupting or restoring predicted secondary structural elements within the M-box RNA element were constructed to demonstrate the importance of both the aptamer and expression platform domains. Each of these mutations exhibited varying degrees of defects in Mg^{2+} responsiveness according to *lacZ* reporter gene fusion data. Therefore, the degree to which some of these mutants affected the apparent EC_{50} for the structural transition observed upon titrating in increasing Mg^{2+} levels was tested. In-line probing of the M1 mutant designed to disrupt a non-conserved portion of the P5 helix resulted in the apparent EC_{50} shifting to a higher Mg^{2+} concentration. When the same experiment was performed with the M2 mutant designed to restore the base pairing in the P5 helix while altering primary sequence, the observed apparent EC_{50} was similar to that of the WT RNA although still slightly higher. This incomplete restoration of function in the M2 mutant was also observed in the *lacZ* reporter fusion assays (Figure 8e). Probing of the M3 mutant, which exhibited the greatest loss of Mg^{2+}-induced transcription attenuation both in vivo and in vitro, revealed that structural features within the L5 terminal loop were acutely disrupted while the remaining structure appeared to be similar to WT RNA in low Mg^{2+} concentrations (Figures 16). However, despite the close overall structural resemblance to WT RNAs, M3 RNAs were incapable of the Mg^{2+}-induced conformational change, suggesting a direct role for L5 in the Mg^{2+}-folded state. The apparent structural similarity between the conformation adopted by M3 RNAs under high Mg^{2+} conditions and the low Mg^{2+} WT RNA conformation made this mutant an ideal control/marker representing the unfolded, non-Mg^{2+}-associated RNAs in subsequent experiments.

Mg^{2+} associates tightly to M-box RNAs in a unique manner as compared to other riboswitch RNAs

To investigate whether Mg^{2+} ions are specifically required for tertiary structure formation or whether the presence of high levels of cations in general will assist in the formation of M-box architecture, the structural probing tests were performed both in the presence of high (2.1 M) monovalent ions and low monovalent ions (0.1 M) (Figure 17). Under high monovalent conditions, these ions are expected to outcompete the loosely-associated divalent ion atmosphere, leaving only high-affinity divalent sites (Das et al., 2005a; Draper et al., 2005). The majority of the probing patterns under high Mg^{2+} conditions were similar between the two conditions (2.1 M and 0.1 M monovalent ions), indicating that divalent ions are specifically required for the formation of most tertiary contacts. Additionally, the EC_{50} values for these Mg^{2+}-induced conformational changes were similar between the two conditions. Therefore, although monovalent ions induce tertiary structure formation for certain RNAs (Takamoto et al., 2004) as well as certain positions within the M-box RNA, divalent ions are specifically required for the final folded M-box tertiary structure.

Due to the fact that all RNA molecules are structurally affected by Mg^{2+} levels to some extent, other RNA species needed to be tested in order to determine whether similar effects may be observed for riboswitches in general. Therefore several unrelated metabolite-sensing RNAs were also subjected to probing assays (Figure 18). In-line probing of an FMN-sensing RNA revealed Mg^{2+}-induced structural changes under the

low monovalent ion conditions. However, this effect was abolished upon addition of molar concentrations of monovalents. Further testing of SAM and TPP-binding riboswitches revealed that they too exhibited a lack of Mg^{2+}-responsive structural changes under high monovalent conditions. In another attempt to compare the M-box RNA to other unrelated but putatively structured RNAs, in-line probing was performed on an RNA element upstream of the *Salmonella enterica spiR* gene (Figure 19). This RNA is thought to function through the action of the PhoP/Q two component system to regulate the expression of virulence genes via the regulation of SpiR (Bijlsma and Groisman, 2005). Since PhoP/Q is a Mg^{2+}-responsive system and this RNA element is predicted by Mfold to have extensive secondary structure, it would be an interesting experiment to determine the structural effects of Mg^{2+} on the *spiR* RNA. Mg^{2+} titrations demonstrated that this RNA is structurally unresponsive to Mg^{2+} under high monovalent conditions. Additionally, this RNA did not appear to be nearly as well structured as the Mfold analyses (Zuker, 2003) would predict. In in-line probing assays, unstructured regions appear as bands on the gel. When this RNA is subjected to in-line probing, the vast majority of nucleotides appear as strong bands. These data both contrast and highlight the complex architecture of the *mgtE* M-box RNA and its specific requirement for Mg^{2+} in the folding process.

As a further demonstration of the special relationship between Mg^{2+} and the tertiary conformation of M-box RNAs, other examples of the M-box element from different organisms were analyzed to determine whether or not they would be structurally altered by changes in Mg^{2+} concentration. For this experiment, two additional M-box RNAs were selected: an M-box RNA from *Bacillus cereus*, a Gram-positive

microorganism fairly closely related to the *Bacillus subtilis* used in all prior assays, and an M-box from *Vibrio cholerae*, an unrelated Gram-negative microorganism. Once again in-line probing was used to determine the structural effects of titrated Mg^{2+} into samples containing these RNAs under high monovalent conditions (Figures 20). These gels cannot be directly compared band for band to the *B. subtilis* M-box gels since these RNAs are radioactively labeled at the 3′ end rather than at the 5′ end, but the overall structural transitions are similar. An interesting future direction would be to compare the EC_{50} for Mg^{2+}-induced compaction of many examples of this element from various unrelated bacterial species. The EC_{50} of these RNA should be set at a threshold around which the physiological levels of intracellular Mg^{2+} naturally fluctuate. If the EC_{50} of RNAs isolated from bacterial species living in different environmental conditions differs, this would indicate that environment factors can impact cytoplasmic contents. However, if the EC_{50} of all of these RNAs is the same, this would indicate that despite different environmental stresses experienced by these organisms, intracellular contents of Mg^{2+} and potentially other metals are evolutionarily fixed at a certain set point. A recent study suggests that the intracellular levels of metals or 'metallome' of most organisms is similar (Barton et al., 2007). However, this and other studies measure total intracellular metal content rather than free metal content. Measurement of total metal content will measure both free metal and sequestered, biologically unavailable metal. The M-box RNA, on the other hand, could be used to directly assess free metal levels.

Additionally, probing reactions on the WT M-box RNA of *B. subtilis* were conducted in the presence of cobalt hexammine, a mimic of fully hydrated Mg^{2+} that can substitute at Mg^{2+} sites when only outer-shell RNA-metal contacts are required (Cowan,

1993) (Figure 21). M-box RNAs were not conformationally modified by increasing cobalt hexammine and the resulting probing patterns closely resembled those obtained under low Mg^{2+} conditions, suggesting that inner-sphere Mg^{2+} coordination is required for tertiary structure formation.

In order to demonstrate that this RNA element would be able to function as a direct sensor of Mg^{2+} inside a cell where the majority of 'free' Mg^{2+} is predicted to be chelated by nucleotide pools, in-line probing titrations were also repeated with Mg^{2+} that had been incubated in the presence of excess nucleotides. The same structural transition was still observed in these assays at a slightly increased, although still physiological EC_{50} in the low mM range (Figure 22).

Mg^{2+}-induced structural rearrangements are associated with the formation of a compact tertiary architecture

Since it is important to assess the formation of RNA structure through multiple independent techniques, *mgtE* RNAs were also subjected to footprinting by hydroxyl radicals, which measures cleavage of the solvent accessible backbone in a manner independent of sequence or secondary structure (Latham and Cech, 1989). This test enriched the SHAPE and in-line probing data by directly assessing whether the Mg^{2+}-induced conformational change correlates with formation of a closely packed internal core, a feature of sophisticated RNA tertiary architecture. Footprinting of the aptamer domain in the presence of 20 μM or 20 mM Mg^{2+} revealed an array of Mg^{2+}-induced protections within internal and terminal loops (Figure 23), demonstrating that indeed a

solvent inaccessible core was formed in the presence of Mg^{2+}. From these probing data, a model emerges wherein the aptamer domain is dominated by secondary structure in the presence of low or no Mg^{2+} but is substantially rearranged upon Mg^{2+} association to include higher order tertiary structure.

To directly measure whether the RNA undergoes a significant change in overall shape in its ligand-bound form, wild-type and M3 aptamer RNAs were subjected to velocity sedimentation by analytical ultracentrifugation (AUC) (Figure 24). During conditions of low (30 μM) or high (10 mM) Mg^{2+}, single RNA species were observed that displayed a sedimentation coefficient of ~5.6 and ~7.0, respectively, revealing that association of Mg^{2+} was indeed accompanied by a significant decrease in hydrodynamic radius. Mg^{2+}-induced compaction was highly cooperative with a Hill coefficient of ~4.29 ± 0.88 and exhibited an EC_{50} of 0.16 mM (Figure 24b). This value is slightly lower than the one obtained by structural probing due to the fact that the RNAs used for AUC assays were initially purified (although this technique was later improved upon for subsequently performed experiments) by size exclusion chromatography without the presence of EDTA to chelate trace amounts of Mg^{2+}. In contrast to the WT RNA aptamer, the M3 mutant demonstrated a sedimentation coefficient of ~5.9 even in the presence of 10 mM Mg^{2+}, confirming that it was incapable of forming the compacted structure and was confined to a secondary structure-dominated state. Based on these data, it seemed reasonable that the extended and compacted conformations might be separated by size exclusion chromatography. Indeed, aptamer RNA equilibrated in 30 μM Mg^{2+} eluted at an earlier retention time relative to 10 mM Mg^{2+} (Figures 25a). However, M3 RNAs equilibrated in 10 mM Mg^{2+} exhibited a retention time nearly identical to wild-type

RNAs that had been equilibrated in 30 μM Mg^{2+}. This result supports the hypothesis that the M3 RNA is a reliable control for the extended conformational state. Therefore, this capability was exploited in order to compare the retention times for M3 and wild-type RNAs equilibrated in divalent metals other than Mg^{2+} (Figure 25b).

Other divalent ions induce M-box compaction in vitro

Previous studies have demonstrated that metalloregulatory proteins and transport proteins are capable of discriminating between chemically similar metal ions typically by evolving metal coordination sites that sense specific ions as dictated by coordination chemistry (Geidroc and Arunkumar, 2007). However, preliminary assays revealed that divalent ions other than Mg^{2+} can also elicit the compacted tertiary conformation (Figure 25b). In-line probing confirmed that at least one of these alternative metals, Ca^{2+}, induced a structural rearrangement identical to Mg^{2+} at the same approximate EC_{50} value (Figure 26a). In-line probing assays with Mn^{2+} titrations were also attempted. The Mn^{2+} appeared to elicit the same conformational change but the EC_{50} for this transition could not be determined since high levels of Mn^{2+} resulted in severe smearing of the RNA bands on the gel presumably due to the oxidized Mn^{2+} that formed after incubation at room temperature for ~40 hours. Additionally, SHAPE probing of M-box RNAs in the presence of 5 mM Ca^{2+} or Mn^{2+} revealed these alternative divalent ions could induce terminator helix formation in a manner similar to that of Mg^{2+} (Figure 26b-d). Together these data indicated that the M-box RNA structure is likely to be a general sensor for

divalent metals in vitro, although the compacted tertiary conformation is specifically tuned to in vivo Mg^{2+} levels.

More extensive titrations with these alternative divalents using AUC were performed in order to investigate whether the EC_{50} value is tuned to a concentration that is reached only by intracellular Mg^{2+} and not other divalent ions (Figure 27). Interestingly, while Mg^{2+} and Ca^{2+} were almost identical in their ability to induce M-box compaction, Mn^{2+} appeared to have approximately 4-5 fold greater affinity to this RNA element. Based upon this observation, one could imagine this RNA element may be used to selectively sense Mn^{2+} in some other organisms even though it has been demonstrated that in vivo this element is only sensitive to Mg^{2+} depletion in *B. subtilis*. We hypothesize that the reason the *B. subtilis* M-box is only responsive to Mg^{2+} in vivo despite the greater affinity between the RNA and Mn^{2+} is that the total metal ion concentrations of bacteria (at least those reported for *E. coli*) predict that Mg^{2+} is the only divalent ion present at a level sufficient to trigger the M-box riboswitch (Outten and O'Halloran, 2001). While the intracellular metal concentrations or 'metallome' of the vast majority of organisms remains unstudied, it is believed that these levels remain relatively constant across a wide range of species (Geidroc and Arunkumar, 2007). Therefore, the studies in *E. coli* are likely to reflect the intracellular metal levels found in *B. subtilis* and many other organisms. These estimates place total intracellular Mn^{2+} levels at ~1-10 μM. These intracellular ions are weakly chelated to thiolate ligands; therefore, the free metal concentration may be up to 0.1 μM, a concentration far too low to be sensed by the *B. subtilis* M-box RNA whose EC_{50} for Mn^{2+} appears to be ~50-100 μM (Figure 27).

However, in order for the cell to detect potentially toxic levels of Mn^{2+}, Mn^{2+}-sensing metalloregulatory proteins tend to operate at a set point in the μM range or higher (Geidroc and Arunkumar; 2007 and references therein). This range is slightly closer to the ~50-100 μM EC_{50} measured for the Mn^{2+}-induced compaction of the *B. subtilis* M-box. Additionally, some studies indicate that in certain species such as *Deinococcus radiodurans* and *Enterococcus faecium* the Mn^{2+}/Fe^{2+} ratio has evolved to be significantly higher in order to protect against radiation damage (Daly et al., 2004). Therefore, certain organisms might have evolved to contain intracellular Mn^{2+} levels high enough to be sensed by the M-box RNA.

In support of the idea of a Mn^{2+}-responsive M-box, M-box RNAs reside in the 5′ UTR of putative manganese transporters in organisms such as *Caldicullulisisruptor saccharolyticus, Carboxydothermus hydrogenoformans,* several *Clostridium* spp., *Geobacter uraniumreducens,* several *Mycobacterium* spp., *Pelobacter propinoicus*, and *Syntrophobacter fumaroxidans* (Table 1; Ramesh and Winkler, 2009). In order for the RNAs from these organisms to be specifically sensing intracellular Mn^{2+} levels rather than Mg^{2+}, one of two possible requirements must be met. Either the specificity of these RNAs must be switched so that they are no longer responsive to Mg^{2+}, or the intracellular concentration of Mn^{2+} is increased in these organisms to the point that Mn^{2+} rather than Mg^{2+} levels fluctuate around the set point of this RNA. Preliminary experiments were performed to test the theory that these RNA elements might be specifically sensing Mn^{2+}. The two M-box elements from *C. acetobutylicum* were chosen for these analyses since they both appear to be regulating Mn^{2+} transport systems. In order to determine whether or not the specificity of these RNAs had been switched from Mg^{2+} to Mn^{2+}, the ability of

these RNAs to fold into the compacted conformation in the presence of Mg^{2+} was assayed (Figure 28). Since Mg^{2+} alone can induce compaction of these RNAs, the only way for these M-box RNAs to be specific Mn^{2+} sensors is for the either intracellular levels of Mn^{2+} to be increased in these organisms relative to other bacteria.

In most organisms, Mg^{2+} is the most abundant divalent cation at ~100 mM total concentration with free ion concentration in the low mM range. The total concentration of divalent ions other than Mg^{2+} are no greater than 200 μM, with free (unbound) concentrations presumed to be substantially lower. Thus, while the *B. subtilis* M-box RNA is not specific for a particular divalent ion in vitro, Mg^{2+} is the only ion that the RNA is likely to respond to in the vast majority of organisms as indicated by the fact that most M-box-regulated genes are Mg^{2+} transporters. After all, others have observed that intrinsic metal ion affinity is a poor predictor of biological metal specificity since some metals are found at much higher free intracellular concentrations than others (Geidroc and Arunkumar, 2007).

Mg^{2+}-responsive transcription attenuation can be recapitulated in vitro

While the combination of all of the previous in vivo and in vitro data strongly suggested that the M-box RNA is a direct sensor of intracellular Mg^{2+} levels, these data would be benefited by the recapitulation of the proposed Mg^{2+}-responsive transcription attenuation mechanism in vitro. The recapitulation of transcription attenuation mechanisms in vitro has been accomplished for a variety of metabolite-sensing RNAs (e.g., Mironov et al., 2002; reviewed in Winkler and Breaker, 2005). These experiments

typically include the addition of only RNA polymerase, ribonucleotides, and DNA templates encompassing the regulatory RNA including the ligand-binding aptamer domain, the expression platform containing base-pairing potential for both terminator and anti-terminator helices, and some additional sequence at the 3′ terminus. A specific effect upon termination through addition of the putative ligand is taken as strong evidence that accessory protein factors and additional unidentified metabolites are not required for attenuation. Therefore the effect of Mg^{2+} on transcription termination within the *mgtE* 5′ UTR in vitro was measured. These reactions result in transcription run-off product (RO), representative of the full length gene product that would be observed in vivo, and transcripts truncated at the terminator helix (T) (Figure 29a).

Data collected at several Mg^{2+} concentrations reveal a modest ~15% increase in termination in response to Mg^{2+} (Figure 29a-b). Typically, intrinsic terminators consist of a G:C-enriched helix followed by a polyuridine tract (de Hoon et al., 2005; Gusarov and Nudler, 1999). Unique within the current members of its riboswitch class and rare compared to all intrinsic terminators, the *mgtE* regulatory RNA has an unpaired nucleotide between its terminator helix and polyuridine stretch (C221). Deletion of C221 improved the dynamic range for Mg^{2+}-induced termination in vitro to ~40% overall (Figure 29a,c). However, since the naturally occurring terminator appears to be imperfect, the question "Why would evolution have favored an imperfect terminator?" can be raised. Perhaps *B. subtilis* might prefer a high level of default *mgtE* expression; therefore, a leaky terminator could have been evolved for control of *ykoK* (*mgtE*) expression. Intriguingly, the effect of introducing this mutation in vivo exhibits unexpected results that possibly warrant further study. When the *lacZ* fusion construct

containing this mutation was created, I expected to observe increased Mg^{2+}-induced repression of reporter gene expression as compared to the WT construct. However, the opposite effect was observed. Under low Mg^{2+} conditions (the 'on' state), *lacZ* gene expression reached levels comparable to what was observed for the WT construct, but under Mg^{2+}-replete conditions the repression effect is only ~4-fold versus the ~8-fold repression observed for the WT construct (Figure 30b). Therefore, deletion of the putatively unpaired residue antagonized magnesium-responsive control of transcription termination. Additionally, when the in vitro transcription termination assays were repeated with *B. subtilis* RNA polymerase rather than the *E. coli* RNA polymerase no Mg^{2+}-induced termination could be observed for this construct (Figure 30a). Another example of a *B. subtilis* terminator element in which the G:C rich helix and the polyuridine tract are separated by an unpaired C can be found in the *queC* $preQ_1$-binding RNA (Roth et al., 2007) and a predicted terminator element in the *eps* operon (Irnov and Winkler, unpublished data). Perhaps future studies on these observations might lead to a different understanding of the mechanism of transcription attenuation in *B. subtilis* and possibly other Gram-positive microorganisms.

In order to demonstrate that the Mg^{2+}-responsive transcription attenuation observed for both the WT and ΔC221 constructs requires an intact aptamer domain, in vitro transcription termination assays were repeated using the M3 mutant construct. Transcription termination by the M3 mutant was unresponsive to Mg^{2+} levels (Figure 29a-b). As a control for aptamer-independent effects of Mg^{2+} on termination, the *mgtE* promoter was fused immediately upstream of the terminator helix. Increased Mg^{2+} had little effect on termination for this construct, demonstrating that the terminator alone is

not stabilized by Mg^{2+} (Figure 29a,c). To further validate Mg^{2+}-responsive transcription attenuation as the general genetic control mechanism for M-box RNAs, the region upstream of *B. cereus* bc4140 (*mgtA*) containing another putative M-box class RNA was assayed under these conditions (Figure 29a,d). This M-box RNA exhibited Mg^{2+}-induced termination with a dynamic range of ~70%. In contrast, varying Mg^{2+} had no effect on termination for an established SAM-responsive transcription attenuator (Figure 29a,d). Additionally, representative transcription termination experiments were performed in low (2.5 mM) versus high (5.0 mM) Mg^{2+} concentrations comparing both *E. coli* and *B. subtilis* RNA polymerase (Figure 29e). Despite the inability of *B. subtilis* RNA polymerase to function on the ΔC221 construct, this polymerase was capable of Mg^{2+}-induced transcription attenuation in conjunction with the WT *B. subtilis* and *B. cereus* M-box RNAs in a manner comparable to that of *E. coli* RNA polymerase. Therefore, M-box RNAs are direct divalent metal sensors that promote transcription attenuation in the absence of accessory proteins and additional small molecules.

Structural model of the Mg^{2+}-bound M-box RNA

Genetic, biochemical, and biophysical analyses of the *mgtE* UTR strongly supported a role for this RNA as a direct Mg^{2+} sensor, but these data would benefit greatly from a visualization of the mechanism of metal sensing at the molecular level. Due to ability of riboswitch RNAs to bind and sense the presence of intracellular metabolites in order to elicit control of gene expression, these RNAs possess exquisite structural sophistication. Therefore, a number of these RNAs such as the adenine-,

guanine-, SAM-, TPP-, lysine-, preQ$_1$-, and FMN-binding riboswitches have been subjected to structural analyses (Serganov et al., 2004; Montange and Batey, 2006; Serganov et al., 2006; Garst et al., 2008; Spitale, et al., 2009; Serganov et al., 2009). In most of these instances, crystallography and NMR of riboswitch RNAs has been utilized to elucidate the mechanisms by which these RNAs recognize and bind to their cognate ligand (Figure 31). To this end, collaboration with a crystallographer, Dr. Charles E. Dann III, was initiated in order to obtain the 2.7 Å crystal structure of the Mg^{2+} associated M-box RNA (Figure 32; Dann et al., 2007). Previously determined three dimensional structures of other RNAs offered a preview of the architectural features that might be expected for M-box RNAs (Batey, 2006). Multiple RNAs have structures that employ specific metal binding sites, hinting at mechanisms that could be used for metal sensing. For example, specific Mg^{2+} sites have been identified for TPP-binding RNAs, group I introns, and the ribosome, wherein the metal ions participate in ligand stabilization, chemical catalysis, and structural stabilization, respectively (DeRose, 2003; Hougland et al., 2005; Klein et al., 2004; Selmer et al., 2006; Serganov et al., 2006; Vicens and Cech, 2006). These different structures coordinate Mg^{2+} through at least one inner-sphere contact, most often via the nonbridging oxygen of a phosphate group, although they also interact with nucleobase functional groups and ribose oxygens. The fact that partially dehydrated Mg^{2+} appeared to be required by M-box RNAs due to the fact that cobalt hexamine cannot induce similar conformational changes (Figure 21) suggested M-box Mg^{2+}-binding pockets will share similar features with these other sites.

RNA crystals were grown in the presence of ~5 mM Mg^{2+}, a concentration that resulted in a structure in the ligand-bound state (Figure 32). The RNA adopts a complex

three-dimensional architecture, mostly comprised of three closely packed, nearly parallel helices (Figure 32b-e). The secondary structure of this crystallographic model largely agrees with that based on covariation and conservation; however, the crystal structure allowed us to reassess the depiction of the P2 region which in actuality appears to be dominated by non-Watson:Crick base-pairing interactions that could not be predicted by comparative sequence alignments (Figure 32a). In total, the structure agrees well with the chemical probing data, including the positioning of hydroxyl radical protections within a closely packed internal core (Figure 32e). Additionally, the majority of positions that demonstrated decreased internucleotide flexibility in response to Mg^{2+} (Figure 13) are located at a region where the three parallel helices converge via a network of long range contacts (Figures 32b-e). A most exciting aspect of the structural model involves the presence of six Mg^{2+} that also reside predominantly in this region of tertiary contacts (Figures 32b-c and 33). Extensive inner- and outer-sphere contacts to Mg^{2+} occur for nucleotides in the P2, L4, and L5 regions of the RNA (Figures 33 and 34). Mg^{2+} mediate multiple long-range interactions, leading to stabilization of the compact three-helical tertiary structure. Together, the chemical probing data, measurements of changes in hydrodynamic radius, and the tertiary structure combine to suggest a model wherein the extended P4 and P5 helices clamp against P2 to form the compacted three parallel helix structure in response to binding of Mg^{2+}.

In addition to the Mg^{2+}, four K^{+} are also modeled in the structure (Figure 35a). To assess the role of K^{+} in the tertiary structure of the RNA, SHAPE probing assays were repeated in the presence or absence of Mg^{2+} and a range of monovalent concentrations (Figure 35b). While in the absence of Mg^{2+} the overall NMIA reactivity of all positions

appears to be affected by the titration of K^+, no structural rearrangements appear to occur in increasing K^+ concentrations. These tests revealed that the Mg^{2+}-bound conformation could be induced even in the absence of monovalent ions. Furthermore, sedimentation velocity measurements of Mg^{2+}-associated RNAs were identical in the absence or presence of K^+ (Figure 35c). In contrast, the presence of high monovalent ions in the absence of Mg^{2+} did not result in formation of the compact tertiary conformation (Figure 17). Therefore, tertiary folding of M-box RNAs is strictly dependent upon divalent ions, despite the presence of four highly coordinated K^+ in the structural model.

Of the six Mg^{2+}, Mg1 exhibits the most inner-shell RNA contacts, consisting of three contacts to L5 nonbridging phosphate oxygens (G100, C102, and A103) and a contact to U104 O4 (Figure 34a). These interactions appear to mediate a sharp bend in the L5 loop. The same Mg^{2+} mediates multiple outer-shell RNA contacts to L5 positions through two associated waters. Although there is precedence for RNA-bound Mg^{2+} with four inner-sphere RNA contacts (Type IV) they are rare in published structures (Cate et al., 1997; Klein et al., 2004). The low incidence of type IV Mg^{2+} in RNA structures and the fact that Mg1 organizes the L5 structure for docking with P2 and L4 indicate that Mg1 likely constitutes a key component in establishing the tertiary architecture.

Mg2 also appears to be important for the tertiary conformation as it coordinates nonbridging phosphate oxygens of G100 (L5) and U23 (P2), acting to bring the parallel P5-P2 helices together (Figures 34b).

A third Mg^{2+}, Mg3, helps position the P2 noncanonical A25-C160 pair and the highly conserved U24 position, which flips out to stack together with A155, a base that is itself flipped out from the P2 helix as a result of a near-classical UAA/GAN motif

(Figures 34c) (Lee et al., 2006). As a result of the Mg3-stabilized local structure, both U24 and A155 can form long-range base interactions to L5 positions. Many long range pairing interactions within the RNA structure involve residues within L5 and P2; therefore, given the intimate involvement of Mg^{2+} (Mg1-3) with these regions, they are likely to comprise the key metals for the tertiary conformation and metal sensing (Figure 33).

Mg4 also appeared to be important in the formation of the compact state observed in the crystal structure as it contacts the L4 tetraloop structure, which also mediates several long-range base interactions to P2 and L5 (Figure 34d). Interestingly, despite its sequence, CAAA, the L4 tetraloop superimposes well with a common RNA tertiary interaction motif, the GNRA tetraloop (Figure 36).

The two remaining Mg^{2+}, Mg5 and Mg6, associate to the modestly conserved P4 internal loop (Figure 34e-f). Due to the peripheral location of these ions and the fact that they do not appear to mediate any long range interactions outside of the P4 internal loop, these ions were initially believed to be relatively unimportant. However, potential roles for these Mg^{2+} are discussed in Chapter 5.

Multiple long range base interactions connect L4, L5, P2, and J2/1, regions in proximity to both Mg^{2+} Mg1-4 and antiterminator nucleotides (Figures 33 and 37). A common method for formation of long range contacts is through the use of A-minor motifs, wherein an adenosine fits into the minor groove of a Watson:Crick base-pair (Nissen et al., 2001). There are four such motifs in the M-box model (Figures 32a and 33a) between distant residues in the primary structure that have remained responsibly well conserved throughout evolution. These motifs likely play a significant role in

reducing the energetic cost of bringing the parallel helices in close proximity by constructing a network of helix-helix interactions. For example, A88-G151-C33 affixes the J4/5 interhelical region near P2 while A117-G83-C57 appears to assist in coaxial organization of P3 and P4. The A71-G22-C163 motif links L4 to P2, which induces a near-continuous array of base stacking between P1, P2, L4, and L5 nucleotides (Figure 37c). Interestingly, this base stacking involves positions within the antiterminator sequence, suggesting that stacking energies are involved in the preferential sequestration of these nucleotides into P1 in response to Mg^{2+}, rather than the antiterminator helix. Finally, the A155- G107-C99 motif anchors P5 to P2 via the UAA/GAN motif and L5 residues (Lee et al., 2006).

Mutational analyses verify essential tertiary contacts predicted by the structural model

In an attempt to determine the individual importance of these A-minor motifs, constructs were created in which either A71, A117, or A155 were mutated to cytidines. The preliminary analyses of the effects of these point mutations were performed in vitro using AUC (Figure 38). While none of these RNAs exhibit structural transitions to the compacted state in exactly the same manner as the WT M-box, some of the mutations are notably more severe than others. A71C, a disruption of the A-minor motif formed by the L4 tetraloop, is very similar to WT RNA. A155C, a disruption of the most highly conserved A-minor motif, is unsurprisingly the most severe mutant.

Although A-minor motifs are some of the most commonly used tertiary contacts in complex RNA structures and do appear to be important for the formation of the M-box compact state, additional equally important-looking long-range base interactions occur within the M-box. These interactions occur in the interhelical junctions that mediate sharp directional changes between adjacent helices or at the base of the molecule where antiterminator nucleotides, P2, L5, and L4 converge (Figure 37). For example, noncanonical pairings are present at the apex of the structure in a region between P6 and P2, facilitating the sharp angle between P2 and P3 helices. Near Mg3 in P2, the conserved U24 nucleobase forms a base-triple with G100 and A106 of L5. Additionally, an antiterminator nucleotide, U167, pairs to the Hoogsteen edge of A101 as well as the A103 Watson:Crick edge. Both of these interactions fasten L5 to the lower region of P2. An additional base-pair between A72 and A105 adds a direct contact between L4 and L5 and completes the network of P2-L5-L4 connections.

In addition to the point mutations that were designed to eliminate A-minor motifs, constructs assessing the individual importance of some of these other long-range tertiary contacts were created. In an attempt to assess the importance of Mg1, U104 was mutated to an A to eliminate the inner sphere contact between this nucleobase and Mg1, in theory eliminating the entire Mg1 binding site. A mutation to convert U24 to an A in order to disrupt the long range interaction with A106 was also created. Preliminary studies using AUC showed that, while these mutants affected the ability for the RNA to fold into the Mg^{2+}-induced compacted state, they did not affect the RNA as severely as the well characterized M3 mutant (Table 3). However, in vivo reporter gene fusion data

demonstrated that the minor effect in vitro was enough the virtually eliminate Mg^{2+}-induced repression of gene expression.

After the crystallographic model of the M-box RNA revealed the structure of the L4 tetraloop, mutants in L4 tetraloop were also constructed. As previously stated, this tetraloop appears to be unique in that the structure overlays perfectly with that of a classic GNRA tetraloop despite the CNRA sequence (Figure 36). The Mg^{2+} within this loop appears to be located in a position that could not be occupied in a traditional GNRA tetraloop due to steric hindrance by the bulkier purine base of at the G position. This ion may actually be mediating the formation of the GNRA-like tertiary architecture in the CAAA tetraloop. Therefore, it seemed possible that this new CAAA tetraloop constituted a Mg^{2+}-dependent GNRA tetraloop. The formation of this tetraloop structure is likely to be functionally important due to the fact that the bend in the backbone positions nucleotides at ideal angles to form tertiary contacts with both the P2 helix and nucleotides within the L5 terminal loop, stabilizing the base of the M-box three-helix bundle in a manner similar to that of Mg1-3.

Interestingly, only approximately 1/4 to 1/3 of the M-box sequences have a CNRA sequence while the remainder has the traditional GNRA sequence. If the theory was correct, the M-box RNAs with a GNRA sequence would have one less Mg^{2+} binding site than the RNAs with the CNRA sequence. Therefore, when Mg^{2+} titrations are performed and the concentrations required for structural transitions are measured, it might be evident that these two types of M-box RNAs have a differently shaped curve and possibly a different EC_{50}. These titrations were performed comparing WT *B. subtilis* RNAs (CAAA) to a GAAA mutant and a WT *B. cereus* RNA with a GAGA classical

'GNRA' tetraloop sequence (Figure 39). Interestingly, the sedimentation coefficients of the WT RNAs from both *B. subtilis* and *B. cereus* overlaid perfectly. However, the GAAA mutant did not appear to function as efficiently as these other RNAs. Since the comparison of the *B. subtilis* and *B. cereus* indicates that either a G or a C at that position can function, there might be other compensatory differences between these RNAs to allow the substitution of a G for a C in most of the M-box sequences but prevent the substitution in the *B. subtilis* RNA.

DISCUSSION

In total, these data indicate that the function of the M-box metalloregulatory RNA is to correlate intracellular Mg^{2+} concentration with formation of terminator or antiterminator helices. The convergence of the long-range base interactions and A-minor motifs to the region of the structure encompassing Mg1-4 and antiterminator nucleotides is likely to preferentially stabilize P1 rather than the mutually exclusive antiterminator helix. Specifically, in addition to Watson:Crick base pairing, certain antiterminator positions are sequestered into tertiary structural features in the Mg^{2+}-bound state, including base triples and base stacking within the tripartite structure formed by the Mg^{2+}-mediated convergence of L4, L5, and P2. These observations reveal a simple, yet elegant mechanism for metal-mediated occlusion of antiterminator nucleotides. Consistent with this model, many of the most highly conserved positions in the M-box RNA are the nucleotides that interact with metals and assist in coordinating the L4-L5-P2 tripartite structure. Therefore, the structural features exhibited by the *B. subtilis mgtE*

aptamer domain are expected to be general features for the other members of the M-box riboswitch class.

The M-box RNA structure is unique in many ways but is built upon principles observed previously for other RNAs. First, the regulatory RNA is remarkably similar in character to riboswitch RNAs, complete with a large evolutionarily conserved aptamer domain that controls formation of mutually exclusive terminator and antiterminator helices. Given the ease to which divalent metals interact with RNA polymers, it is reasonable to ask why such a large, highly conserved RNA structure would be required for metal sensing. The results of a battery of biochemical and biophysical tests suggest that a potential explanation emanates from the cooperative binding of metal ligands. Where other riboswitch aptamer domains have evolved to bind a single ligand, the M-box RNA associates with multiple metal ion ligands and tertiary formation is highly cooperative, perhaps demanding the increased information content. The only other example of cooperative binding of riboswitch ligands was observed for glycine-sensing RNAs, which utilize a tandem arrangement of glycine-binding aptamer domains (Mandal, et al., 2004). This feature of these RNAs may also allow sensing of small changes in ligand concentration, thereby offering genetic control over a more narrow range than for noncooperative regulatory RNAs. Another key principle for RNA structures that the M-box employs is the use of Mg^{2+} to transition into a fully folded tertiary conformation (Draper et al., 2005; Sigel and Pyle, 2007; Woodson, 2005).

Two features of M-box RNAs render them as metal sensors rather than simply another structured RNA region. First, metal ions promote a compacted RNA tertiary structure for the explicit purpose of governing the accessibility of a short nucleotide tract.

The availability of this signaling sequence can then be directly correlated to formation of terminator or antiterminator helices. Second, the apparent K_D for the metal-induced tertiary conformation is tuned to an appropriate intracellular concentration, thereby imparting sensory function to the overall structure.

The last two chapters of this work have described an RNA-based divalent cation sensor using a variety of in vivo and in vitro techniques. As such, a comparison to the previously characterized *S. enterica mgtA* RNA element, a proposed cation sensor, could highlight general rules for RNA-based metal sensors (Cromie et al., 2006). Unfortunately, as this RNA and M-box RNAs share no similarity at the primary or secondary structure level, a direct structural comparison is not particularly useful. The commonality between these two RNA classes reside in the fact that both appear to utilize a transcription attenuation-like genetic control mechanism through predicted Mg^{2+}-regulated switching between mutually exclusive helices. However, while the M-box RNAs described herein contain identifiable intrinsic terminator sequences, the novel mechanism by which helix switching in the *S. enterica mgtA* RNA influences transcription is not fully resolved. In addition to the transcription attenuation mechanism utilized by the *mgtA* RNA, both the *S. enterica mgtA* and *mgtCB* UTRs appear to be conferring Mg^{2+}-responsive gene expression by targeting the mRNA for degradation by RNase E (Spinelli, et al., 2008).

In a preliminary effort to compare the M-box and *mgtA* RNAs, the same Mg^{2+} titration in-line probing assays used for the M-box RNA were performed on the *S. enterica* RNA in an attempt to determine the structural affects of Mg^{2+} on this unrelated class of metalloregulatory RNA. When these assays were performed in the presence of

high monovalent ions, no significant structural transitions were observed (Figure 40). This result highlights the already apparent differences in the mechanisms of these two unrelated Mg^{2+}-sensing RNAs.

Regardless of the significant differences between these elements, the presence of multiple distinct classes of metal-sensing RNAs underscores the essential nature of metal ion homeostasis at all possible levels of regulation (i.e. RNA and protein) and raises the possibility of undiscovered RNA-based metal sensors. The data presented herein provide a structural and biochemical framework for comparative analyses of such metalloregulatory RNAs.

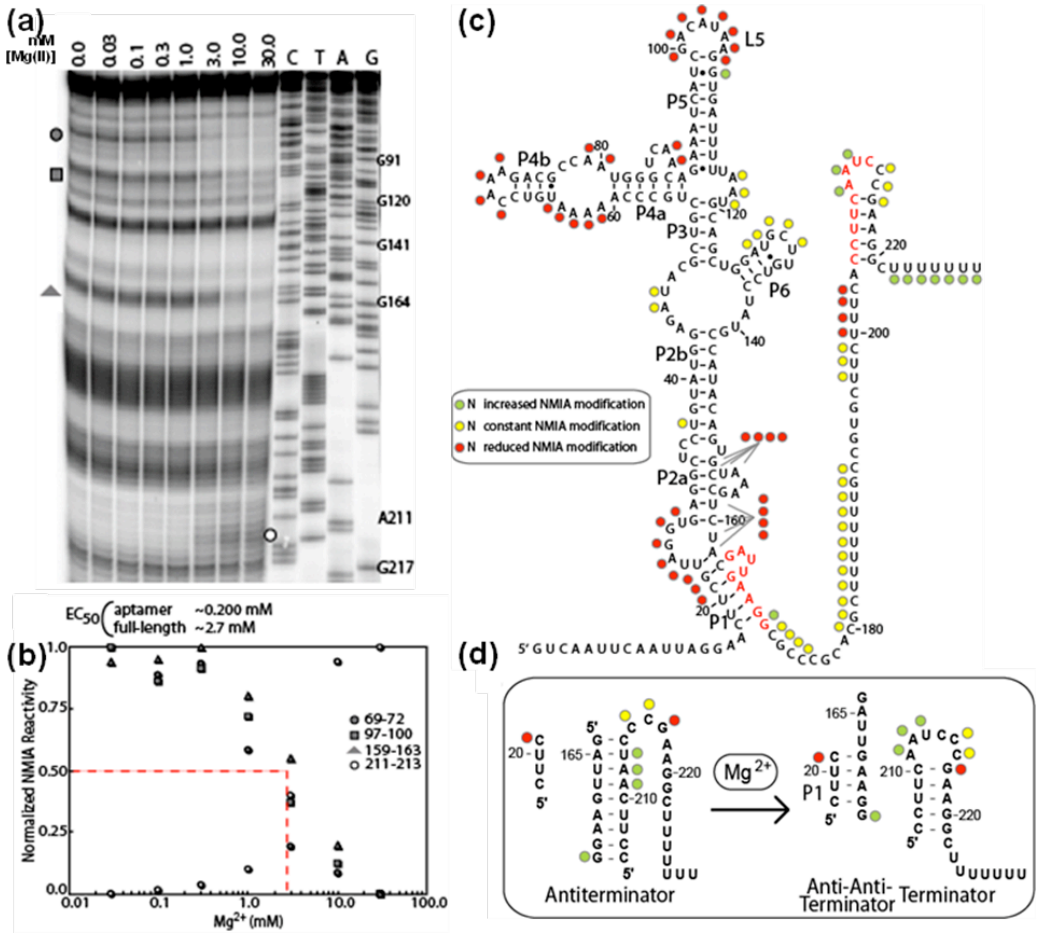


Figure 13: Mg^{2+}-induced structural transitions revealed by SHAPE probing, **(a)** Representative SHAPE probing of the *mgtE* regulatory RNA incubated with increasing Mg^{2+}, adjacent to DNA sequencing ladders. Reactions contained 2.1 M monovalent ions to outcompete loosely associated divalent ions. Filled shapes and the open square in panel indicate positions that decrease or increase in NMIA reactivity, respectively. **(b)** Filled shapes and the open square in panel **(a)** indicate positions that decrease or increase in NMIA reactivity. The decreasing reactivity was normalized and graphed. The increase in reactivity for positions 211–213 was normalized and graphed alongside this composite curve. These data indicate an EC_{50} value of ~2.7 mM for the construct that included the entire regulatory region including the terminator. **(c)** Summary of SHAPE data overlayed with the *mgtE* RNA sequence. Red, green, and yellow circles positioned adjacent to nucleotides denote internucleotide linkages that decrease, increase, or remain unchanged in their overall reactivity to NMIA in response to increased Mg^{2+}, respectively. **(d)** Schematic highlighting the agreement between SHAPE probing data and the helical regions postulated to be involved in genetic control (antiterminator, terminator, P1).

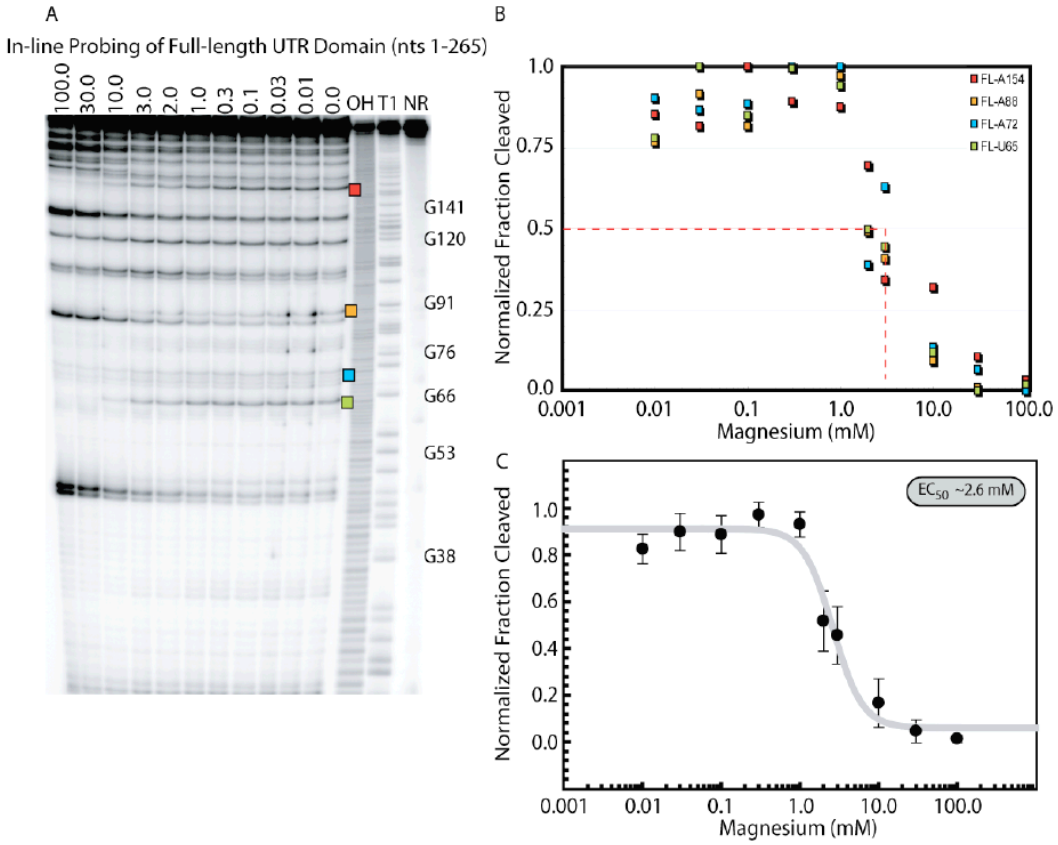


Figure 14: In-line probing analysis of full-length M-box regulatory element. **(a)** In-line probing revealed a concerted structural change of the *B. subtilis mgtE* full-length regulatory RNA that occurred in response to Mg^{2+} titration. The colored circles mark the bands that were quantified in the graphs. The numbers above the gel image display mM Mg^{2+} concentrations used in each reaction. 2 M potassium was added to each reaction to outcompete loosely associated Mg^{2+}. NR = non-reacted RNA. T1 = RNA digested with RNase T1, which cleaves after G. OH = RNA scission after every nucleotide. This probing data is summarized on the secondary structure model shown in Figure 3. **(b)** The normalized fraction of spontaneous backbone scission was graphed for positions denoted in the gel shown in **(a)**. All of the lines show a similar trend, with a similar half-maximal value, indicating that a concerted conformational rearrangement had occurred. The concentration of Mg2+ required for half-maximal scission (EC_{50}) is denoted by the red dotted line or revealed by curve-fitting analysis of the probing data, as shown in **(c)**. Figure taken from Dann, Wakeman et al., 2007.

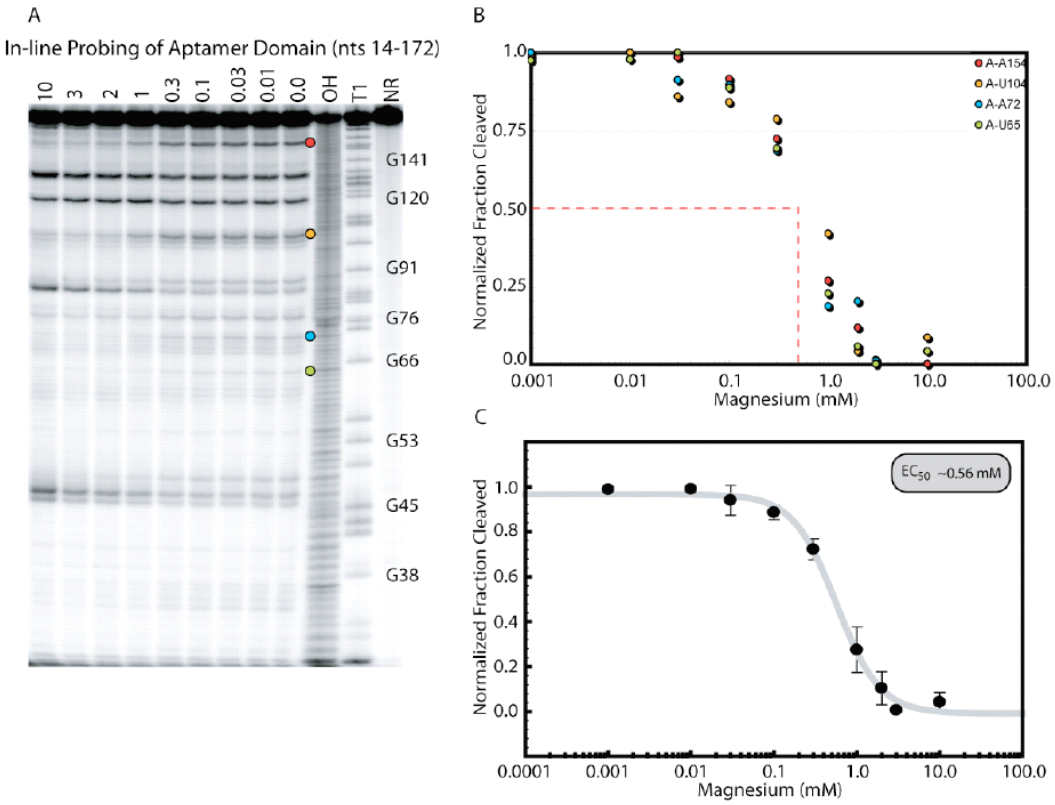


Figure 15. In-line probing analysis of the M-box aptamer domain. **(a)** In-line probing revealed a concerted structural change of the *B. subtilis mgtE* aptamer domain that occurred in response to Mg^{2+} titration. The colored circles mark the bands that were quantified in the graphs. The numbers above the gel image display Mg^{2+} concentrations used in each reaction in mM. 2 M potassium was added to each reaction to outcompete loosely associated Mg^{2+}. NR = non-reacted RNA. T1 = RNA digested with RNase T1, which cleaves after G. OH = RNA scission after every nucleotide. This probing data is summarized on the secondary structural model shown in Figure 3. **(b)** The normalized fraction of spontaneous backbone scission was graphed for positions denoted in the gel shown in **(a)**. All of the lines show a similar trend, with a similar half-maximal value, indicating a concerted conformational rearrangement had occurred. The EC_{50} is denoted by the red dotted line or revealed by curve-fitting analysis of the probing data, as shown in **(c)**. Figure from Dann, Wakeman et al., 2007.

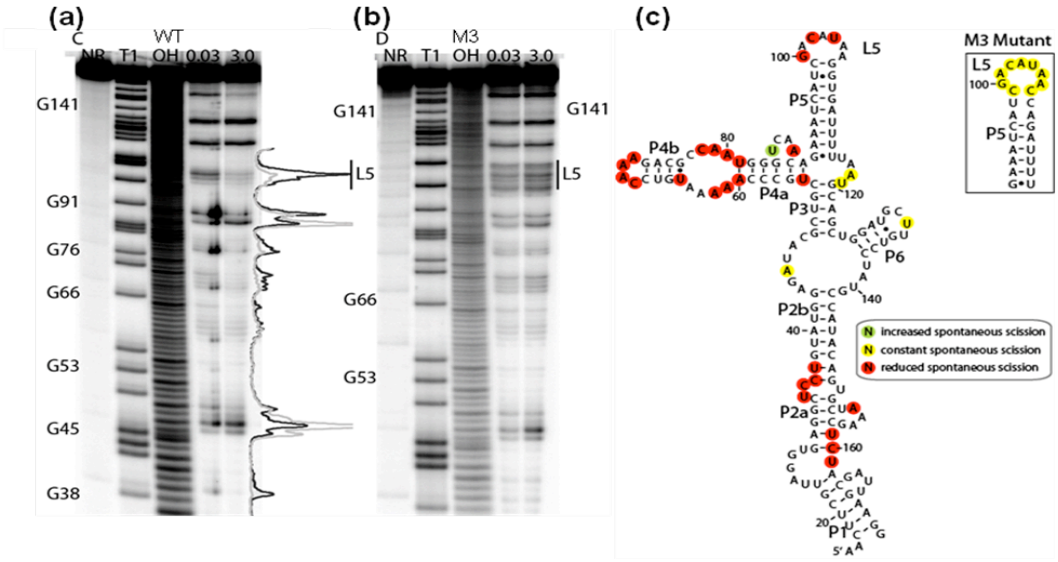


Figure 16: In-line probing of M3 versus WT aptamer domain constructs. **(a)** Structural probing of the wild type *B. subtilis mgtE* RNA aptamer at Mg2+ concentrations above and below the EC50 revealed a significant conformational change that occurred upon Mg2+ association. The structural rearrangements for wild-type *B. subtilis mgtE* RNA aptamer are highlighted by the line traces on the left. A significant number of bands decrease in intensity as the Mg^{2+} concentration is increased from 30 μM (black trace) to 3 mM (gray trace). (NR, nonreacted RNA; OH, hydroxyl-mediated cleavage at all positions; T1, cleavage at G). Reactions contained 2.1 M monovalent ions. **(b)** However, structural probing of the M3 mutant RNA at Mg^{2+} concentrations above and below the EC_{50} revealed no significant conformational change. The banding patterns that resulted from low and high Mg^{2+} conditions for M3 RNAs appeared strikingly similar to the low Mg^{2+} conformation of wild type RNAs, suggesting that M3 is simply incapable of the Mg^{2+}-induced tertiary conformation. However, significant probing differences were observed in the region of the L5 loop, which appears to be disrupted and disordered in M3. Therefore the M3 mutations disrupt L5 structural features and disallow the Mg^{2+}-induced tertiary conformation. These observations suggest that M3 RNAs can be considered reasonable RNA controls for the low-Mg^{2+} conformation exhibited by wild-type RNAs. **(c)** This probing data is summarized on the secondary structure model of the M-box RNA. Red, green, and yellow circles that encircle nucleotides denote internucleotide linkages that exhibit decreased, increased, or unchanged spontaneous cleavage in response to increased Mg^{2+}, respectively. The probing pattern for M3 RNAs was identical to wild-type in low Mg^{2+} except for disruption of L5, as shown in the inset.

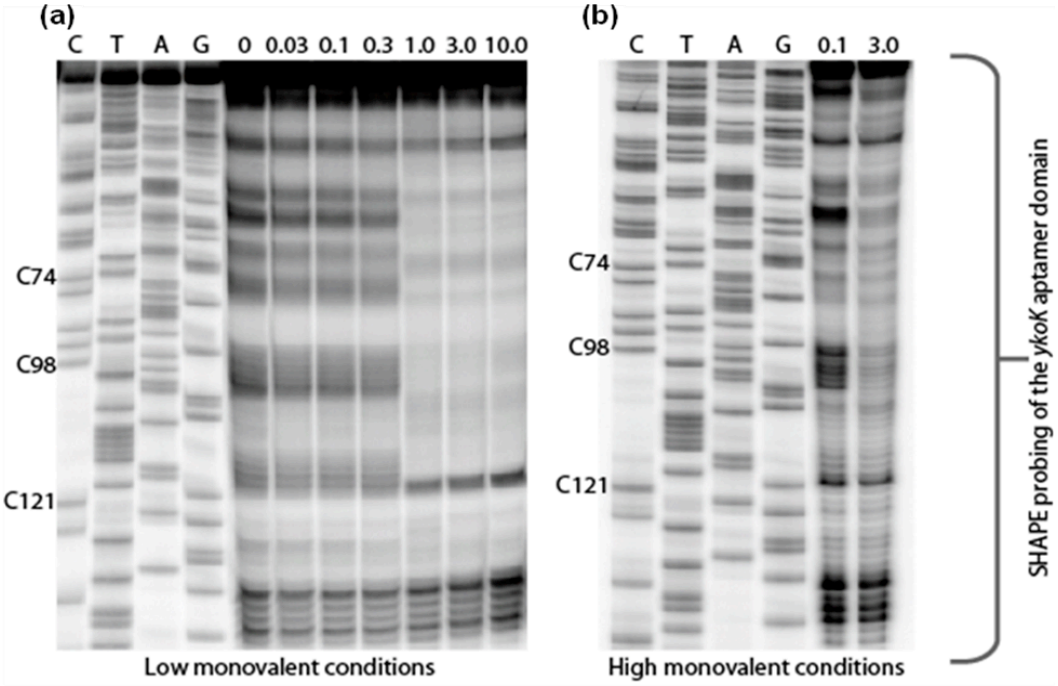


Figure 17: SHAPE probing revealed a concerted structural change of the *B. subtilis mgtE* RNA element that occurred in response to Mg^{2+} under both **(a)** low monovalent conditions (100 mM KCl) and **(b)** high monovalent conditions (2 M NaCl, 100 mM KCl). The numbers above the gel image display Mg^{2+} concentrations used in each reaction (mM). These data indicate that divalent ions are specifically required for a significant, specific conformational change in the *B. subtilis mgtE* RNA.

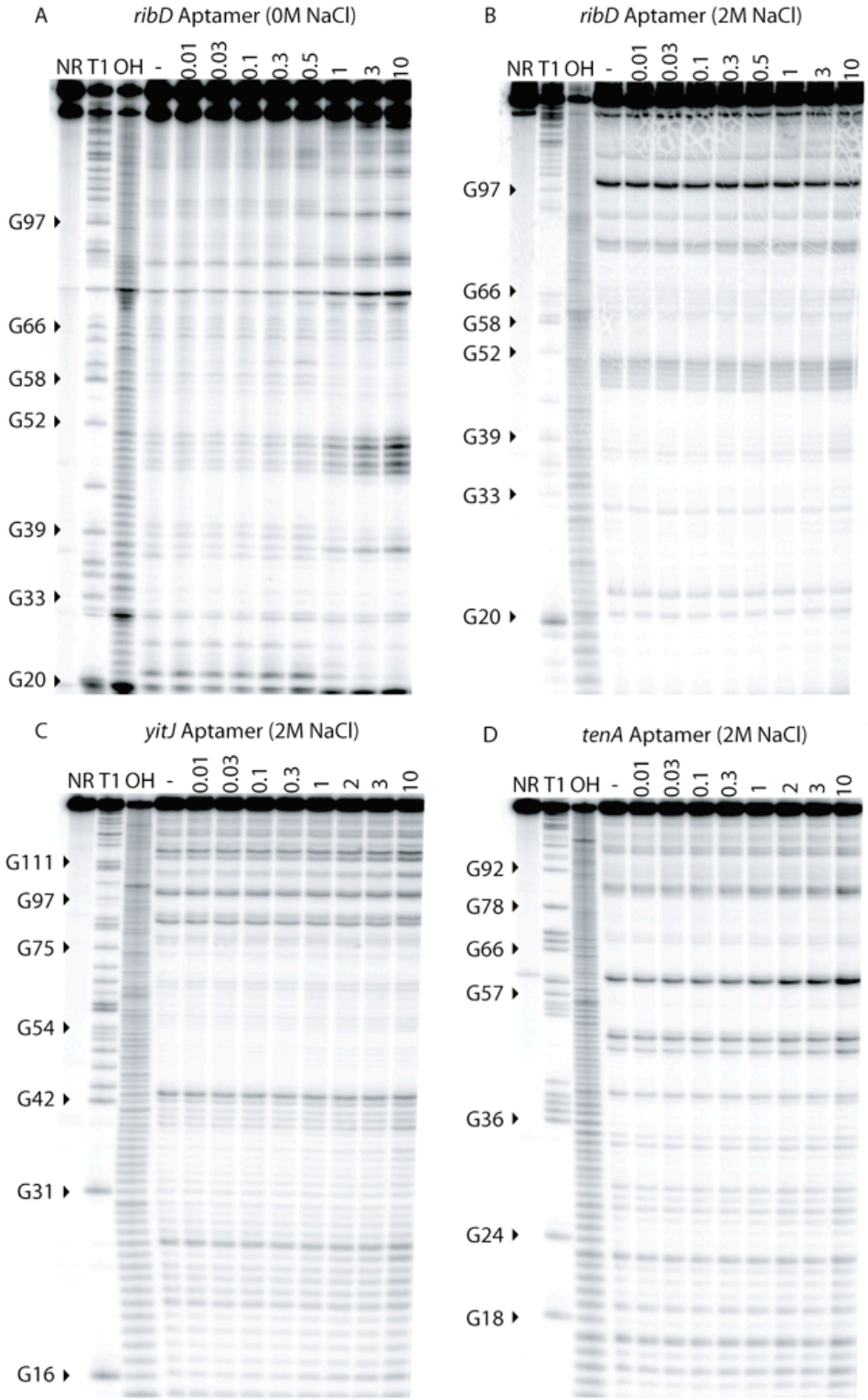

A
ribD Aptamer (0M NaCl)
NR T1 OH - 0.01 0.03 0.1 0.3 0.5 1 3 10
G97
G66
G58
G52
G39
G33
G20
B
ribD Aptamer (2M NaCl)
NR T1 OH - 0.01 0.03 0.1 0.3 0.5 1 3 10
G97
G66
G58
G52
G39
G33
G20
C
yitJ Aptamer (2M NaCl)
NR T1 OH - 0.01 0.03 0.1 0.3 1 2 3 10
G111
G97
G75
G54
G42
G31
G16
D
tenA Aptamer (2M NaCl)
NR T1 OH - 0.01 0.03 0.1 0.3 1 2 3 10
G92
G78
G66
G57
G36
G24
G18

Figure 18: In-line probing with unrelated riboswitch RNAs. **(a)** In-line probing of *B. subtilis ribD*, a flavin mononucleotide-sensing RNA, revealed that significant changes in the probing pattern occurred between 500 μM and 1 mM Mg2+ in the presence of 0.1 M monovalent ions. This observation suggests that Mg^{2+} alters the overall conformational state when the RNA is exposed to moderate monovalent ion levels. The RNA construct in this experiment was identical to '165-*ribD*', as described in Winkler et al., 2002. **(b)** However, further probing revealed that the Mg^{2+}-induced conformational change is ablated by the presence of 2.1 M monovalent ions. Interestingly, the probing pattern under these conditions was strikingly similar to the low monovalent, high Mg^{2+} condition, indicating that divalent ions are not specifically required for the conformational change that occurred with lowered monovalent ions. **(c)** In-line probing of *B. subtilis yitJ* (an *S*-adenosylmethionine-sensing RNA) under high monovalent conditions revealed no conformational change in response to increased Mg^{2+}. The RNA construct in this experiment was identical to '124-*yitJ*', as described in Winkler et al., 2003. **(d)** In-line probing of *B. subtilis tenA* (a thiamine pyrophosphate-sensing RNA) under high monovalent conditions revealed no conformational change in response to increased Mg2+. The numbers above the gel images display mM Mg2+ concentrations used in each reaction. The RNA construct in this experiment constituted nucleotides +1 to 124 relative to the transcriptional start site as indicated by Mironov et al., 2002. NR, T1, and OH identify no reaction, partial digestion with nuclease T1, and partial digestion with alkali, respectively.

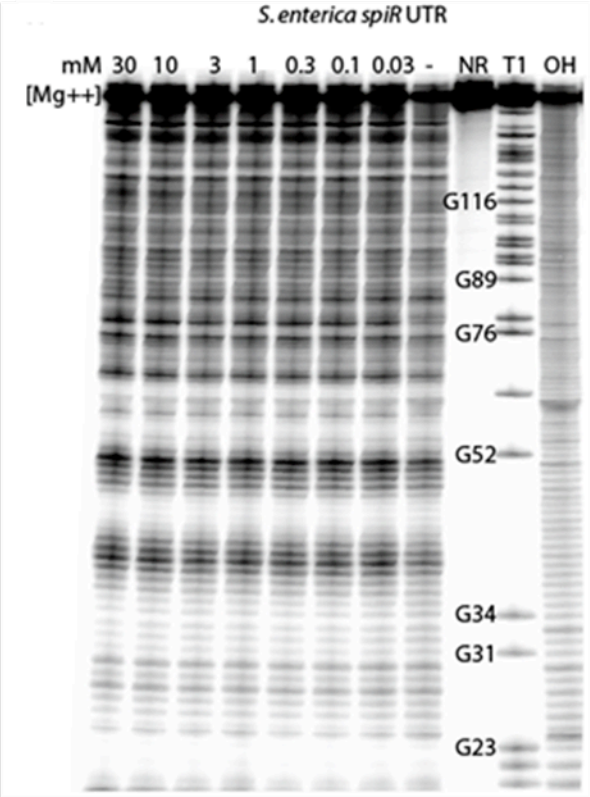


Figure 19: In-line probing with *S. enteric spiR* UTR. Under 2.1 M monovalent conditions, in-line probing reveals that there are no significant structural transitions that occur within this RNA at Mg^{2+} concentrations ranging from 0 to 30 mM. Additionally the large number of bands versus clear zones indicates that this UTR is largely unstructured. NR, T1, and OH identify no reaction, partial digestion with nuclease T1, and partial digestion with alkali, respectively.

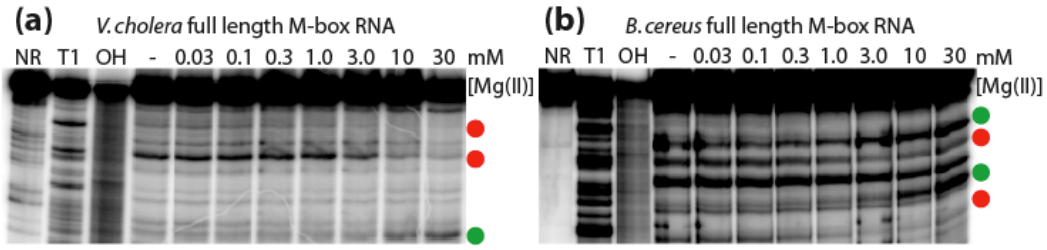


Figure 20: In-line probing of additional M-box RNA elements from *V. cholera* **(a)** and *B. cereus* **(b)** under 2.1 M monovalent conditions. Structural changes are marked by red spheres (increasing bands/regions becoming more structured) and green spheres (decreasing bands/regions becoming more flexible). These changes occur within the aptamer domain but are difficult to map due to their location at the top of the gel. NR, T1, and OH identify no reaction, partial digestion with nuclease T1, and partial digestion with alkali, respectively.

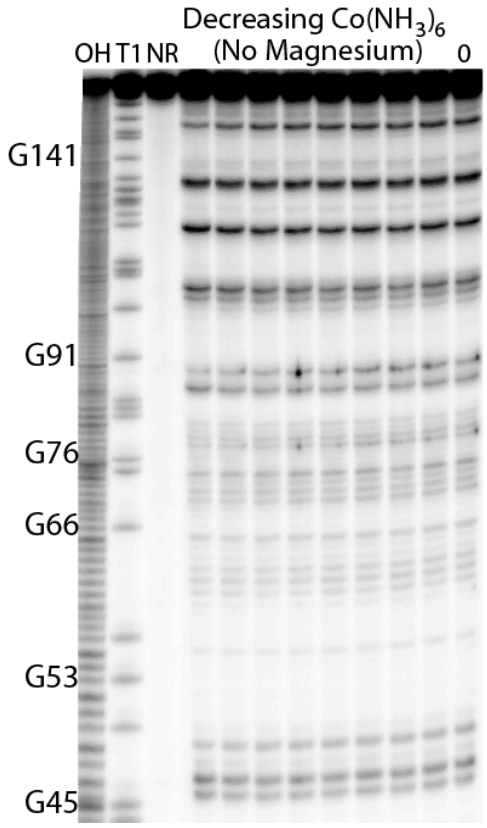


Figure 21. In-line probing with cobalt hexammine under high monovalent conditions. No significant structural transitions are observed as cobalt hexammine concentrations are increased from 0 to 100 mM. NR, T1, and OH identify no reaction, partial digestion with nuclease T1, and partial digestion with alkali, respectively.

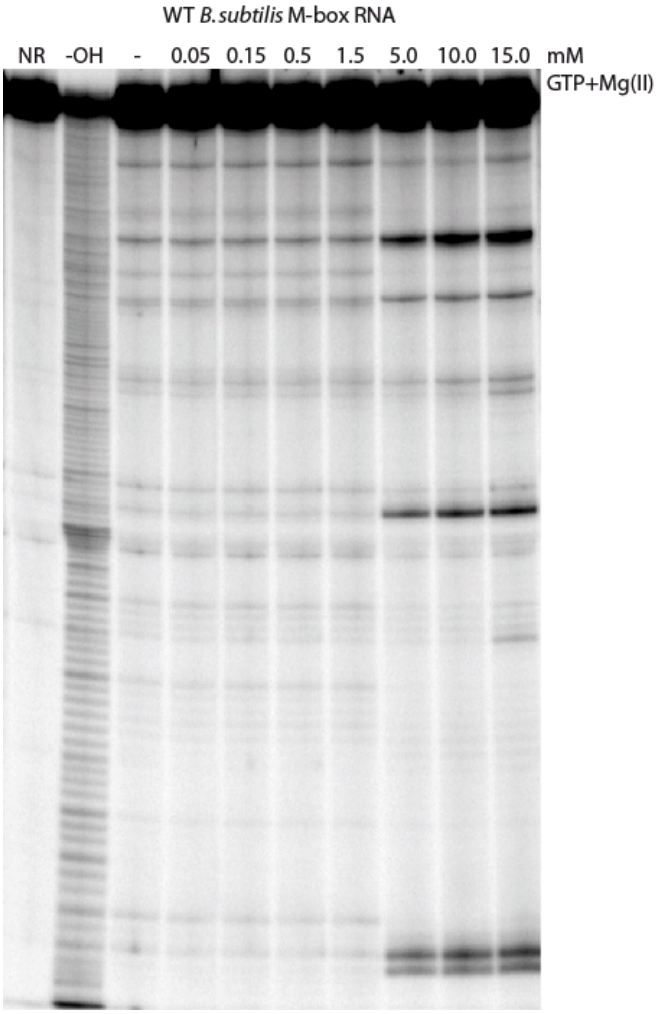


Figure 22: In-line probing with GTP-chelated Mg^{2+}. Concerted structural changes occur for M-box RNAs in the presence of chelated Mg^{2+}. These changes occur at a slightly higher though physiological EC_{50}. NR and OH identify no reaction and partial digestion with alkali, respectively.

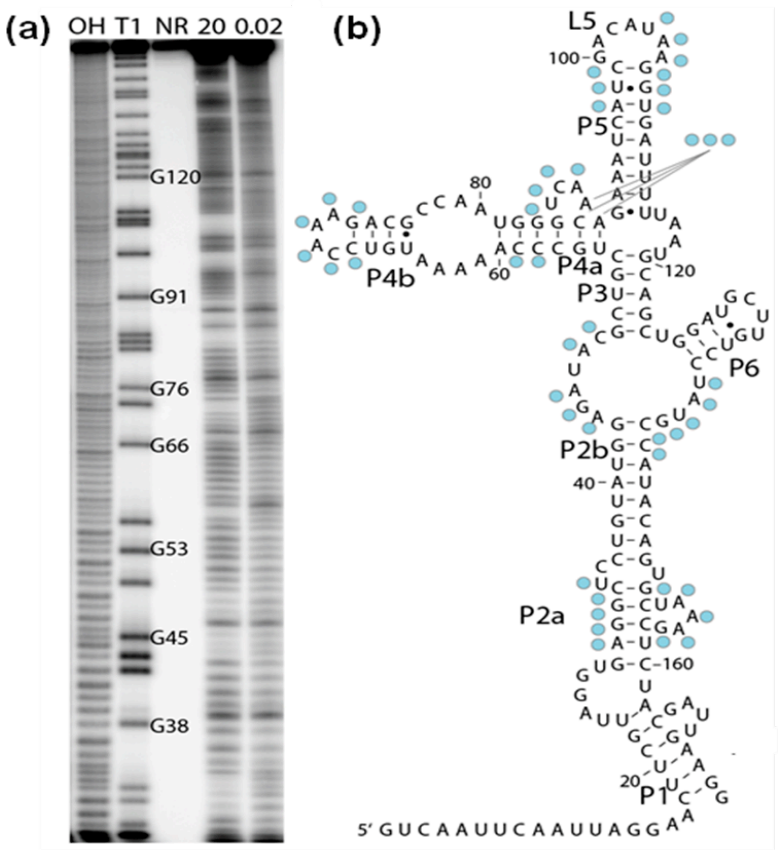
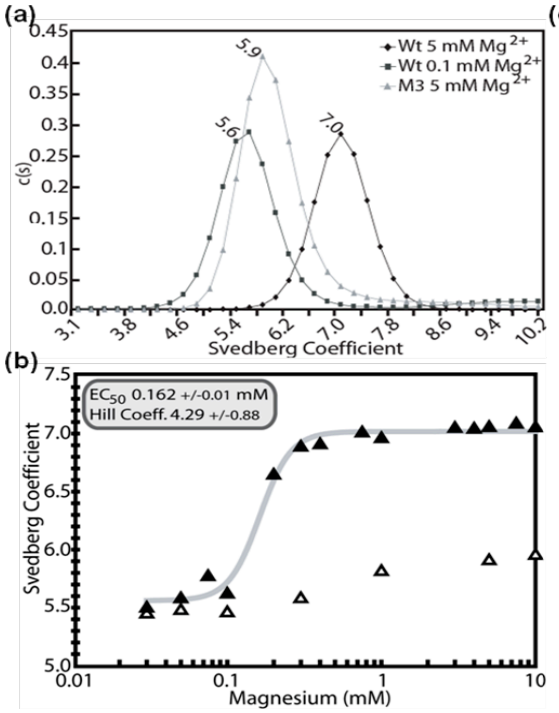


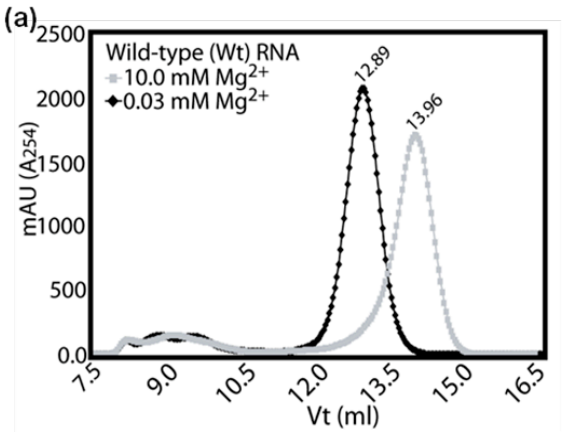
Figure 23: Hydroxyl radical footprinting reveals a Mg^{2+}-induced solvent inaccessible core formed in M-box aptamer domain. **(a)** Hydroxyl radical footprinting of the aptamer domain in the presence of 0.02 mM or 20.0 mM Mg^{2+}. NR, T1, and OH identify no reaction, partial digestion with nuclease T1, and partial digestion with alkali, respectively. **(b)** Positions of Mg^{2+}-induced protection against hydroxyl radical cleavage have been mapped onto the secondary structure of the M-box and are denoted by blue circles.



(c)

mM [Mg^{2+}]	$S_{20,w}$	$f(f_0)$	R_H	a/b	D
10	7.0	1.5	3.1	2.5	$6.9e^{-07}$
7.5	7.1	1.4	3.1	2.5	$7.0e^{-07}$
5	7.1	1.4	3.1	2.5	$7.0e^{-07}$
4	7.0	1.4	3.1	2.6	$7.0e^{-07}$
3	7.0	1.4	3.1	2.5	$7.0e^{-07}$
1	7.0	1.5	3.1	2.8	$6.9e^{-07}$
0.75	7.0	1.5	3.1	2.6	$6.9e^{-07}$
0.4	6.9	1.4	3.1	2.8	$6.8e^{-07}$
0.3	6.9	1.5	3.1	2.9	$6.8e^{-07}$
0.2	6.6	1.5	3.3	3.6	$6.6e^{-07}$
0.1	5.6	1.7	3.9	7.2	$5.6e^{-07}$
0.075	5.8	1.9	3.8	6.5	$5.7e^{-07}$
0.05	5.6	1.9	3.9	7.4	$5.5e^{-07}$
0.03	5.5	2.0	3.9	7.8	$5.4e^{-07}$

Figure 24: Velocity sedimentation (analytical ultracentrifugation) analysis of M-box RNA. **(a)** Wild-type and M3 RNAs were subjected to analytical ultracentrifugation (AUC) in the presence of high and low Mg^{2+} (wild-type in 5 mM Mg^{2+} = filled diamonds; wild-type in 0.1 mM Mg^{2+} = filled squares; M3 in 5 mM Mg^{2+} = filled circles). Wild-type aptamer RNA exhibited a change in sedimentation coefficient when equilibrated in high Mg^{2+} relative to low Mg^{2+}. However, M3 RNAs demonstrated a sedimentation coefficient most similar to low Mg^{2+} conditions, despite the presence of increased Mg^{2+}. **(b)** AUC of wild-type aptamer RNAs (filled triangles) revealed a highly cooperative change in sedimentation velocity in response to Mg^{2+}. In contrast, M3 RNAs (open triangles) underwent a considerably smaller change in hydrodynamic radius under identical conditions. **(c)** Additional parameters determined for the AUC experiment shown in (b). Sedimentation coefficients were calculated using SedFit v9.4. The hydrodynamic parameter, $S_{20,w}$, was obtained by multiplying the sedimentation coefficient by a conversion factor. Additional parameters (diffusion coefficient (*D*), frictional ratio (*f(f0)*), Stokes radius (R_H), and the axial ratio (*a/b*) of a prolate ellipsoid) were calculated for each sedimentation coefficient using the SedFit program. R_H and *D* are expressed in nm and cm^2 sec^{-1}, respectively. A partial specific volume of 0.53 cm^3 gm^{-1} and a hydration value of 0.59 g/g for the RNA were used for these calculations.



(b)

Retention Time (ml)			
	0.03 mM	10 mM	Difference
Wt	12.9	14.0	1.1
M3	12.9	13.0	0.1

10 mM	Wt	M3	Difference
Mg^{2+}	14.0	13.0	1.0
Mn^{2+}	14.1	13.2	0.9
Ca^{2+}	14.2	13.2	1.0
Sr^{2+}	14.2	13.1	1.1
Co^{2+}	14.2	13.5	0.7

Figure 25: Size exclusion chromatography confirms that Mg^{2+} induces a significant change in hydrodynamic radius. **(a)** The extended, low Mg^{2+} conformation was demonstrated to be separable from the compact, high Mg^{2+} conformation via size exclusion chromatography. **(b)** Given that M3 RNA serves as a control for the low Mg^{2+}, extended conformation, retention times were compared for wild-type and M3 RNAs equilibrated in alternate divalent metals. These data suggest that alternate divalent ions can also induce the compacted conformation.

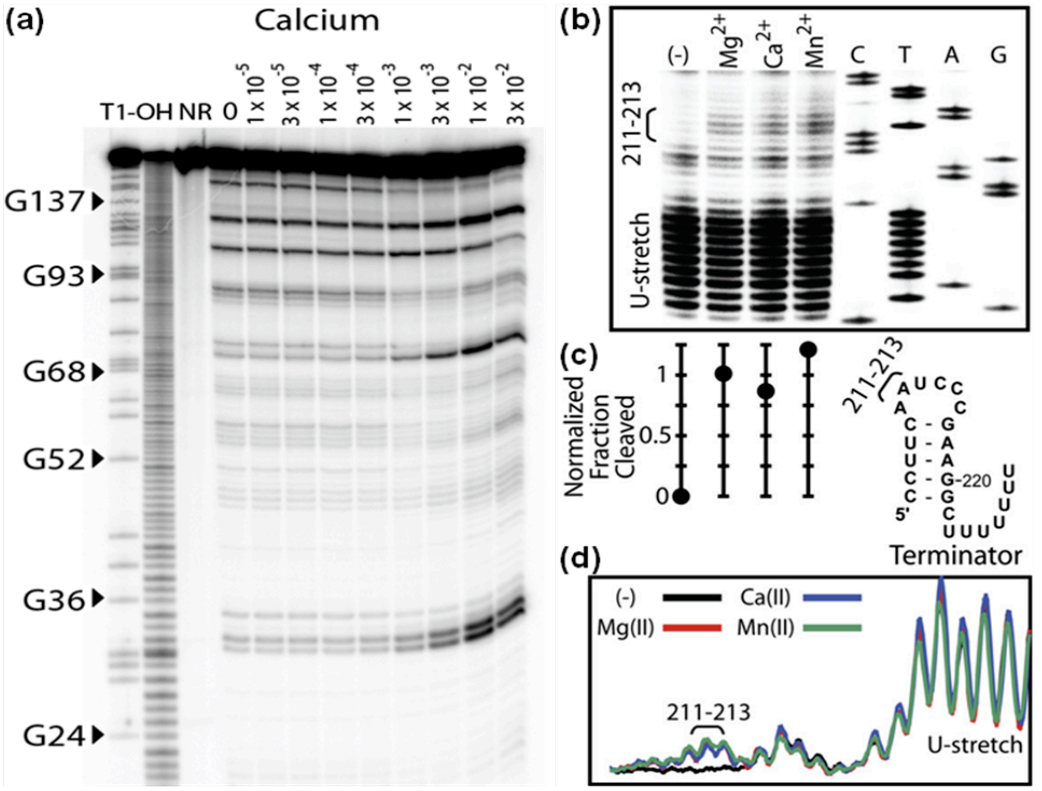


Figure 26: Other divalent ions elicit the same conformational changes in the M-box RNA as Mg^{2+}. **(a)** In-line probing assays in which Ca^{2+} is titrated into the reactions in place of Mg^{2+} result the exact same probing changes and have the exact same apparent EC_{50}. NR, T1, and OH identify no reaction, partial digestion with nuclease T1, and partial digestion with alkali, respectively. **(b)** SHAPE probing of the terminator region of the full length M-box element supports the formation of the terminated state in the presence of Mg^{2+}, Ca^{2+}, or Mn^{2+}. **(c)** Quantification of band intensity at positions 211-213. **(d)** Line traces to highlight this change in band intensity only occurs at positions 211-213 as compared to other surrounding regions of the gel.

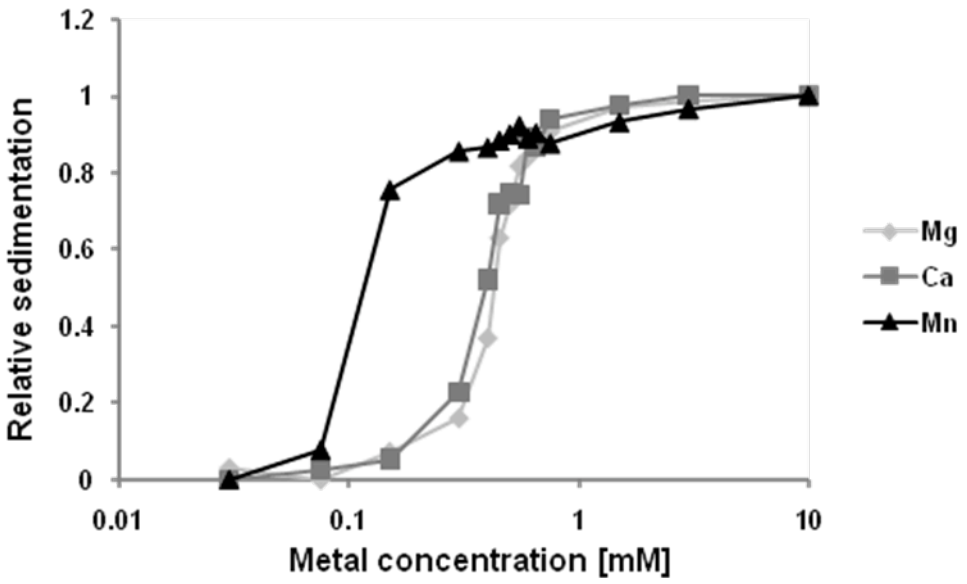


Figure 27: AUC analysis with different divalent ions. As the probing data indicated, Mg^{2+} and Ca^{2+} induce compaction of the M-box RNA at the same concentration. However, the affinity of Mn^{2+} for the M-box RNA appears to be significantly higher, inducing compaction at approximately a 4-5 fold lower concentration.

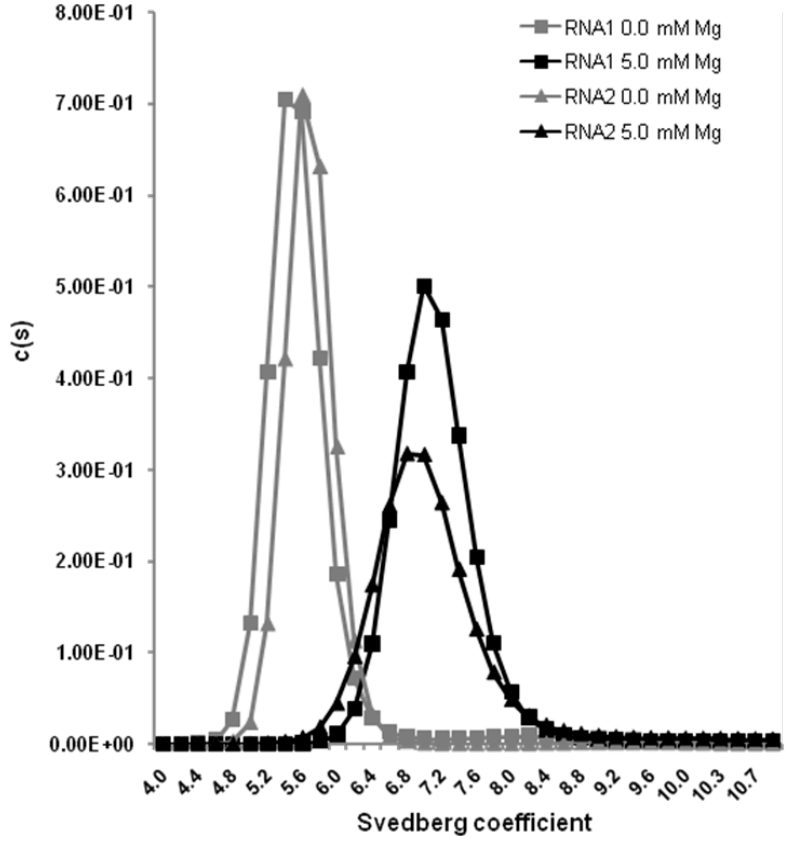


Figure 28: AUC analysis of *C. acetobutylicum* M-box RNAs. Two RNAs from *C. acetobutylicum,* labeled here as RNA1 and RNA2, appear to regulate the gene expression of Mn^{2+} transporters. While both of these RNAs regulate Mn^{2+} transport, they both appear to be responsive to Mg^{2+} in vitro as indicated by their ability to transition from sedimentation values of ~5.5 to 7.0 upon the addition of 5 mM Mg^{2+}.

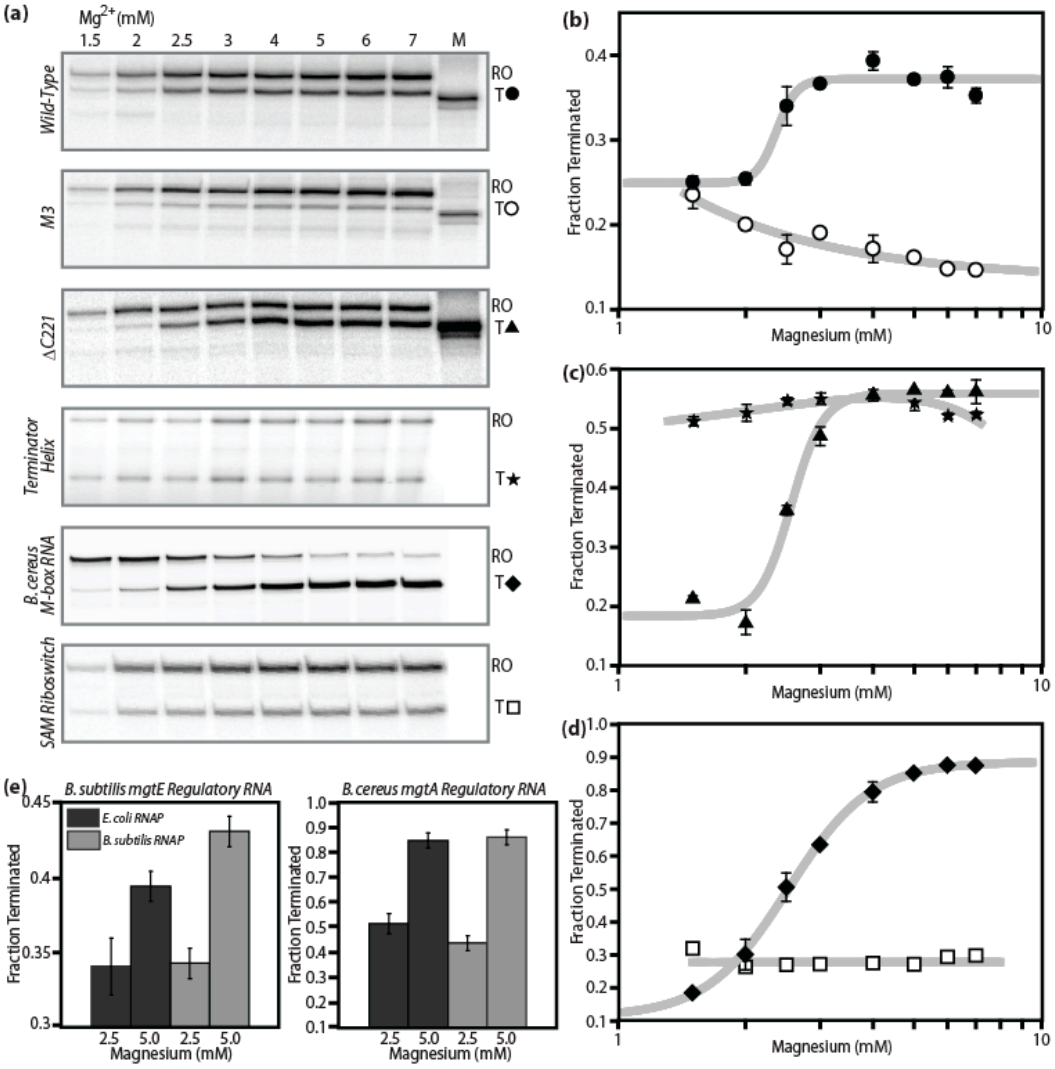


Figure 29: Mg^{2+} induces transcription termination in vitro. **(a)** Transcription assays using templates encompassing wild-type, M3, ΔC221, or terminator sequences from the *B. subtilis mgtE* 5′ UTR, a *B. cereus* M-box RNA, and a SAM riboswitch at multiple [Mg^{2+}]. Lower and upper bands correspond to termination at the intrinsic terminator (T) or run-off transcription (RO), respectively. A size marker for terminated transcripts is shown in the last lane of the first three panels. Only M-box RNAs with an intact aptamer and terminator show Mg^{2+}-dependent transcription termination. Mg^{2+} levels do not affect the termination of M3, terminator alone, or the SAM riboswitch transcript. **(b-d)** The fraction termination is shown plotted against Mg^{2+} concentration. **(e)** Assays were conducted using *B. subtilis* and *E. coli* RNA polymerase with DNA templates for *B. subtilis mgtE* and *B. cereus mgtA*. Figure from Dann, Wakeman et al., 2007.

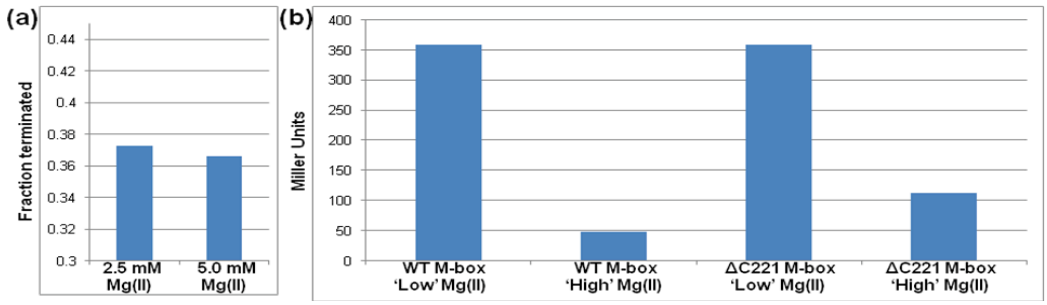


Figure 30: ΔC221 construct appears to interfere with *B. subtilis* RNA polymerase function. **(a)** In vitro transcription termination assays using *B. subtilis* RNA polymerase show that increasing Mg^{2+} concentrations does not lead to increased termination for this construct. However, this construct was able to function to increase transcription termination in a Mg^{2+}-dependent manner when *E. coli* RNA polymerase was used (see Figure 29 a and c). **(b)** Reporter gene fusions with WT and ΔC221 M-box RNAs show that both constructs reach approximately the same level of gene expression under growth conditions containing 5 μM Mg^{2+}. However, under 2.5 mM Mg^{2+} growth conditions, the WT M-box RNA is able to repress gene expression to a greater extent.

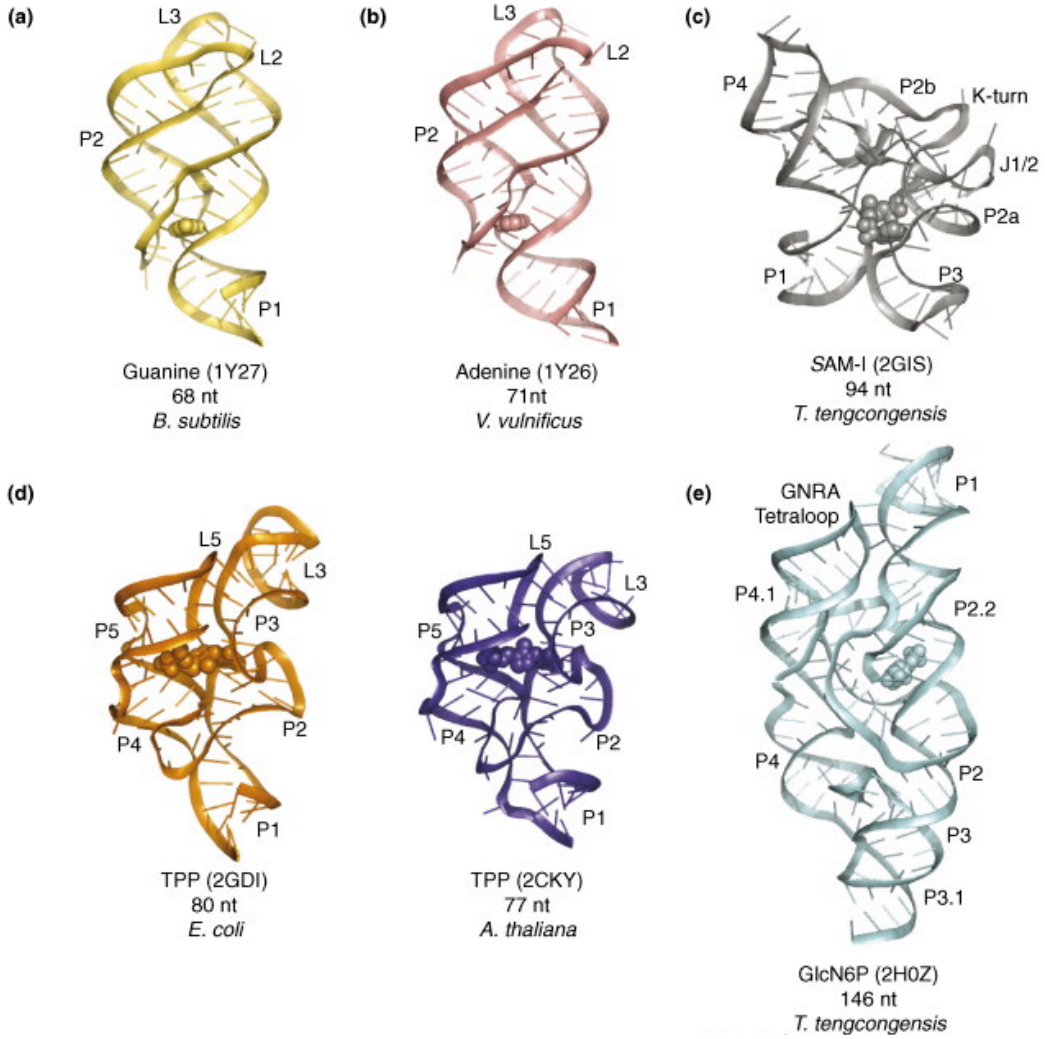


Figure 31: Global visualization of riboswitch structure and ligand binding. **(a)** Guanine riboswitch. **(b)** Adenine riboswitch. **(c)** SAM-I riboswitch. **(d)** TPP riboswitch. **(e)** *glmS* ribozyme (GlcN6P riboswitch). The Protein Data Bank (PDB) accession codes are given in parentheses, along with number of nucleotides and source organism. RNAs shown are bound to their natural ligand (spheres) except for the GlcN6P-binding ribozyme **(e)**, which is bound to glucose-6-phosphate. Figure from Wakeman et al., 2007.

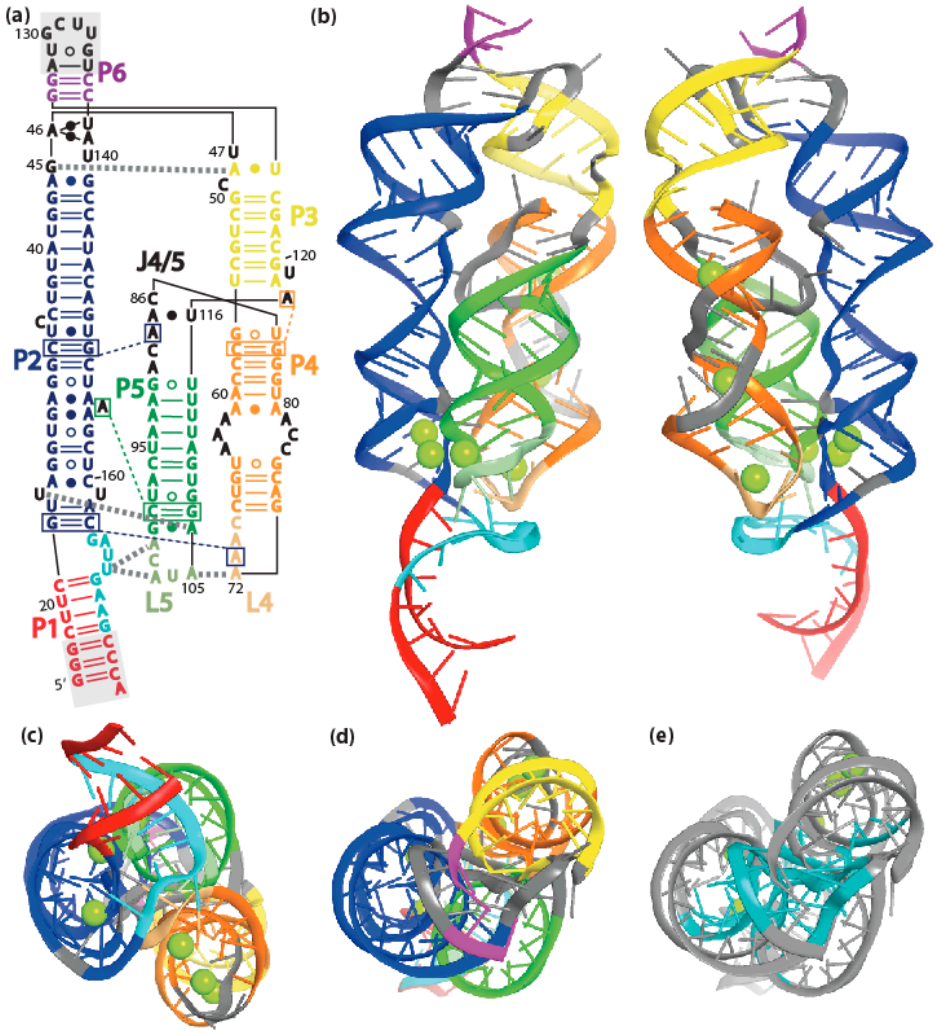


Figure 32: **(a)** Secondary structure diagram based on the tertiary structure: P1-red; P2-blue; P3-yellow; P4/L4-orange; P5/L5-green; P6-magenta; joining (J) regions-black; antiterminator-cyan. Open and closed circles mark G:U and noncanonical pairs, respectively. Long-range base pairs are connected by heavy dashed lines. Four A-minor motifs are boxed and connected by thin dashed lines. Gray shaded residues indicate positions added for crystallization or positions that are disordered in the structure. **(b)** Two views of the aptamer domain model are shown with six Mg^{2+} depicted as green spheres. Colors are as in **(a)**. **(c)** A bottom-up view. **(d)** Top-down view. **(e)** Nucleotides that are protected based on hydroxyl radical footprinting in the presence of Mg^{2+} are highlighted in teal, showing correlation between protected residues (Figure 23) the inner core of the RNA structure. Image from Dann, Wakeman et al., 2007.

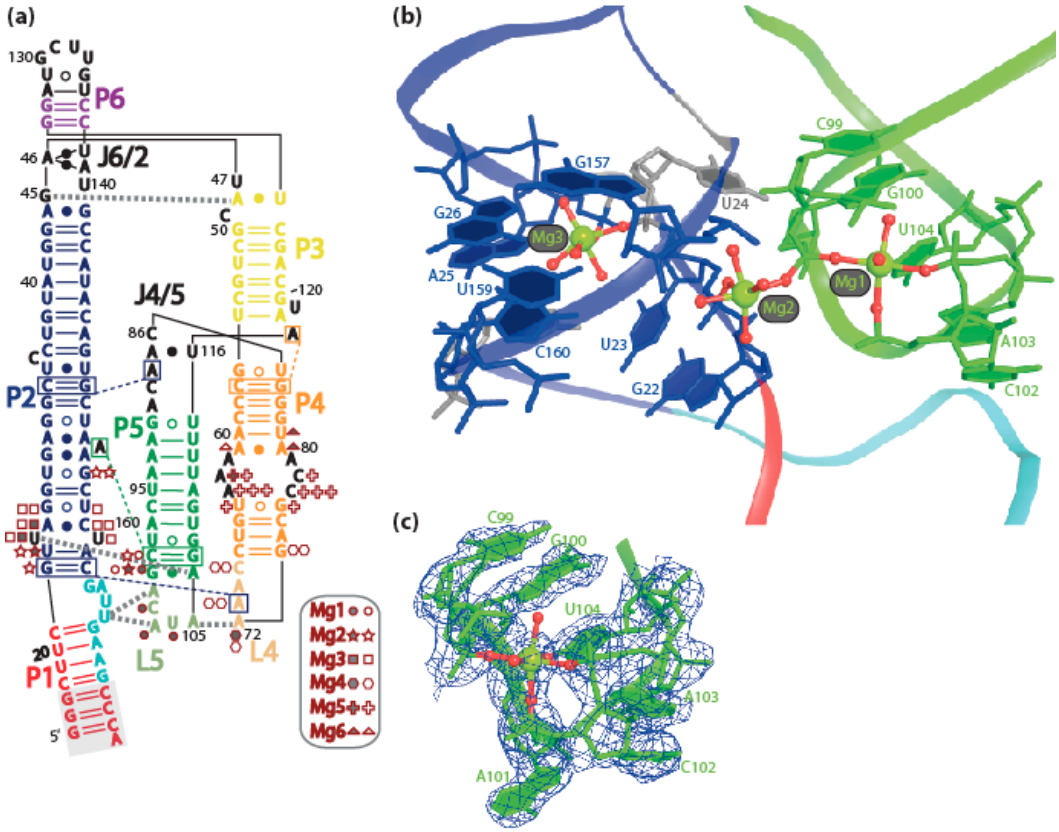


Figure 33: **(a)** Secondary structure shows the positions of inner- and outer-sphere contacts using shaded and open symbols, respectively, for the six Mg^{2+} coordinated to the RNA. Each Mg^{2+} has a unique symbol denoted in the legend. **(b)** Nucleotides that contact Mg1-3 via inner- and outer-sphere interactions are labeled and shown as sticks on a ribbon model of the phosphate backbone. P4 and L4, which would be located in the background of this image, have been omitted for clarity. A total of 4, 2, and 2 direct RNA oxygen atoms are coordinated to Mg1, Mg2, and Mg3, respectively. Only one of these RNA contacts (U104 O4) does not involve a nonbridging phosphate oxygen. Residue G100 is flanked by Mg1 and Mg2 coordinating to its two nonbridging phosphate oxygens. Residues that contact Mg1-3 cluster near the tripartite region of long-range tertiary contacts formed by L4, L5, and P2. **(c)** The experimental electron density map is shown for Mg1 and nucleotides contacted through inner- and outer-sphere interactions. Image from Dann et al. 2007.

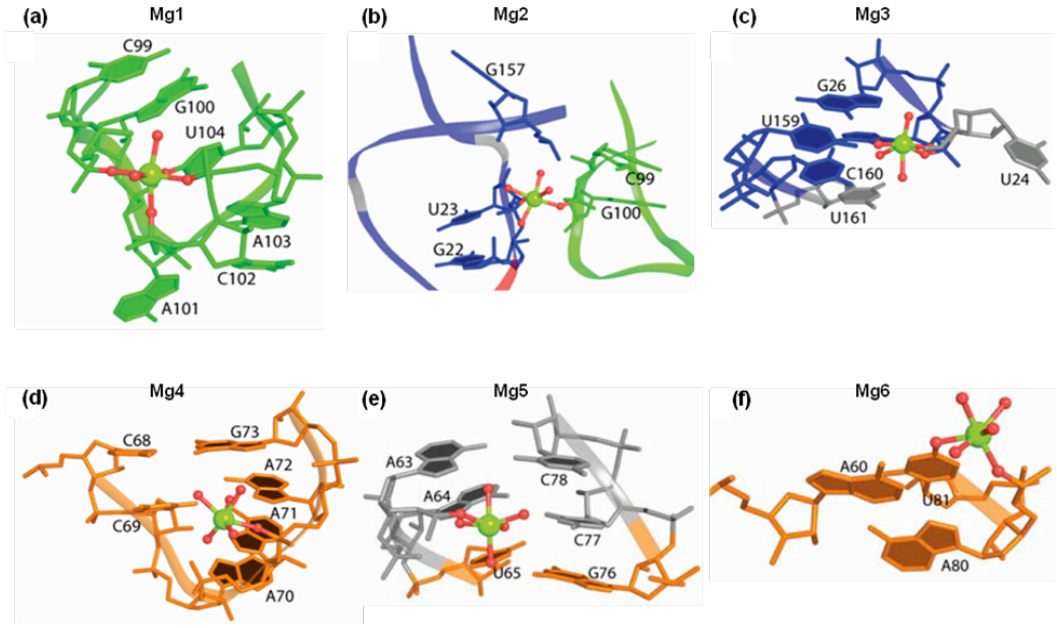


Figure 34. Binding sites of Mg1-6. Mg^{2+} are shown as green spheres. Bonds between Mg^{2+} and RNA or H_2O are shown in red.

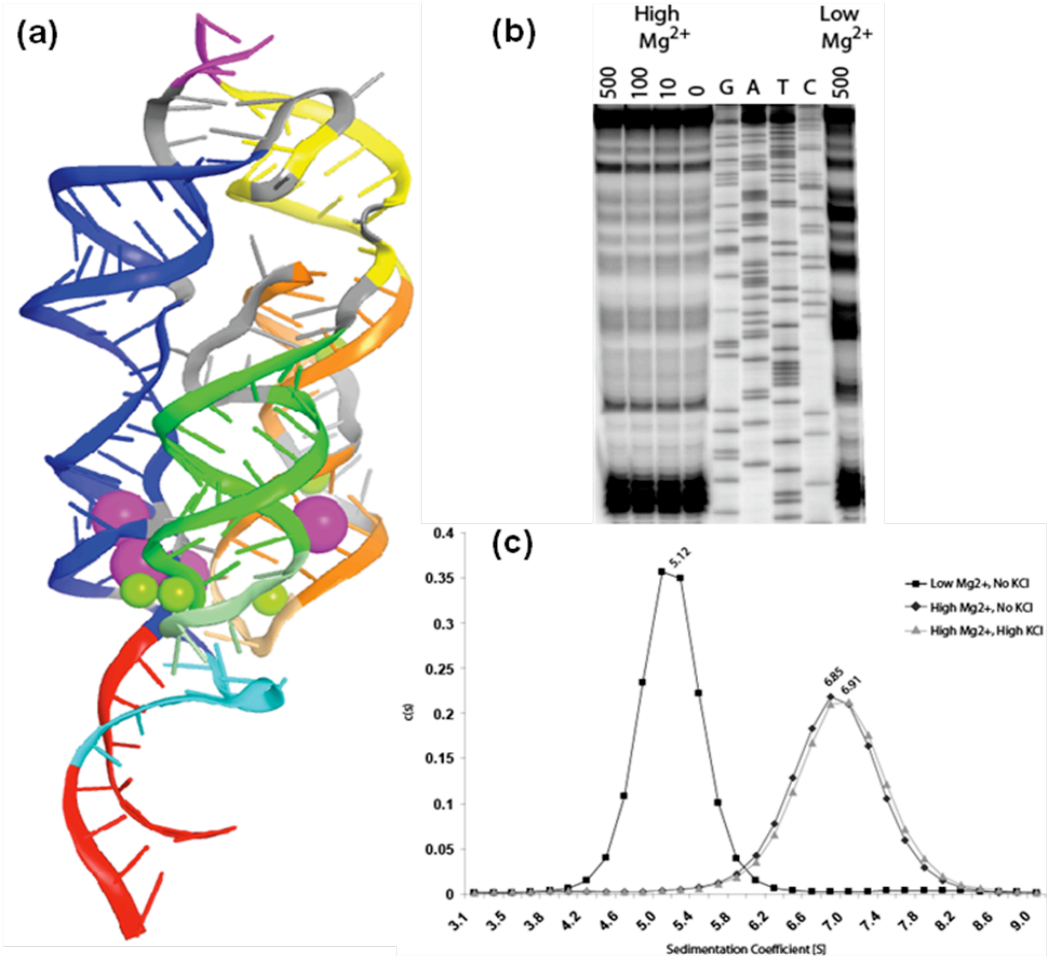


Figure 35: K^+ is not required for Mg^{2+} induced tertiary formation. **(a)** Three dimensional structural model of the M-box RNA showing the locations of six bound Mg^{2+} (green) and four bound K^+ (magenta). **(b)** SHAPE probing in 10 mM Mg^{2+} and with increasing K^+ revealed that the characteristic Mg^{2+}-induced probing changes occurred in the presence or absence of K^+. The numbers above the gel image display KCl concentrations (mM) used in each reaction. **(c)** AUC demonstrates that the M-box compacts to same extent in the presence or absence of K^+.

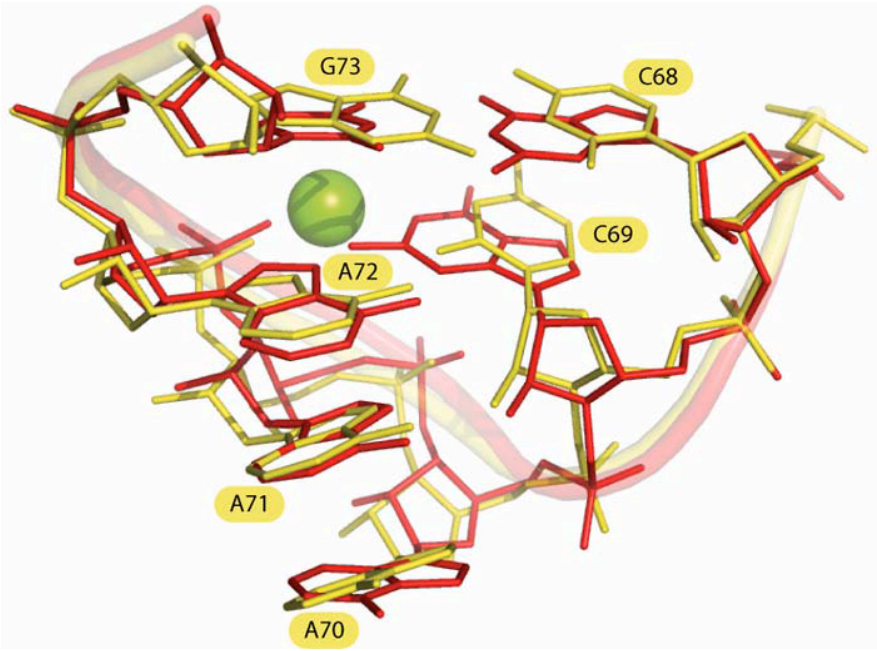


Figure 36: Superimposition of the L4 tetraloop with the P4-P6 GNRA tetraloop. The L4 CAAA tetraloop (yellow) was superimposed with the well-characterized GAAA tetraloop from the P4-P6 domain of the *Tetrahymena thermophila* group I intron (red) from Cate et al., 1996. In the latter structure a sheared G:A pair is formed between the 5' and 3' nucleotides which then stack upon the adjacent helix. The *Tetrahymena tetraloop* docks into a conserved 11 nucleotide receptor domain in the final tertiary conformation. It is interesting that the *B. subtilis mgtE* L4 tetraloop is structurally similar to the GNRA tetraloop despite the fact its sequence (CAAA) violates the consensus pattern of that motif. However, both tetraloops are responsible for forming long-range interactions. From superimposing the two tetraloop structures it is apparent that Mg4, which interacts with L4 residues, occupies part of the location of the 5' G of the group I intron GAAA tetraloop. This metal is located in different position relative to the Mg^{2+} present in some examples of the classical GNRA tetraloops. Image from Dann, Wakeman et al., 2007.

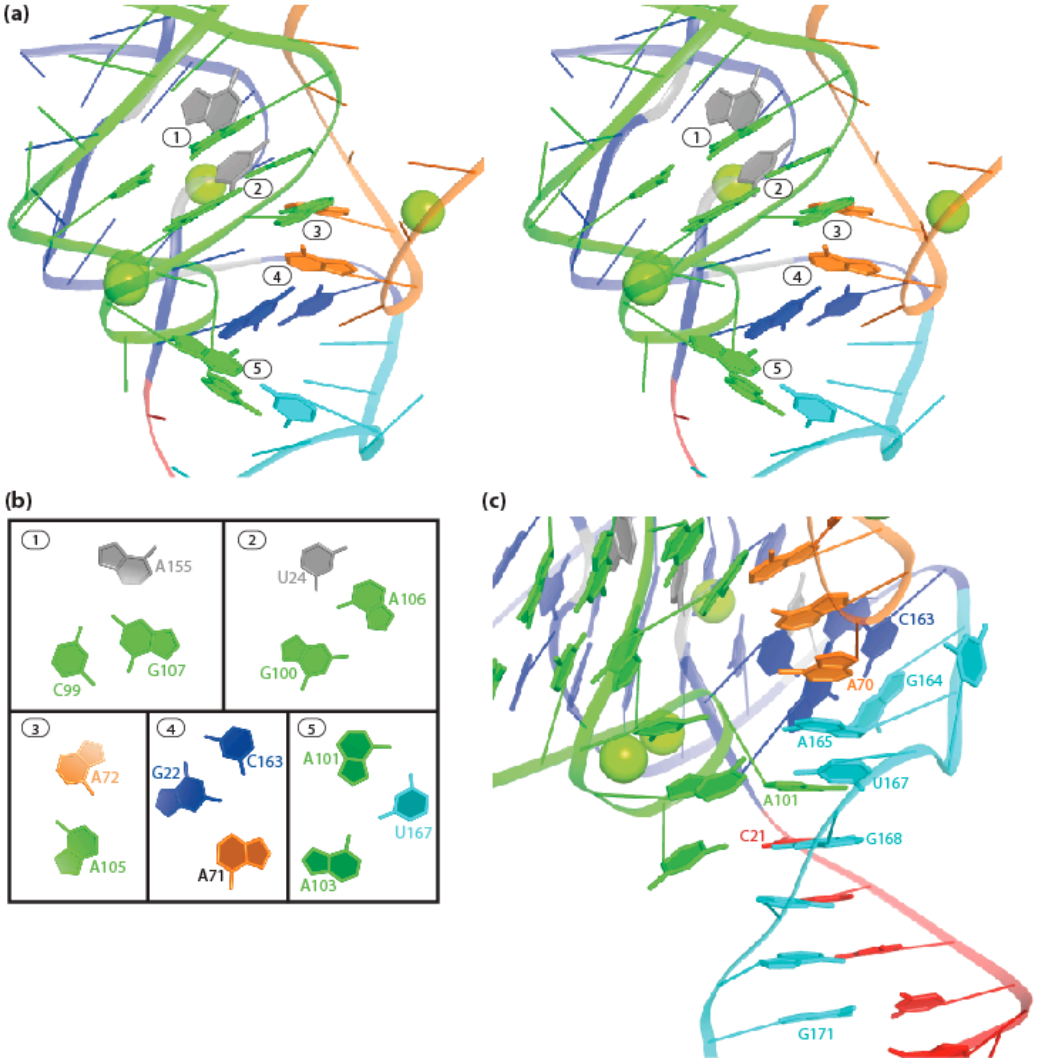


Figure 37: Sequestration of antiterminator nucleotides via long-range base interactions and base stacking. **(a)** A stereo view comprising the region of the molecule where P2, L5, L4, and antiterminator nucleotides converge. Nucleotides involved in long-range base interactions are shown as filled bases. A total of five such interactions stabilize the region of the molecule adjacent to antiterminator nucleotides. Base pairs involving G168–G171 complete the sequestration of the antiterminator. **(b)** The interactions numbered in **(a)** are presented individually from a top-down perspective. **(c)** In addition to interactions highlighted in **(a)** and **(b)**, this region of the molecule is enriched in stacking between P1, P2, L4, and L5 residues. These interactions are likely to be important in stabilizing the overall compacted tertiary conformation and for sequestration of antiterminator nucleotides (cyan).

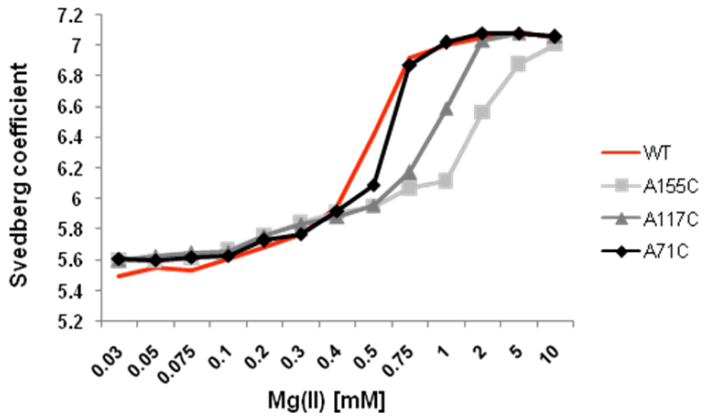


Figure 38: AUC analysis of A-minor motif mutants. When Mg^{2+} is titrated into samples containing M-box RNAs, these RNAs transition from an extended state sedimenting at ~5.5 S to a compacted state sedimenting at ~7.0 S. While none of the A-minor motif mutants are able to transition as efficiently as WT M-box, the impact of some mutations are clearly more severe than others.

	Svedberg coefficient at 2.5 mM Mg^{2+}	Fold repression of gene expression in vivo
WT M-box	7.0 S	7.5
M3	5.82 S	1.4
U104A	6.85 S	2.1
U24A	6.85 S	1.9

Table 3: AUC and *lacZ* reporter fusion analysis of M-box point mutants. While AUC analysis indicates that these point mutations are more similar to WT RNAs than M3 mutant RNAs, the effect of these mutations in vivo is severe, rendering the U104A and U24A mutants almost as inactive as the M3 mutant.

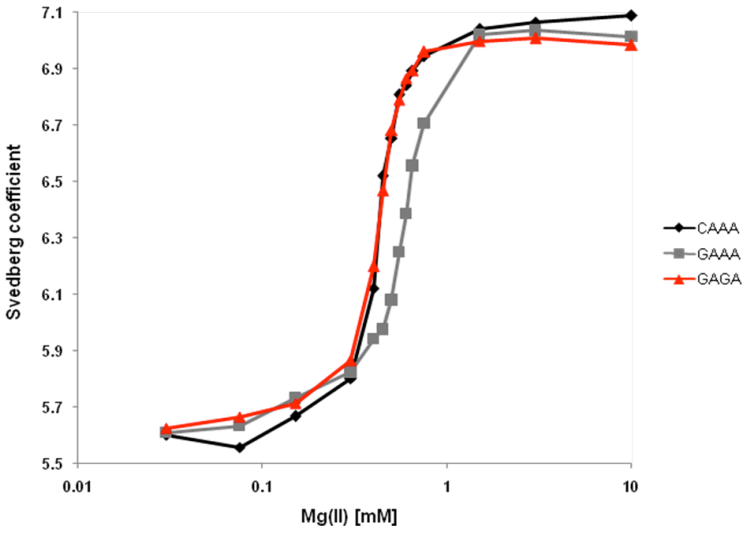


Figure 39: Tetraloop analysis using AUC. Velocity sedimentation shows that a Mg^{2+}-induced structural change occurs at the same Mg^{2+} concentration for both WT *B. subtilis* M-box (CAAA) and WT *B. cereus* M-box (GAGA) but not for a *B. subtilis* mutant (GAAA).

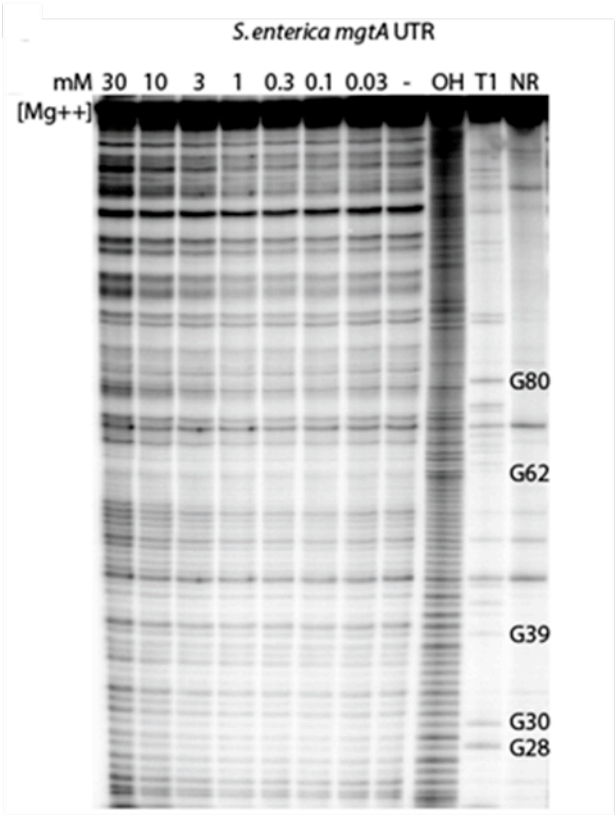


Figure 40: In-line probing of the *S. enteric mgtA* RNA in high monovalent ion concentrations. NR, T1, and OH identify no reaction, partial digestion with nuclease T1, and partial digestion with alkali, respectively. Numbers denote Mg^{2+} concentration in mM.

CHAPTER FIVE

IDENTIFICATION OF THREE METAL-BINDING CORES

INTRODUCTION

Typically, riboswitch aptamer domains bind to single small molecule metabolites such as enzymatic cofactors, amino acids, and nucleobase-containing ligands (Dambach and Winkler, 2009). In contrast, the X-ray crystallographic structural model of the M-box riboswitch aptamer revealed the absence of an organic metabolite ligand but the presence of at least six tightly associated Mg^{2+} (Dann et al., 2007). Additional biochemical and biophysical analyses of this RNA described in Chapter 4 confirmed that binding of divalent ions leads to formation of a compact tertiary architecture. This Mg^{2+}-bound compacted conformation results in sequestration of an oligonucleotide stretch within the aptamer domain, thereby preventing the formation of an antiterminator helix and allowing instead an intrinsic terminator helix to decrease downstream mRNA synthesis (Figure 41a). These data agreed well with the proposed role of the M-box riboswitch in functioning as a sensor of intracellular Mg^{2+} that was initially based on the observation that expression levels of M-box-regulated genes appeared to fluctuate in a Mg^{2+}-responsive manner. However, even if this RNA is indeed a direct sensor of Mg^{2+}, some additional nonspecific metal interactions are also undoubtedly required for these purposes since, in general, all RNAs require metal ions to promote folding. Presumably, the presence of multiple ligand-binding sites in this RNA allows for cooperative binding of

Mg^{2+} and may assist genetic control over a more narrow range of ligand concentration than for noncooperative riboswitches. Cooperative binding of riboswitch ligands has also been observed for glycine-sensing RNAs, which utilize a tandem arrangement of glycine-binding aptamer domains (Mandal, et al., 2004). However, unlike the glycine riboswitch M-box RNAs bind multiple ligand molecules within a single aptamer domain. Understanding the molecular basis for this metal ion network is critical for understanding M-box metallosensory function.

Although most structurally complex RNAs also bind to Mg^{2+}, at least 2 other unrelated classes of regulatory RNAs, specifically the glycine- and the GlcN6P-sensing riboswitches, appear to be conformationally altered by Mg^{2+} over a ~100-fold range, likely making these RNAs relatively insensitive to small fluctuations in intracellular Mg^{2+} (Draper, 2008). This trend will likely apply to most other known riboswitch classes as well. In contrast, M-box RNAs appear to be capable of responding to small (2-3 fold) changes in Mg^{2+} concentration (Dann et al., 2007). Additionally, despite the general requirement of cations for the folding of structured RNAs, most high resolution RNA crystal structures do not contain even a single metal ion with multiple inner sphere contacts to the RNA (Draper, 2008), a stark contrast to the M-box riboswitch. While the structural model derived by X-ray crystallography gives some insight into the molecular features required for this unique relationship between the M-box RNA and its Mg^{2+} ligands, it is still likely to represent an incomplete portrayal. Indeed, the structural model provides a useful but static three-dimensional image of the final, ligand-bound molecule, thereby yielding little information of the structural and molecular requirements during RNA folding. Therefore, to gain greater insight into their molecular mechanisms, M-box

RNAs were subjected to analysis by nucleotide analog interference mapping (NAIM), a chemogenetic approach for studying functionally required features at an atomic level (Strobel and Shetty, 1997; Ryder and Strobel, 1999; Waldsich, 2008).

NAIM analyses are performed in three steps. First, nucleotide analogs are randomly incorporated into RNA molecules. This random incorporation is accomplished through the addition of a low level of the nucleotide analog in an otherwise typical in vitro transcription reaction. Many different nucleotide analogs are regularly used for this purpose so the ratio of nucleotide analog/nucleotide required for introduction of a low level of analog incorporation per RNA molecule is well known for a number of these analogs. Typically, a single incorporation per RNA molecule is ideal (Christian and Yarus, 1992). Most of these analogs can be incorporated with WT T7 RNA polymerase, but sometimes a Y639F mutant is required to incorporate some nucleotide analogs in order to eliminate discrimination between certain functional groups (Ryder and Strobel, 1999). Once the nucleotide analog has been incorporated, the second step of NAIM analysis involves a functional separation of the RNA molecules based on their ability to perform a certain task. This selection can be as complex as assaying the ability of the RNA to self cleave or splice or it can be as simple as monitoring the ability of the molecule to fold in a compacted state (Waldsich, 2008). Finally, the third step of NAIM analysis is to visualize sites of nucleotide analog incorporation via cleavage with iodine (Gish and Eckstein, 1988). These iodine cleavage reactions will appear similar to sequencing gels when resolved by PAGE. Selected pools are compared to non-selected or ‘parental’ pools in order to visualize positions at which the nucleotide analog caused interference or enhancement in function. Bands representing positions of interference will

appear lighter as compared to the parental band and bands representing positions of enhancement in function will appear darker. A schematic detailing the three steps of a NAIM analysis in which molecules were separated based on their ability to fold can be found in Figure 42.

Most inner sphere interactions to Mg^{2+} occur through non-bridging phosphate oxygens. Therefore, random atomic substitution of phosphate oxygens for sulfur can be used to abrogate binding of Mg^{2+} due to the fact that 'hard' metals (*e.g.,* Mg^{2+}) prefer to interact with 'hard' rather than 'soft' ligands (*e.g.,* sulfur) (Eckstein, 1985). Sulfur substitutions at metal sites that are important for the compact conformation of M-box RNAs will result in RNA molecules trapped in the 'extended' state. Nucleotides containing a sulfur group in place of a non-bridging phosphate oxygen of the α phosphate group are known as phosphorothioate nucleotide analogs. While in some cases interferences may simply occur due to steric effects caused by incorporation of the bulkier sulfur group (Basu and Strobel, 1999; Waring, 1989), phosphorothioate modifications are considered to be generally useful in studying essential Mg^{2+}-RNA contacts. Herein several phosphorothioate interference sites for an M-box RNA from *B. subtilis* are identified. These sites cluster into three separate regions within the RNA. Two of these regions have been shown by prior structural analyses to be metal binding cores as described in Chapter 4 and are functionally confirmed by data herein. A third potential metal binding core was unidentified by the prior structural studies but revealed as important for tertiary structure formation by phosphorothioate interference mapping. In total, phosphorothioate interference analyses suggest that M-box RNAs utilize a complex network of direct metal binding sites in order to exert metalloregulatory control.

RESULTS

Separation of extended and compacted M-box RNAs

Nucleotide analog interference mapping (NAIM) was used in this study to assess the relationship between divalent ions and nonbridging phosphate oxygens for M-box RNAs. Specifically, populations of RNA molecules containing random phosphorothioate substitutions were synthesized and functionally separated into pools of Mg^{2+}-bound and unbound RNAs based on their ability to adopt the previously described, Mg^{2+}-induced compacted architecture determined in Chapter 4 and summarized in the schematic in Figure 41a.

While many RNA sequences contain the requisite information for folding of intricate three-dimensional structures, a number of these noncoding RNAs adopt structurally distinct and functionally specialized conformations in response to specific cellular signals such as the phosphorylation state of accessory protein factors (ex. LicT), temperature/physical status (ex. the *pfrA* RNA), and the binding of small organic molecules (reviewed in Winkler, 2005a). Additionally, most of these RNAs will exhibit large structural changes upon titration of cations. These conformational transitions are oftentimes accompanied by changes in hydrodynamic radii. Therefore, experimental methods that measure changes in hydrodynamic radius can be employed for study of signal-induced RNA conformational changes.

Previous studies have demonstrated that the M-box riboswitch undergoes molecular compaction upon association with Mg^{2+}, resulting in a significant reduction in molecular radius in the Mg^{2+}-bound state (Dann et al., 2007). This Mg^{2+}-induced change in molecular shape for the separation of extended and compacted complexes was exploited for separation by size exclusion chromatography. Specifically, single phosphorothioate modifications were randomly incorporated into M-box RNAs during in vitro transcription and the resulting RNA molecules were separated on a size exclusion column in the presence of Mg^{2+} (Figure 41b). Most phosphorothioate substitutions appeared to be neutral in that they did not affect the ability of the RNA to adopt the Mg^{2+}-induced compacted state. However, an overlapping peak was observed that exhibited a retention volume matching RNAs trapped in the extended conformation. RNA molecules corresponding to both peaks were collected, radioactively labeled at the 5' terminus, and further resolved by nondenaturing polyacrylamide gel electrophoresis in the presence of Mg^{2+} (*i.e.,* native PAGE) (Figure 41c). During native PAGE, the RNA pools were electrophoresed alongside wild-type RNAs to mark the location of the Mg^{2+}-bound, compact conformation and alongside a previously characterized mutant, 'M3' (Dann et al., 2007), to mark the location of the extended conformation. These experiments were separately performed for RNA molecules incubated in the presence of 0.5 mM Mg^{2+}, which is close to the Mg^{2+} concentration that triggers half-maximal compaction (Dann et al., 2007), and 1.0 mM Mg^{2+}, a saturating Mg^{2+} concentration at which the majority of the WT RNAs should adopt the final compacted state.

The fact that the M3 mutant does not appear to experience any significant compaction under saturating Mg^{2+} concentrations can be seen as a testament to the

requirement of specific Mg^{2+}-binding sites for the overall tertiary structure of the RNA. In general, increasing cation concentrations allow for the compaction of RNAs due to the neutralization of the negatively charged phosphate backbone. However, a point mutation in one conserved region of this ~160 nucleotide-long RNA molecule (Figures 8a and 16b,c) almost completely eliminates any Mg^{2+}-induced structural effects. Upon visualization of the three-dimensional structural model, it became apparent that this mutation is likely to be affecting the formation of the Mg1 binding pocket, the Mg^{2+} site originally predicted to be the most likely to be essential due to its four inner sphere contacts to the M-box RNA. Therefore, disruption of a single Mg^{2+}-binding site appears to be sufficient to trap the RNA in the extended state.

Another interesting selection condition would be to test the 0.5 mM and 1.0 mM Mg^{2+} concentrations in a background of high concentrations of monovalent ions (1 – 2 M NaCl or KCl versus 0.1 – 0.2 M). It has been shown previously that high levels of monovalent ions can sometimes supplant the need for divalent ions in certain RNA structures (Das et al., 2005a). Therefore, a subset of the interferences observed under low monovalent concentrations might disappear, indicating that while they are essential for compaction, they are not strictly divalent binding sites. Preliminary attempts have been made in order to separate RNAs on a native gel under these conditions (data not shown); however, even after 8 hours of electrophoresis, the RNA barely enters into the gel.

Phosphorothioate interferences at previously established Mg^{2+}-binding sites

To identify sites of sulfur substitution, phosphorothioate linkages were cleaved via iodine oxidation (Gish and Eckstein, 1988) and the RNA fragments were resolved by denaturing PAGE for both the 0.5 mM Mg^{2+} and the 1.0 mM Mg^{2+} conditions (Figures 43 and 44). Sulfur modifications that are detrimental to a particular conformational state result in fainter bands as compared to the parent pool of RNA molecules, in which the extended and compacted species were not separated. Conversely, sites of phosphorothioate enhancement exhibit greater band intensity as compared to the unselected pool. For analysis of phosphorothioate-substituted M-box RNAs, interference values were calculated from the ratio of band intensities for functionally-separated RNAs and the unseparated, parent pool. These values were then normalized to account for minor loading differences (Figure 45). These interference values are similar to the κ-values used in most NAIM analyses (Ryder and Strobel, 1999). In this study, interferences were grouped into 3 categories: strong interferences (interference values >10), moderate interferences (interference values = 5-10), and weak but reproducible interferences (interference values = 2-5). Nucleotide positions at which phosphorothioate substitutions have virtually no effect have an interference value of approximately 1 in this analysis. Enhancements have interference values of 0.5 or less.

The 6 metal binding sites (labeled as Mg1-Mg6) identified by the prior structural analyses of the *B. subtilis mgtE* M-box RNA detailed in previous chapters were proposed to incorporate a total of 12 inner sphere interactions between divalent ions and RNA oxygens, occurring mostly between Mg^{2+} and pro-R_p non-bridging phosphate oxygens.

Based upon the functional analysis provided by the NAIM data, 7 of these inner sphere interactions, representing all of the known pro-R_p oxygen inner sphere coordinations, were detected as sites of phosphorothioate interference when the RNA molecules were separated into extended and compacted pools in the presence of 0.5 mM Mg^{2+} (Figure 46). Moreover, at least one deleterious sulfur substitution was observed for each of the 6 metal sites, indicating that all of the crystallographically observed divalents are required for metalloregulatory function. Particularly strong interferences were observed at phosphate oxygens located 5' of U24, G100, C102, A103, and G157, which together interact with 3 of the 6 ions (Mg1-Mg3). Notably, all three of the phosphate oxygen contacts to Mg1 are sites of strong interference (G100, C102, A103), implicating Mg1 as a particularly important ion for the compacted state. Mg2 is coordinated by G157 through a Mg^{2+}-chelated water molecule. While phosphorothioate substitutions typically interfere with inner sphere coordination of Mg^{2+}-oxygen interactions, phosphorothioate disruption of outer sphere Mg^{2+}-RNA interactions have also been observed (Basu and Strobel, 1999), offering an explanation for the strong interference at G157. U23 and G100 form inner sphere contacts to Mg2 via their pro-S_p phosphate oxygens but the importance of these contacts could not be investigated in these experiments due to the fact that T7 RNA polymerase preferentially inserts phosphorothioate modifications in a pro-R_p configuration (Griffiths et al., 1987). Moderate interferences were observed for 5' phosphates of A72 and A63 contacting Mg4 and Mg5, respectively, suggesting an important but less critical role for these Mg^{2+}. Finally, a weak but reproducible interference was observed at the 5' phosphate of A80, which contacts Mg6. In total, these

biochemical data confirm the importance of all of the metal sites that were predicted by prior structural analyses.

When the RNA molecules were separated into extended and compacted pools in the presence of 1.0 mM Mg^{2+}, interferences at positions coordinating only 3 of the 6 known Mg^{2+} were observed (Mg1, Mg3, Mg4), suggesting that metal excess can bypass the requirement for certain Mg^{2+} sites (Figure 44). At this concentration, strong interferences were observed at linkages responsible for coordinating Mg1 (G100, C102, and A103) and Mg3 (U24). Also, a weak but reproducible interference occurred for the 5' phosphate of A72, which contacts Mg4. In general, positions that are identified as sites of interference in the compacted pool show up as sites of enhancement in the extended pools, indicating that phosphorothioate substitutions at these positions result in molecules that can no longer adopt the compacted conformation and are instead trapped in the extended state.

Sites of additional phosphorothioate interferences

In addition to phosphorothioate interferences corroborating crystallographically predicted Mg^{2+}-binding sites, strong interferences were also measured for compact RNAs at the 5' phosphates of C35, C36, G91, and A101. Moderate interferences were also measured for C89 and A105, and a weak but reproducible interference was measured for U119 when the RNA pools were separated under 0.5 mM Mg^{2+} conditions (Figure 47). Additionally, A101 appears as a strong interference even under 1.0 mM Mg^{2+} separation conditions. None of these sites correlate with known metal binding sites, yet their

locations within the structural model have allowed us to propose explanations for their importance in metal-mediated compaction. One exception is a weak but reproducible interference that appears in the extended RNA pool at position C33 that is more difficult to explain. Previous studies have noted that individual phosphorothioate substitutions can stabilize formation of certain, unfavored RNA structures (Horton et al., 2000). The interference at C33 for extended RNAs may similarly result from stabilization of an alternate, unidentified state that is slightly favored over the extended conformation upon incorporation of the C33 phosphorothioate.

Interferences at A101 and A105 occur within the L5 terminal loop, which is responsible for construction of the Mg1 binding pocket as well as multiple long-range contacts to distal parts of the RNA molecule. It appears that interferences at A101 and A105 do not result from disruption of metal coordinations but instead from perturbation of contacts between these residues and other M-box regions by the slightly bulkier sulfur groups. The pro-R_p oxygen of A101 is positioned within 2.7 Å of the A165 N6 exocyclic amine as well as 2.6 Å to the guanosine 2' hydroxyl of G22 (Figure 47b). Disruption of the proper positioning of G22 via phosporothioate incorporation might be particularly detrimental due to the fact that this nucleotide also makes long-range contacts to A71 of L4. Similarly, the phosphorothioate interference at A105 may be due to disruption of an interaction between the A105 phosphate oxygen and N1 of G100, an interaction that is undoubtedly significant for organization of the unique, Mg^{2+}-bound L5 structural fold. Therefore, disruption of the A101 and A105 phosphate interactions are likely to affect both the formation of the critical Mg1 binding site and long-range interactions between L5, L4 and P2 in a manner similar to the M3 point mutation.

The remaining sites of phosphorothioate interferences (C35, C36, G91, C89, and U119) are likely to result from factors other than simple disruption of important hydrogen bonding contacts. Intriguingly, these phosphate oxygens cluster together in a common locus within the RNA molecule, near to several important interhelical junctions within the apical portion of the molecule (Figure 47c-e). Moreover, these phosphate oxygens are oriented in three-dimensional space with neighboring functional groups in a manner that may be consistent with construction of metal binding pockets. It can be speculated from these data that these new sites of interferences are likely to be important for three previously unidentified metal binding sites, designated herein as M7-9.

The candidate M7 binding site is formed in part by the U119 pro-R_p oxygen and the pro-S_p oxygens of A117 and A118 (Figure 47c,f). All three of these phosphate oxygens are oriented in a manner that should allow for the coordination of a metal ion. Interestingly, the distances between these phosphate oxygens are similar to the distances observed between inner sphere Mg^{2+} ligands of other metal binding sites. For example, the average bond distance for inner sphere ligands of the Mg1-6 metal sites is 2.168 Å (Dann et al., 2007), oriented at approximately 90º angles from one another. Therefore, the approximate distance between adjacent inner sphere ligands in the Mg1-6 binding sites is 3.066 Å and the approximate distance of inner sphere ligands on opposite sides of the Mg^{2+} ion is 4.336 Å. The pro-S_p oxygen of A118 is 3.36 Å and 2.57 Å from phosphate oxygens of A117 and U119, respectively, and the A117 and U119 phosphate oxygens are placed in a convergent orientation, separated by 4.116 Å. There are not other obvious candidate ligands within the putative M7 binding site and therefore the remaining three ligands would be expected to be water molecules, if indeed a Mg^{2+} can be chelated to this

locus. The region corresponding to these residues has been highlighted in a surface representation of the M-box RNA, colored by electrostatic potential (Figure 47f). The relatively higher electronegativity of this pocket aids in speculation that this region might constitute a metal binding site.

The putative M8 site would be predicted to encompass the C89 and G91 pro-R_p oxygens, which appear to be oriented toward a central cavity. It may also be noteworthy that A90 and A87 ribose 2' oxygens also appear to be oriented toward this pocket, although 2' ribose oxygens are rare inner-sphere ligands of Mg^{2+} (Figure 47d,g). While this site in the crystal structure contains a large empty space that would seem likely to accommodate a Mg^{2+} and contains more potential coordinating ligands and stronger interference values than the phosphorothioate substitutions at the M7 site, the angles between the potential ligands do not appear to be as convincing. The C89 and G91 non-bridging phosphate oxygens point in slightly different directions, and while this observation does not eliminate the possibility that one or both of these functional groups are coordinating an unseen Mg^{2+} via outer sphere coordination, it does indicate that inner sphere coordination between both of these oxygens and a single Mg^{2+} is unlikely. Another possible explanation for these interferences is that this central cavity represents a transiently occupied metal-binding site that is created during the folding process but not properly formed in the final folded state due to the fact that some of the coordinating ligands have been locked out of proper positioning when certain tertiary contacts formed. One of the strongest observations in support of the M8 site as a putative metal-binding site is the highly negative charge found in this particular location (Figure 47g).

The putative M9 site involves the pro-R_p oxygens of both C35 and C36, which angle towards each other and toward a well-ordered water molecule in the structural model (Figure 47e,h). It is possible that the interaction with this water could itself be essential or that this molecule could be mediating an outer sphere coordination between these phosphate oxygens and a missing metal. It is also possible that the water molecule replaces a metal ion that is present transiently during the folding of the RNA molecule. Consistent with the role of these residues as part of an electronegative metal binding pocket, phosphate groups of 142-144 are also oriented above and point toward the C35-C36 phosphate oxygens. Electrostatic surface potential for this region of the RNA also appears to be highly negatively charged (Figure 47h).

In total, these data suggest that M-box RNAs utilize a previously unidentified layer of complexity for their metalloregulatory functions. If it is true that a subset of the phosphorothioate interferences occur within several putative metal binding pockets (M7-9), there are multiple potential explanations for the absence of metal ions at these locations in the structural model. It is possible that they are only transiently required during RNA folding. Alternatively, their local structural environment may have been perturbed enough during crystallization to result in a loss of metal ion occupancy. In general, this part of the crystal structure was simply not as well ordered as the portions known to contain Mg^{2+}-binding sites as evidenced by the higher B-factor values observed at the M7-M9 sites as compared to the Mg1-6 positions (Figure 48). Regardless, the observation that phosphorothioate interferences cluster within these three loci provides functional evidence for their role in formation of the compact conformation of M-box RNAs.

Phosphorothioate substitutions in the presence of thiophilic metals

In general, phosphorothioate modifications disrupt Mg^{2+}-binding sites due to the fact that hard metals prefer to interact with hard ligands rather than soft ligands (*e.g.,* sulfur) (Eckstein, 1985). Thiophilic metals prefer interactions with soft ligands but are still usually capable of adopting the octahedral coordination geometry typical for Mg^{2+} ions. Therefore, it is a common technique to reassess phosphorothioate interference patterns using RNA molecules that have been incubated in the presence of more 'thiophilic' metals (Christian and Yarus, 1993). Rescue of phosphorothioate interferences by thiophilic metals is generally then taken as evidence of restoration of critical metal-RNA contact sites. However, addition of thiophilic metals can also rescue interferences at positions that are not predicted by structural analyses to associate directly with metal ions, suggesting that interpretation of thiophilic 'rescue' data may not always be straightforward (Basu and Strobel, 1999).

While the most commonly used metal for thiophilic rescue experiments is Mn^{2+}, other ions such as Co^{2+}, Cd^{2+}, Ni^{2+}, and Zn^{2+} are often used (Basu and Strobel, 1999; Weinstein et al., 1997). These thiophilic rescue experiments were first attempted with Mn^{2+} but in the native gel solution, Mn^{2+} quickly oxidizes and forms a brown precipitate that results in an undefined smear of RNA rather than distinct bands. Therefore, separation of extended and compacted M-box RNAs was performed instead in a mixture of 0.5 mM Cd^{2+} and 0.5 mM Mg^{2+} and analyzed relative to the 1.0 mM Mg^{2+} data. A single rescue of a phosphorothioate interference site was observed for G100, one of the

positions responsible for coordination of Mg1 (Figure 49). However, sulfur substitutions at the remaining two phosphate oxygens that contact Mg1 (C102 and A103) were not rescued by cadmium. In general, the inability of thiophilic metal ions to rescue sites of phosphorothioate-induced interference cannot be taken as evidence against the presence of a Mg^{2+}-binding site. It has been previously noted that phosphorothioate substitutions at established Mg^{2+} sites are not always rescued by thiophilic metal ions, possibly due to metal ion exclusion by the bulkier sulfur group (Brautigam and Steitz, 1998; Ruffner and Uhlenbeck, 1990). It has also been suggested that highly-coordinated Mg^{2+} are not easily replaced by thiophilic metals (Basu and Strobel, 1999; Brautigam and Steitz, 1998). Therefore, the fact that most of the interferences are not rescued by thiophilic metals could be a testament to the high degree to which each individual metal ion is sensed by the M-box RNA.

In some instances, the analysis of thiophilic rescue experiments can actually be complicated by the appearance of new phosphorothioate interferences upon the inclusion of thiophilic metals (Christian and Yarus, 1993). The appearance of these Cd^{2+}-induced interferences were observed at positions U47, A48, G50, A60, C78, A88, A111, A117, and A118 (Figure 50). Additionally, interferences were observed at C35, C36, A63, A80, C89, U119, and G157 when the 0.5 mM Cd^{2+}/0.5 mM Mg^{2+} data set is compared to the 1.0 mM Mg^{2+} data set but not the 0.5 mM Mg^{2+} data set.

In addition to the thiophilic rescue observed at G100, a very slight degree of rescue was observed at position A101 (Figure 51). This partial rescue decreases the strong interference of 12.8 observed at A101 to a moderate interference of 7.1. This rescue was particularly surprising since this site is not predicted to coordinate a Mg^{2+}

based on the crystallographic model of the Mg^{2+}-bound M-box RNA. This could indicate that a transiently bound Mg^{2+} is coordinated by this position. However, as previously stated, the interference at A101 appears to occur due to the disruption of critical tertiary contacts. While this type of thiophilic rescue data is uncommon, thiophilic rescue at non-metal-coordinating positions has been previously observed (Basu and Strobel, 1999).

DISCUSSION

Mg^{2+}-binding sites can be grouped into three 'core' regions

The structural model of the *B. subtilis* M-box RNA previously revealed the presence of six well-ordered Mg^{2+} (Dann et al., 2007) but could not address their relative importance for metalloregulatory control. It is not unusual to observe RNA-chelated Mg^{2+}; indeed, RNA tertiary structures are oftentimes dependent on Mg^{2+} ions that assist in bridging long-range interactions or positioning key residues for tertiary contacts (Woodson, 2005; Draper, 2008; Klein et al., 2004). Therefore, it was unclear from the M-box structural model alone whether all or a subset of these metals were intricately involved in signal-responsiveness. Moreover, other than direct sequestration of nucleotides required for genetic control it was unclear what features of M-box RNA tertiary structure are unique for their function as metalloregulatory agents. From their large size (>150 nt) and phylogenetic patterns (*e.g.,* extensive primary sequence conservation and widespread phylogenetic distribution) it can be assumed that M-box RNAs exhibit unique structural features that render them as specialized metal sensors as

opposed to simple, common structural folds. The intent in analyzing M-box RNAs was to elucidate these basic mechanistic features and correlate them with the general understanding of bacterial Mg^{2+} homeostasis.

The observation that six divalents are directly chelated to M-box RNAs suggested the possibility that intracellular responsiveness to Mg^{2+} by M-box RNAs might be exquisitely controlled via a coordinated progression of Mg^{2+} binding sites and metal-induced structural features. Although extensive solution-based experimentation is required in order to definitively make a statement about cooperativity in the M-box RNA, the Hill coefficient measured for metal-induced compaction of this RNA upon Mg^{2+} titration suggests the possibility of cooperative binding (Figure 24b). The typical riboswitch responds to a single metabolite ligand over a ~100-fold range in concentration, although the cooperative binding of multiple ligands to the glycine riboswitch remains a notable exception (Welz and Breaker, 2007). However, a sharper change in gene expression might be elicited for M-box RNAs in response to even minor Mg^{2+} fluctuations if tertiary structure formation is tightly coupled to a highly cooperative network of direct and indirectly associated Mg^{2+} ligands. Therefore, identification of metal sites that are essential for function is an important part in elucidating of basic features of M-box regulatory RNAs and their role in control of intracellular Mg^{2+}.

Reassessment of the individual roles and importance of the six divalents observed in the structural model (Mg1-6), prompted by the biochemical data presented herein, suggests that they can be grouped into two distinct, functionally-related clusters. These metal-binding ‘core’ regions were defined based upon both co-localization of Mg^{2+} sites and the degree of phosphorothioate interferences for positions within each cluster.

The current utilization of this term resembles in concept the metal binding core located within the P4-P6 domain of the *Tetrahymena* group I intron, which is important for organization of the P5abc subdomain and the subsequent folding pathway (Cate et al., 1997). These separate metal-binding cores within the M-box RNA may be working in a coordinated manner to accomplish common structural tasks. Mg1-Mg4, located within the tri-partite junction in which distal regions of the RNA (P2, L4, L5) are brought into close proximity with each other, form Core 1, while Mg5 and Mg6, located in a peripheral region of the RNA within a P4 internal loop, constitute Core 2 (Figure 52).

The proximity of Core 1 to the antiterminator nucleotides and the structural features that sequester them was the basis for the initial belief discussed in Chapter 4 that it might be more significant than Core 2 for formation of the final metal-bound, compacted conformation. Indeed, equally strong phosphorothioate interferences were observed at all but one of the positions involved in formation of the Core 1 binding sites, even at Mg^{2+} concentrations above the apparent K_D (1.0 mM Mg^{2+}). One explanation for the lack of interference at 1.0 mM Mg^{2+} observed at G157, the single exception within Core 1, is possibly due to the fact that this position is coordinating the Mg^{2+} via outer sphere coordination rather than the inner sphere coordination that is usually identified in phosphorothioate interferences. The observation that, in general, the individual interference strength is similar for each of the positions in Core 1 suggested that they are functionally intertwined (Figure 53). Consistent with this prediction, nucleotides surrounding each Mg^{2+} binding site of Core 1 (Mg1-4) appear to work together to mediate the formation of multiple long-range contacts (Figure 37). Mg4 and Mg1 bring the P4 and P5 terminal loops (L4 and L5) together by positioning A72 and A105,

respectively, which appear to form hydrogen-bonding interactions between their respective N1 and N6 functional groups. Mg1 also positions L5 residues, A101 and A103, for hydrogen-bonding interactions with U167, located within the P2 helix. Together, these interactions anchor L5 and L4 to the base of the molecule. The observation that the A101 phosphorothioate interference, disrupting some of these distal hydrogen-bonding interactions, appears to follow the same trend as the Core 1 interferences disrupting Mg^{2+}-binding sites also supports the idea that the Mg^{2+} sites of this region of the M-box are functionally connected through a network of long range interactions. These interactions are further bolstered by an A-minor motif (Nissen et al., 2001) formed between A71 and a P2 G:C base pair, an interaction that is likely to be dependent on the presence of Mg4. Also, Mg2 directly contacts phosphate oxygens from both P2 and L5 while Mg3 assists positioning of P2 residues for long-range hydrogen-bonding interactions to L4. Therefore, the Core 1 metal sites appear to be crucially linked to one another and the final folded architecture of the metal-bound M-box RNA.

Just as Core 1 metal-binding residues shared a similar overall degree of phosphorothioate interferences, so to do Core 2 positions (Figure 53). Due to the peripheral location of Core 2, this region was not expected to be as essential as Core 1. Indeed, Core 2 interferences were all classified as weak interferences for 0.5 mM Mg^{2+} and were not observed at 1.0 mM Mg^{2+}, suggesting that their functional role(s) may be partially supplanted during conditions of metal excess. Under 0.5 mM Mg^{2+} conditions, the A63 contact to Mg5 just meets the criteria for a moderate interference, while the A80 contact to Mg6, the most tangentially associated Mg^{2+}, appears as a weak though reproducible interference. However, these Mg^{2+} sites are both located within an internal

loop within the P4 helix where they appear to stabilize base stacking interactions between residues that stack against the top and bottom half of the P4 helix, respectively (Figure 52c). A modest kink is introduced between these helices which in turn allows for proper orientation of the P4 terminal loop for long-range interactions to L5 and P2. All M-box RNAs that have been identified (Griffiths-Jones et al., 2005; Barrick et al., 2004) contain the A-rich P4 internal loop that constitutes Core 2, bolstering the presumption that this structural element and the Mg^{2+} that promote it are essential for M-box metalloregulatory function.

The additional sites of phosphorothioate-induced interferences that could not be explained by the existing structural model did not appear to fit into either of these two Mg^{2+} cores. Instead, they are all located near the apex of the M-box structure in a cluster that may constitute a third metal-binding core. More specifically, these phosphorothioate interferences cluster together into three separate foci that may suggest the presence of three new metal binding pockets. As with Cores 1 and 2, the interference values within the putative Core 3 share a common intensity, suggesting that individual metal sites may be working in conjunction with one another (Figure 53). Correspondingly, analysis of the Core 3 putative metal binding sites revealed that, as in Core 1, surrounding nucleotides appear to be positioned for formation of long-range base contacts. While Core 1 functions to stabilize the tripartite interaction between the three parallel P2, P4, and P5 helices at the base of the molecule, Core 3 would be predicted to stabilize the tripartite interaction at the opposite end of the molecule (Figure 52d). Each of the individual putative metal sites (M7-9) within Core 3 appears to individually promote formation of essential long-range contacts. For example, the putative M8 binding site that is formed in part by C89

and G91 appears to assist in positioning the adjacent nucleotides (A87 and A88) such that they splay out in opposite directions, thereby allowing long-range hydrogen bonding interactions with nucleotides that participate in formation of the other two putative metal sites, M7 and M9. The adenine nucleobase of A117, one of the nucleotides whose phosphate oxygen participates in formation of the M7 site, forms a base-stacking interaction with A87 while simultaneously forming an A-minor motif interaction with a P4 G:C base pair, thereby linking P4 and P5. A88 forms base-stacking interactions with its adjacent nucleotides C89 and A90 as well as C35, a bulged nucleotide within the P2 helix whose phosphate oxygen participates with formation of the putative M9 site in conjunction with C36. This contact forms one link between the P2 and the P5 helices while yet another link is created by the formation of an A-minor motif in which A88 docks into a base pair of the P2 helix. Indeed, all four A-minor motif interactions and most other long-range interactions observed in the structural model appear to be dependent upon metal binding sites in Core 1 and the putative Core 3 (Figure 52a). Therefore, these data suggest that the network of three-helical, long-range contacts required for stabilization of the compact RNA conformation may be mediated primarily by metal binding sites at the base of the molecule (Core 1) and at the top of the molecule (Core 3). Presumably, it is the coordinated action of these structural features that has in part engendered the basic physical properties of the M-box tertiary fold to allow it to function as a dedicated metal sensor in bacteria.

As a preliminary test of the importance of Core 3, an A87 to C mutant was generated. This position was shown in the crystallographic model to make contacts mediated by bends in the phosphate backbone that are possibly induced by the putative

metal sites of Core 3. The nucleobase of A87 is splayed out by the putative M8 site to form base stacking interactions with A117 positioned by the putative M7 site. Using velocity sedimentation by AUC, this construct was analyzed for its ability to adopt the Mg^{2+}-induced compacted state (Figure 54). This mutant is severely inhibited in its ability to undergo Mg^{2+}-induced compaction confirming that the tertiary interactions of Core 3 are essential to the final folded state of the M-box RNA.

While it is tempting to designate the electronegative pockets that contain sites of phosphorothioate interferences as putative metal binding sites, alternative explanations must also be considered. For example, phosphorothioate substitutions can perturb formation of individual hydrogen bonding interactions (Hinrichs et al., 1987; Christian and Yarus, 1992). While this outcome of phosphorothioate substitutions is likely to be responsible for interferences observed at A101 and A105, the structural model does not reveal any obvious interactions that could be disrupted by sulfur substitutions at C35, C36, C89, G91, or U119. However, the possibility cannot be ruled out that these substitutions affect hydrogen bonds that are specifically important during the RNA folding process but are not present in the final native structure. It is also important to note that the charge distribution around the phosphate group is altered upon phosphorothioate substitution (Frey and Sammons, 1985). A greater proportion of the negative charge becomes distributed towards the remaining nonbridging oxygen for phosphorothioate-containing linkages. This could affect folding pathways by altering local electrostatic interactions between backbone phosphates. Finally, the ionic radius of sulfur is larger than that of oxygen and might lead to deleterious steric effects. Further biochemical and structural analyses will be required to fully resolve the role of the putative Core 3

residues during metal-induced folding of M-box RNAs. Nonetheless, it is clear from the data presented herein that backbone oxygens within this region are equally important to the metal binding sites observed in the structural model and within Cores 1 and 2.

Preliminary crystallographic analyses performed by Dr. Arati Ramesh indicate that M7 and M9 are indeed metal-binding sites. She has crystallized the M-box RNA in the presence of Mn^{2+} which is capable of adopting the same coordination geometry as Mg^{2+} but is actually known to bind with higher affinity to the M-box RNA (Figure 27) as well as many other RNAs. Therefore, the metal-binding sites with low occupancy in Mg^{2+} might have higher occupancy with Mn^{2+} and be more easily visualized. Another benefit of using Mn^{2+} for crystallization is the anomalous signal produced by Mn^{2+} which allows Mn^{2+}-occupied positions to be identified. In the Mn^{2+}-bound crystals, M7 and M9 are clearly occupied by Mn^{2+}. However, M8 remains unoccupied, possibly due to the fact that this site is only transiently occupied during folding.

Future directions

While these data are highly reproducible and appear fairly conclusive, there are other methods that can be employed to enumerate the Mg^{2+} associated to the M-box RNA. Another commonly used solution-based technique utilizes a fluorescent dye known as 8-hydroxylquinoline (HQS) (e.g. Das et al., 2005a). Additionally, the NAIM data would benefit from the use of alternative nucleotide analogs to test the importance of other RNA functional groups. Deoxyribonucleotide phosphorothioate analogs would be ideal for the analysis of the putative M8 metal-binding site since two of the Mg^{2+}-

coordinating ligands in this region are predicted to be 2′ hydroxyl groups. Finally, phosphorothioate substitution experiments should also be performed using the *B. cereus* M-box RNA to confirm that the same binding sites predicted for the *B. subtilis* M-box are also found in this RNA with the exception of Mg4, the binding site that should not be occupied by Mg^{2+} when the L4 tetraloop sequence is GNRA rather than CNRA due to steric hindrance by the bulkier purine nucleobase (see Chapter 4).

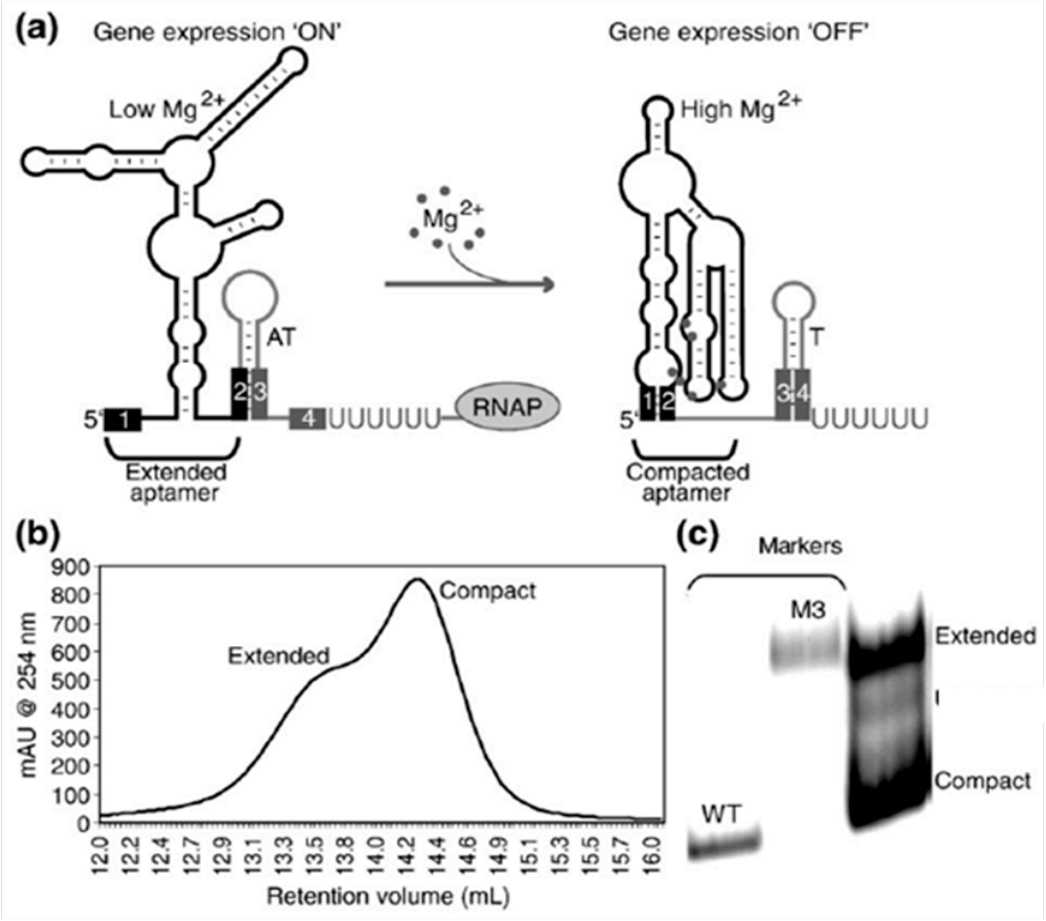


Figure 41: Separation of extended and compact M-box RNAs. **(a)** Schematic showing the transition from extended to compact conformations for the M-box. Nucleotide stretches labeled 1–4 are regions that form base-pairing interactions such that downstream expression is 'on' in the unbound state and 'off' in the magnesium-bound state. Control of gene expression is achieved through mutually exclusive formation of either an intrinsic transcription terminator helix ("T") or an antiterminator helix ("AT"). The portion of the RNA colored black represents the aptamer domain that is responsible for ligand binding. RNAP = RNA polymerase. **(b)** Representative elution profile for size-exclusion chromatography in buffer containing 0.5 mM magnesium for M-box RNAs randomly modified at adenosines with phosphorothioate substitutions. **(c)** Non denaturing gel electrophoresis of extended and compacted M-box species in the presence of 0.5 mM magnesium. WT and M3 mark the compact and extended states, respectively.

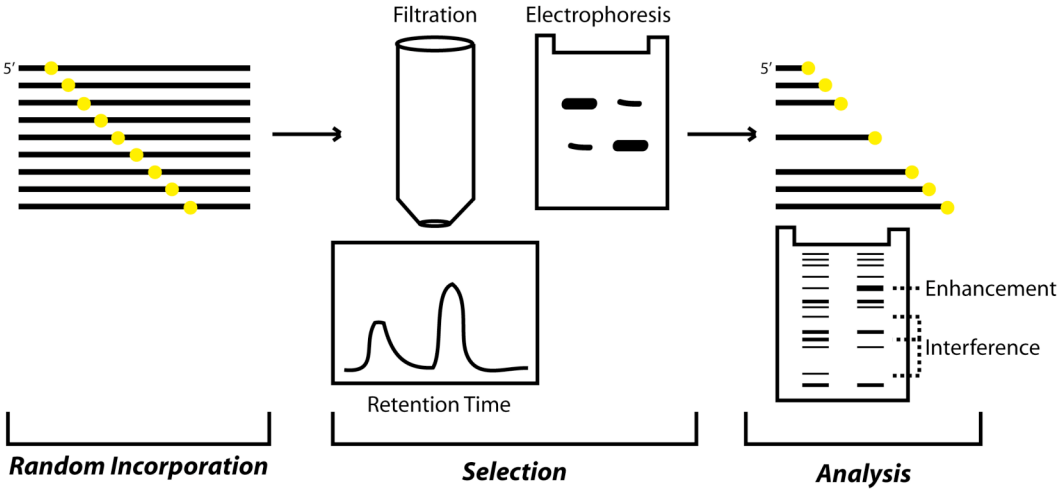


Figure 42: Schematic of NAIM analysis. Sites of nucleotide analog incorporation are depicted in yellow. Random Incorporation: Nucleotides analogs are randomly incorporated by T7 RNA polymerase during transcription. Selection: These molecules are subjected to selection based on their ability to fold. Analysis: Cleavage at sites of analog incorporation is induced by iodine treatment and selected RNA pools are compared to unselected RNA pools on when resolved by PAGE.

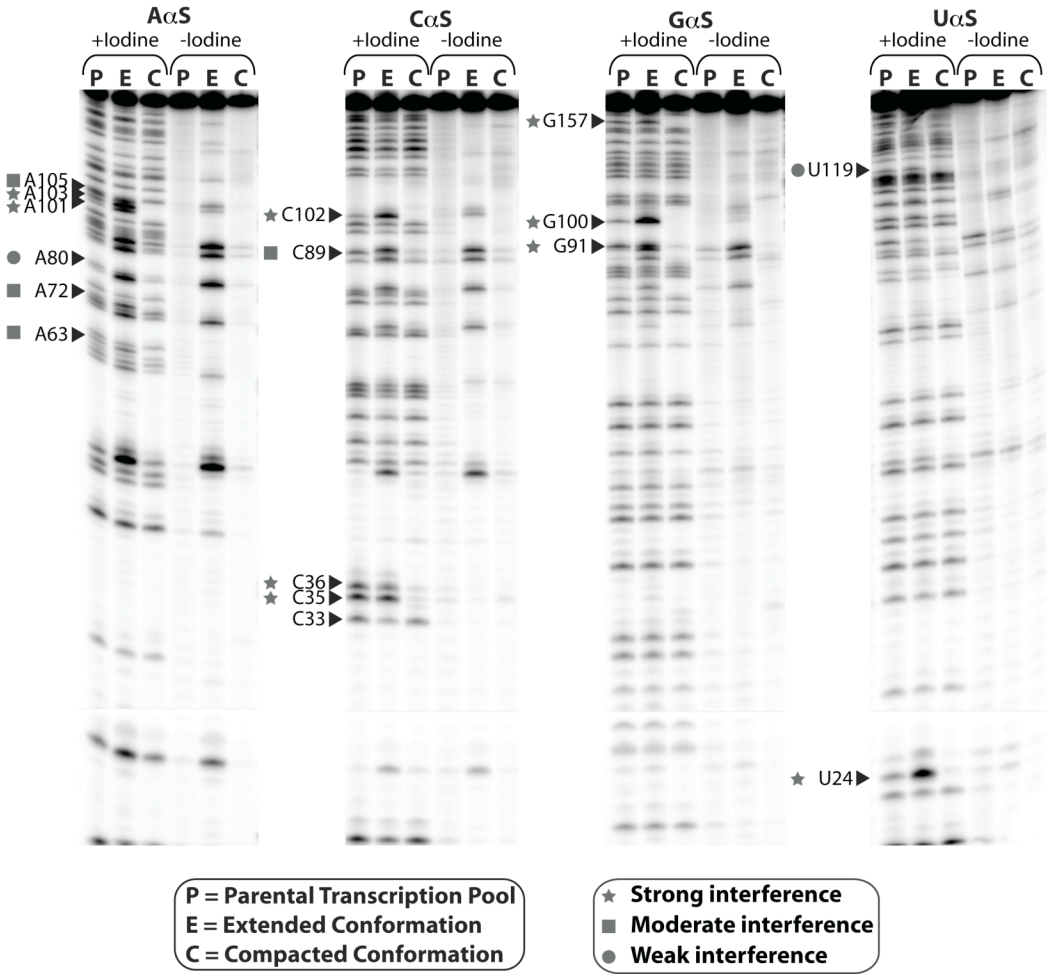


Figure 43: Representative data for iodine-mediated cleavage ladders of RNA molecules resolved in 0.5 mM magnesium. Each site of phosphorothioate interference is labeled and the strength of the interference is denoted by the symbol adjacent to the nucleotide label. P = parental RNA pool, E = extended RNA pool, and C = compacted RNA pool.

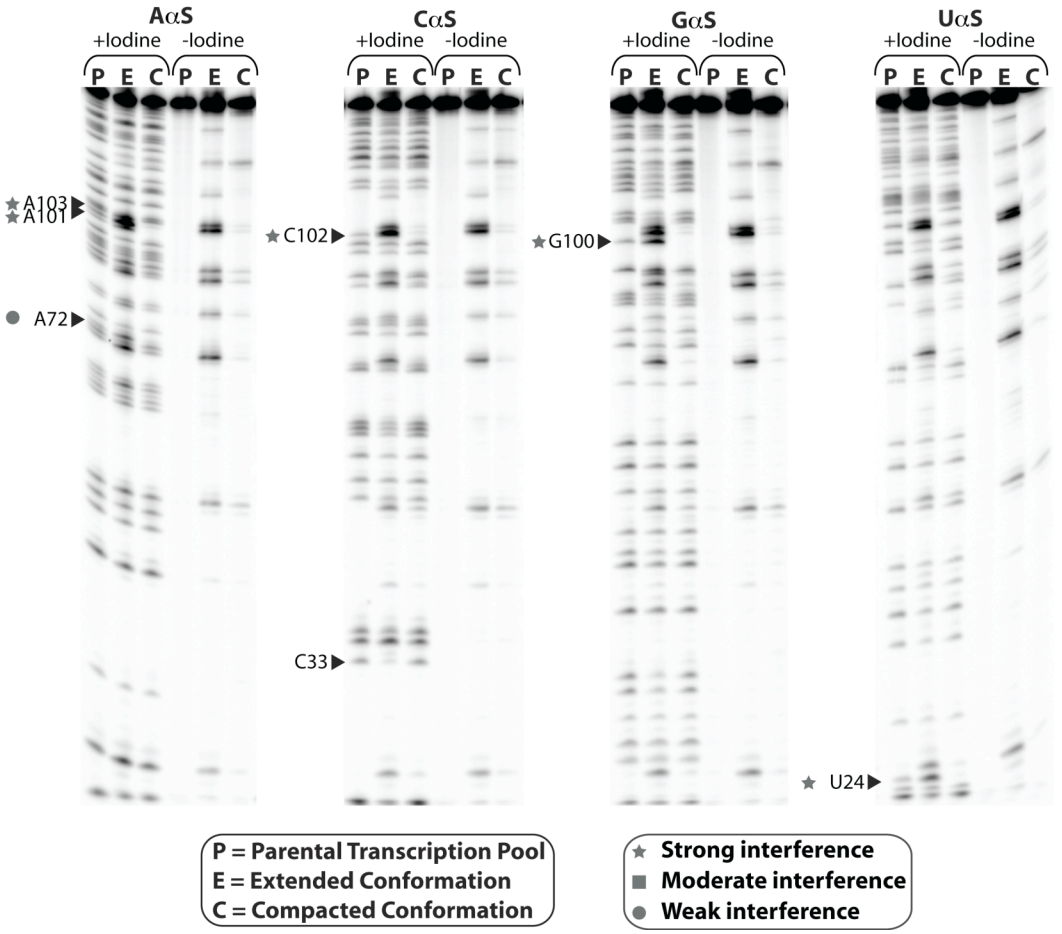


Figure 44: Representative data for iodine-mediated cleavage ladders of RNA molecules resolved in 1.0 mM magnesium. Each site of phosphorothioate interference is labeled and the strength of the interference is denoted by the symbol adjacent to the nucleotide label. P = parental RNA pool, E = extended RNA pool, and C = compacted RNA pool.

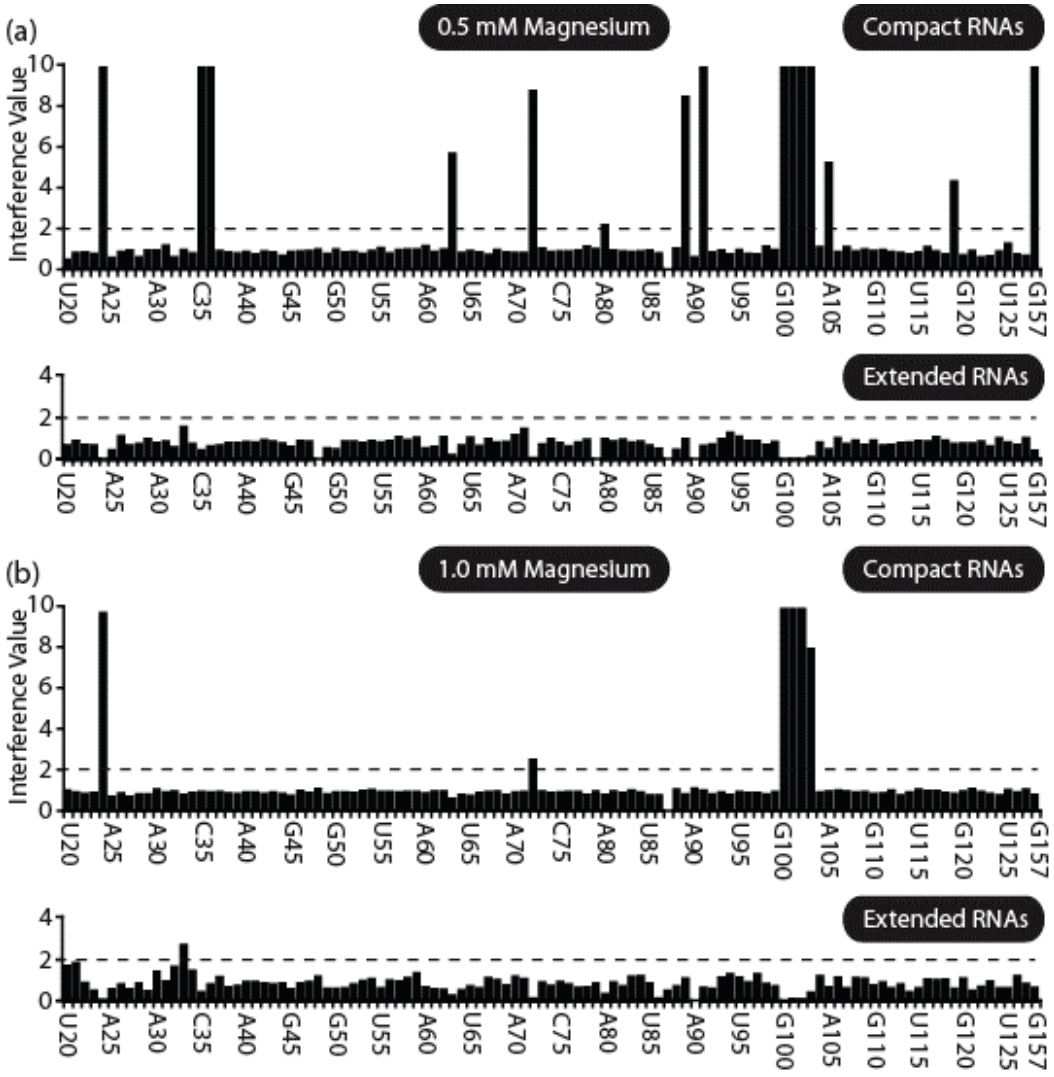


Figure 45: Quantification of interference values for extended and compact M-box RNAs separated under **(a)** 0.5 mM and **(b)** 1.0 mM magnesium conditions. The dashed line indicates an arbitrary cutoff for interference values greater than 2. Maximum interference value shown is cut off at 10 even for positions with greater interference values.

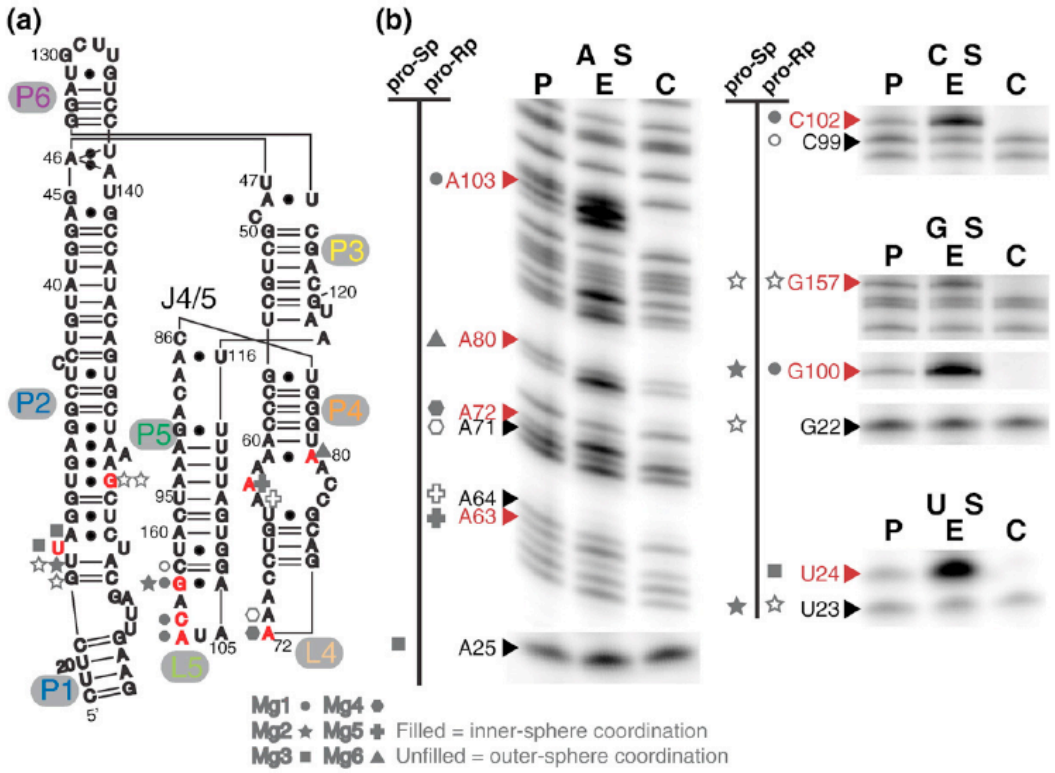


Figure 46: Sites of phosphorothioate interferences at established magnesium-binding sites. **(a)** Secondary-structure diagram for the *B. subtilis mgtE* M-box RNA. Positions identified in the structural model as coordinating to Mg^{2+} (Mg1–Mg6) via phosphate oxygens are denoted by different symbols. The different helical elements and terminal loops are indicated with colored labels according to a color scheme utilized by the other figures in this article. Red letters denote positions of phosphorothioate interferences. **(b)** Representative data showing sites of phosphorothioate interferences for RNA molecules in 0.5 mM magnesium. P, parental, unselected transcription reaction; E, extended RNA molecules; C, compact RNAs. In most cases, individual phosphorothioate interferences for compact RNAs correlated with phosphorothioate enhancements for extended RNAs relative to the parental transcription pool. Symbols corresponding to Mg1–Mg6 are grouped into columns based on whether the metal coordinates to either the pro-S_p or the pro-R_p oxygen in the structural model.

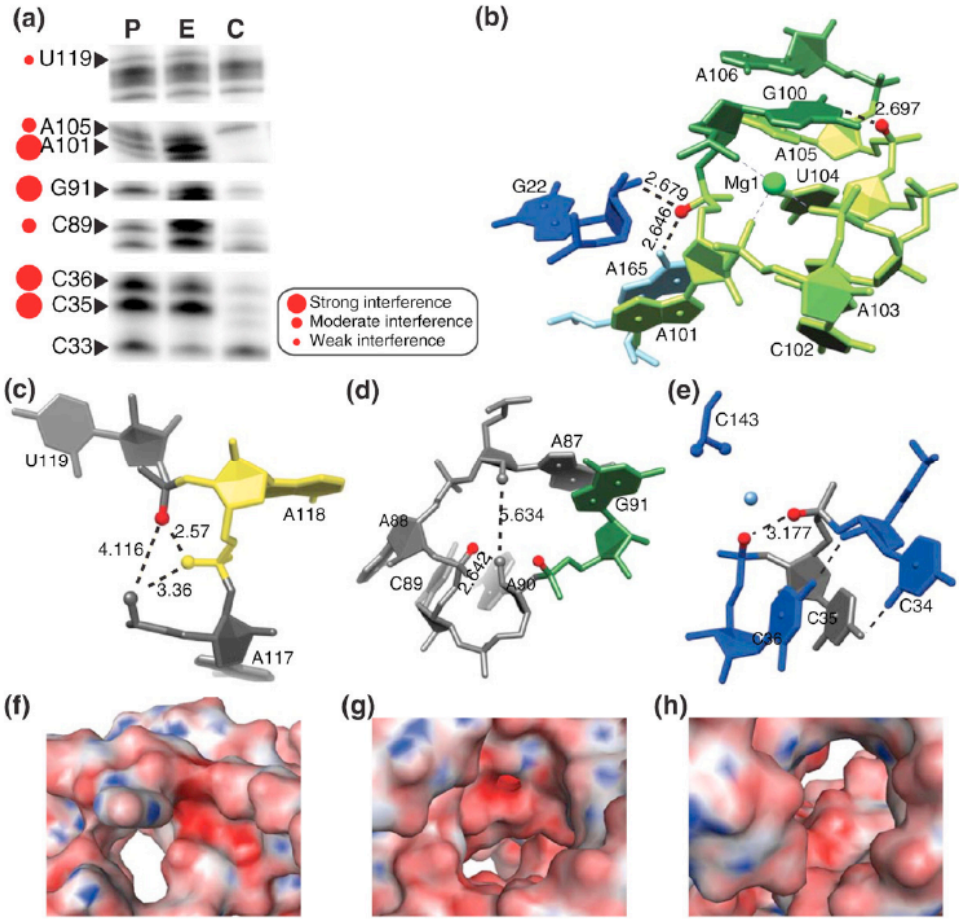


Figure 47: Phosphorothioate interferences at positions not previously identified as Mg^{2+}-binding sites. **(a)** Data showing sites of interferences for RNA molecules in 0.5 mM magnesium. P, parental reaction; E, extended; C, compact. Strong interferences are marked by large circles, moderate interferences by medium-sized circles, and weak interferences by small circles. **(b)** Phosphorothioate modifications (red spheres) at A101 and A105 are to likely interfere with formation of structural features required for the Mg1 binding site and interactions between L5 and P2. Black dashed lines denote putative hydrogen-bonding disrupted by phosphorothioate substitutions. The green sphere represents Mg1. **(c)** Structural context for the interference at the pro-R_p of U119. The pro-S_p phosphate oxygens of A117 and A118 and the U119 pro-R_p oxygen appear to form a candidate metal-binding site (M7). **(d)** Structural context for the interferences at the pro-R_p of C89 and G91, which may participate in formation of a metal-binding site (M8). **(e)** Structural context for interferences at the pro-R_p of C35 and C36, which may participate in formation of a metal-binding site (M9). Phosphate oxygens of C143 also contribute to electronegativity near the C35 and C36 backbone. **(f)–(h)** Surface depictions of electrostatic potential calculations for the regions depicted in **(c)–(e)**, respectively.

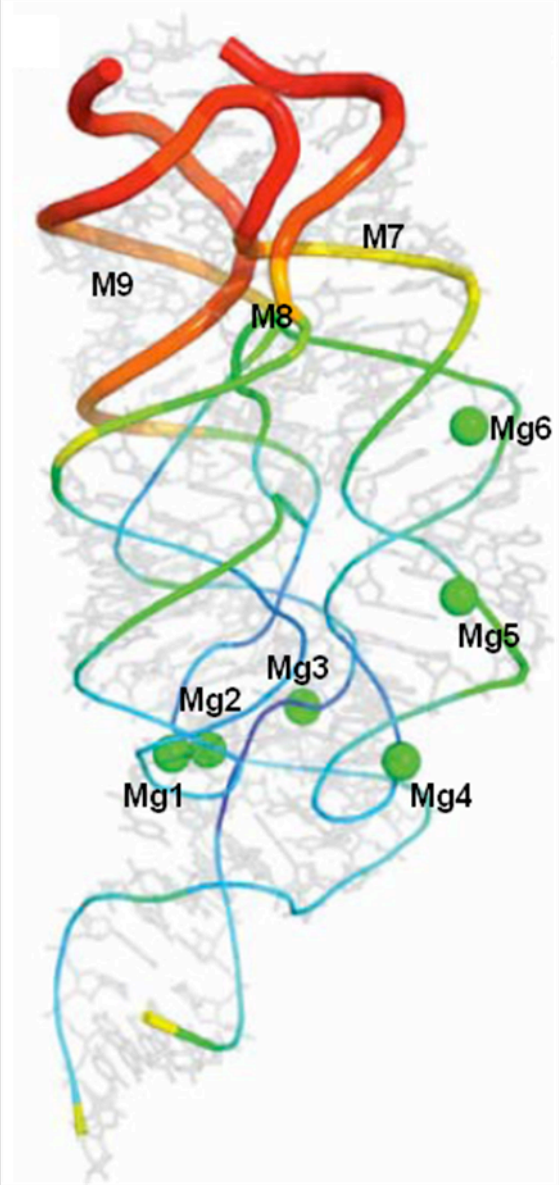


Figure 48: The backbone of the RNA structure is shown in varying color and varying thickness in order to represent B-factor values from the model. The higher the B-factor, the more disorder observed for that particular region of the RNA structure. Approximate B-factor ranges are indicated as follows. Purple = <20. Dark blue = 21-25. Light blue = 26-35. Green = 36-45. Yellow = 46-55. Gold = 56-65. Orange = 66-75. Red = >76. Mg1-Mg6 are shown as green spheres. M7-M9 are also marked.

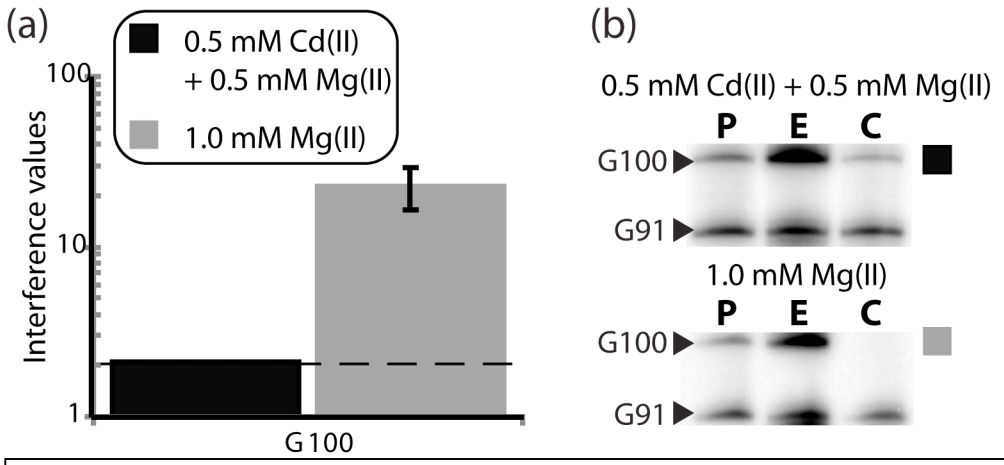


Figure 49: A single site of thiophilic metal rescue was observed at G100 when RNA molecules were resolved in the presence of cadmium. **(a)** Quantification of the partial rescue of the G100 phosphorothioate induced interference. When RNA molecules were separated in conditions containing 0.5 mM cadmium and 0.5 mM magnesium (displayed as a black bar), the interference value at G100 just met the criteria for a weak interference (the minimum value considered a significant interference is marked by the dotted line). When the same pool of RNA molecules was separated under conditions containing 1.0 mM magnesium (displayed as a gray bar), a strong interference was observed for this position (interference value greater than 10). **(b)** Example gel slices highlighting the bands used for this quantification. The lane containing compacted RNAs separated under thiophilic metal conditions has a band at the G100 position marked by a black box, while the lane containing compacted RNAs separated under 1.0 mM magnesium conditions displays the absence of a band at the G100 position marked by a gray box. P = parental RNA pool, E = extended RNA pool, and C = compacted RNA pool.

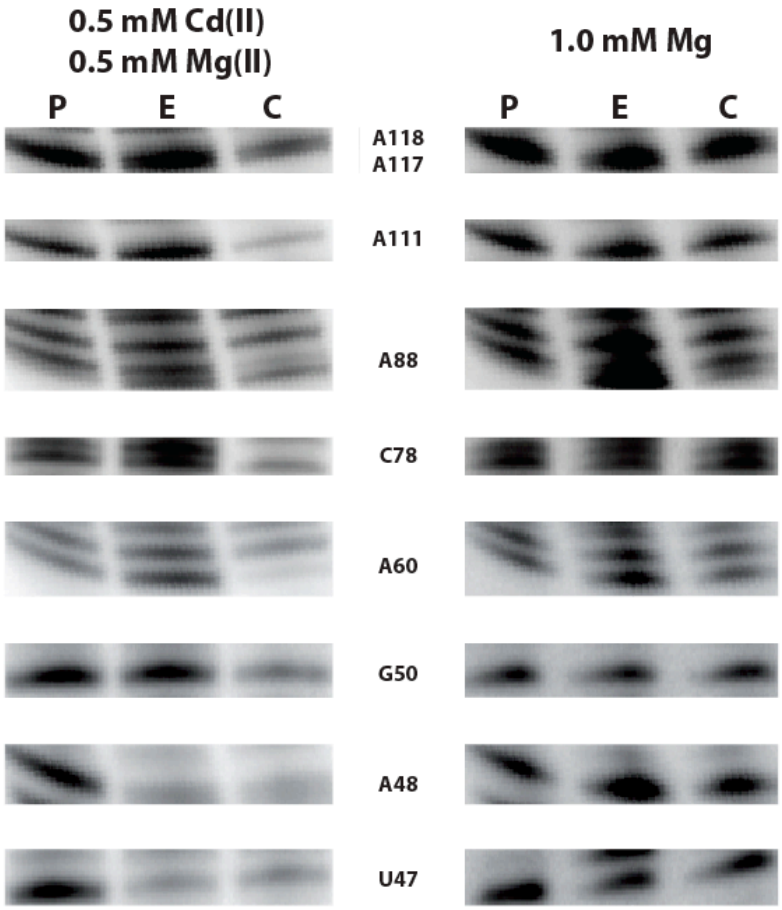


Figure 50: Cd^{2+}-induced phosphorothioate interferences. All positions highlighted exhibit phosphorothioate interference in the compacted conformation only when separated under conditions containing 0.5 mM Cd^{2+} and 0.5 mM Mg^{2+}. Additionally, U47 and A48 exhibit interferences in the extended conformation.

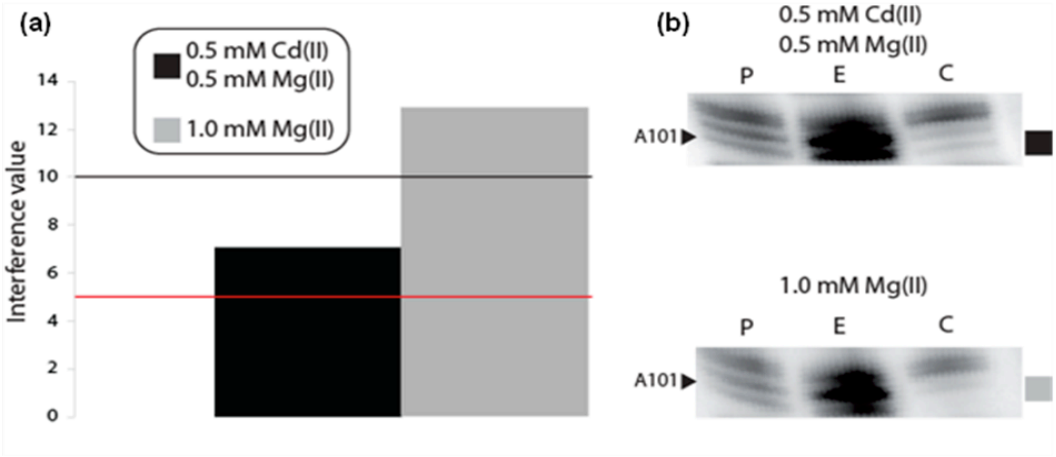


Figure 51: An additional site of partial thiophilic metal rescue was observed at A101 when RNA molecules were resolved in the presence of cadmium. **(a)** Quantification of the partial rescue of the A101 phosphorothioate induced interference. When RNA molecules were separated in conditions containing 0.5 mM cadmium and 0.5 mM magnesium (displayed as a black bar), the interference value at A101 met the criteria for a moderate interference marked as a red line. When the same pool of RNA molecules was separated under conditions containing 1.0 mM magnesium (displayed as a gray bar), a strong interference was observed for this position (interference value greater than 10, marked as black line). **(b)** Example gel slices highlighting the bands used for this quantification. The lane containing compacted RNAs separated under thiophilic metal conditions has a band at the A101 position marked by a black box, while the lane containing compacted RNAs separated under 1.0 mM magnesium conditions displays the absence of a band at the A101 position marked by a gray box. P = parental RNA pool, E = extended RNA pool, and C = compacted RNA pool.

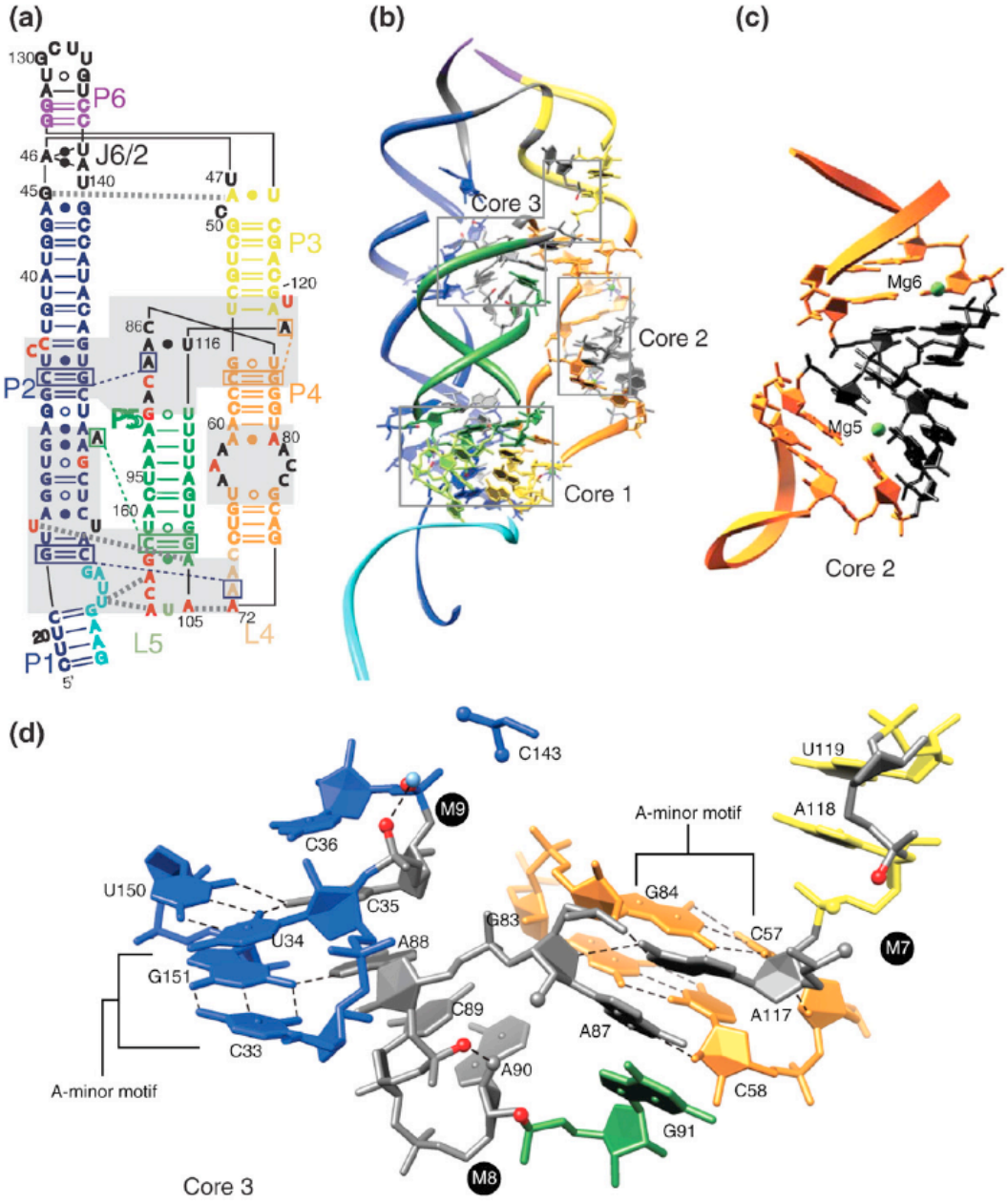

(a)
P6
J6/2
P3
P2
P4
P5
L5
L4
P1
(b)
Core 3
Core 2
Core 1
(c)
Mg6
Mg5
Core 2
(d)
C143
M9
C36
U150
C35
U34
A88
G151
C33
A-minor motif
Core 3
A-minor motif
G84
G83
C57
M7
U119
A118
A117
A87
C89
A90
C58
G91
M8

Figure 52: Metal-binding sites can be grouped into three separate cores. **(a)** Secondary-structure diagram showing long range base interactions (dashed lines), A-minor motif interactions (colored boxes connected by dashed lines), sites of phosphorothioate interferences (red letters), and magnesium cores (gray shaded boxes). **(b)** Residues contacting metals or involved in long-range base interactions are shown within the structural model. These residues cluster within the three magnesium cores postulated herein. **(c)** The Mg^{2+} of Core 2 assist in orienting bases to stack on either the top or the bottom half of the P4 helix while introducing a kinked orientation of the P4 terminal loop (L4). This orientation is necessary for formation of tertiary contacts between L4, L5, and P2. Observation of phosphorothioate interferences at the Mg5 and Mg6 sites indicates that Core 2 is functionally important. **(d)** Residues within Core 3. Specific hydrogen-bonding interactions are highlighted with dashed lines. Key A-minor motifs are labeled accordingly. Regions corresponding to the newly identified candidate metal sites (M7–M9) are indicated. These sites appear to assist formation of key tertiary interactions, interhelical bonds, and stacking interactions.

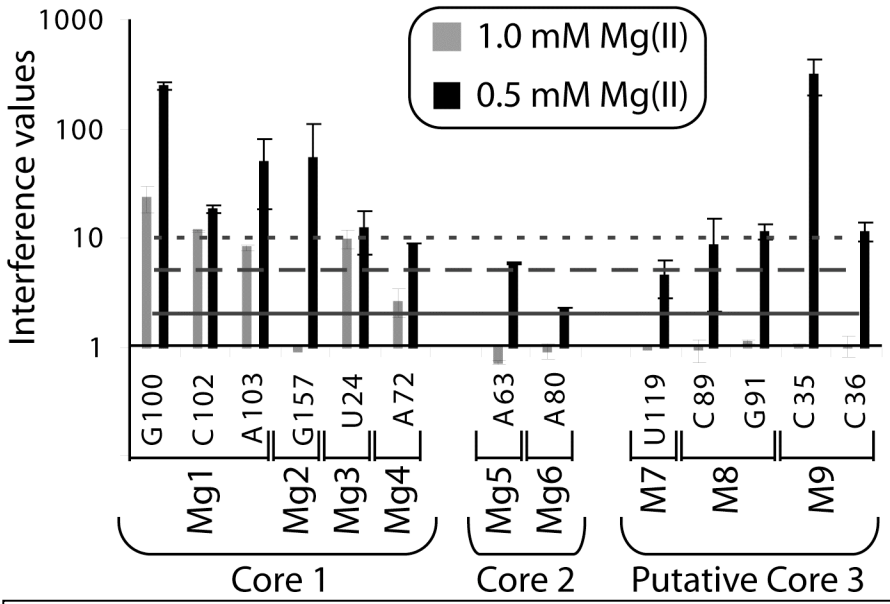


Figure 53: Metal binding sites can be grouped into three cores. A plot of interference values versus position suggests that the degree of phosphorothioate interferences correlates within three separate cores.

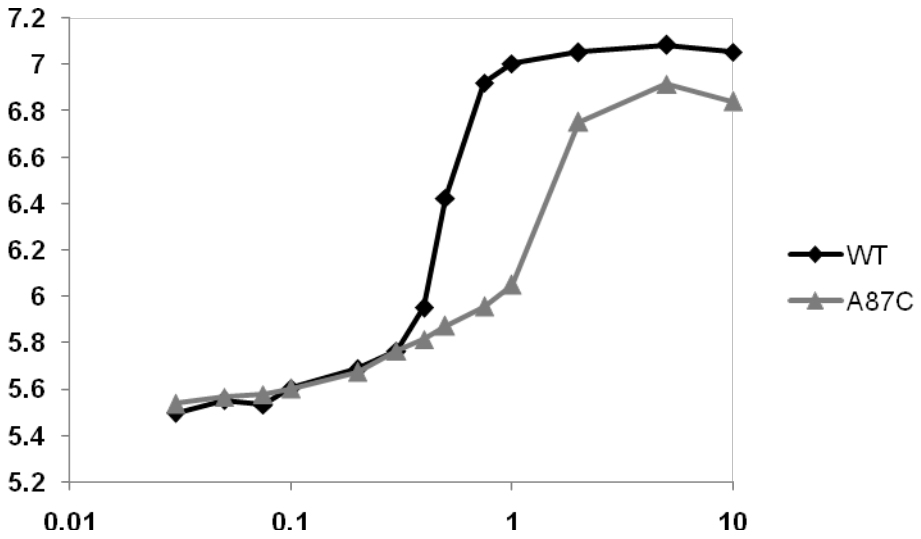


Figure 54: Mutational analysis to determine the importance of tertiary contacts formed by the putative Core 3. The nucleobase of A87 is splayed out by the putative M8 site to form base stacking interactions with A117 positioned by the putative M7 site. When this position is mutated to a cytidine, the EC_{50} for the Mg^{2+}-induced transition is significantly shifted and this construct only reaches a sedimentation coefficient of approximately 6.8 S at saturating Mg^{2+} concentrations as compared to the 7.0 S attained by the WT construct.

CHAPTER SIX

GENERAL DISCUSSION, BROAD IMPLICATIONS AND FUTURE DIRECTIONS

Comparing and contrasting the M-box RNA to unrelated riboswitches

Since all RNAs require the presence of cations to promote folding, the complex architecture of riboswitch RNAs necessitates an intimate relationship with positively charged metal ions. In addition to this general requirement, some riboswitch RNAs appear to have evolved more specialized uses for metal ions and Mg^{2+} in particular. Mg^{2+} assists in the stabilization of ligand binding for the TPP-, GlcN6P-, and FMN-binding riboswitches (Serganov et al., 2006; Cochrane et al., 2007; Serganov et al., 2009). Within the phosphate-rich environment formed by these and all RNA structures, Mg^{2+} helps to neutralize the negative charge provided by the phosphate groups found on the TPP, GlcN6P, and FMN molecules. It is likely that other RNA aptamers that bind ligands possessing negatively charged functional groups will utilize a similar mechanism.

Unlike these metabolite-binding RNAs, the M-box RNA element does not require Mg^{2+} for stabilization of ligand binding. Instead Mg^{2+} itself acts as the ligand. In reality, the role of the Mg^{2+} ions in the M-box RNA can be likened to the general contribution to folding observed for all structured RNAs. However, unlike these other regulatory RNAs, the M-box translates the requirement of cations for folding into a sensory function by meeting two criteria. First, the folded structure adopted in the

presence of Mg^{2+} governs the accessibility of a short nucleotide tract that results in regulatory function by controlling the formation of an intrinsic terminator helix. Second, the EC_{50} for this metal-induced structural transition is tuned to an appropriate physiological intracellular concentration.

Recently, studies have been performed to determine the effect of Mg^{2+} on the structures of other riboswitches. Many of these studies utilize a method known as small angle X-ray scattering (SAXS) to reveal the global architecture of the RNA molecule in a manner similar to that of a crystallographic structure but at significantly lower resolution and in solution rather than in a crystalline state (Garst et al., 2008; Lipfert et al., 2007). Unsurprisingly, these studies demonstrate that the presence of Mg^{2+} has a significant impact on the global structure of each riboswitch RNA that was tested. SAXS data demonstrate that the glycine-sensing riboswitch becomes compacted in the presence of Mg^{2+} and undergoes further compaction upon association with glycine (Lipfert et al., 2007). In contrast, the lysine-sensing riboswitch becomes dramatically compacted in the presence of Mg^{2+} but does not appear to undergo further compaction in the presence of its ligand (Garst et al., 2008). This was a very shocking discovery since it was assumed for all riboswitch RNAs that ligand-binding resulted in significant conformational changes within the aptamer domain in order to elicit conformational changes within the expression platform. The observation that some riboswitch aptamers appear to be fully compacted upon association with Mg^{2+} alone begs the question "Are these other riboswitch RNAs acting as metal-sensors?" The data presented herein would indicate otherwise.

The lysine-sensing riboswitch and other riboswitches capable of associating to small organic ligands are not likely to be functioning as metal-sensors first and foremost due to the fact that they are never found to regulate genes involved in Mg^{2+} transport. Additionally, the lysine riboswitch has proved to be capable of responding to lysine even when bacterial cells are grown under Mg^{2+}-replete conditions (Sudarsan et al., 2003), indicating that the Mg^{2+}-associated structure is not adopting the 'terminated' structure. Instead, the Mg^{2+}-bound state is likely forming a pre-organized binding pocket ready to accept the ligand in a manner similar to but even more organized than the purine riboswitch (Stoddard et al., 2008). Upon binding of lysine, some structural rearrangement must occur within the aptamer to allow for the switch from 'antiterminated' to 'terminated' conformation only upon ligand binding. In-line probing analyses of this RNA support the idea that the bulk of the RNA is pre-organized even in the absence of ligand; however, significant structural rearrangements seem to occur within the region of the P1 helix upon association with lysine (Sudarsan et al., 2003). The conformational switch within the P1 helix is the change that matters most for sensory function since this helix contains the nucleotide stretch responsible for relaying the signal of ligand-binding to the expression platform. Ultimately, while this and many other classes of riboswitch RNAs are Mg^{2+} dependent, they are not Mg^{2+}-sensors because a functional output is only obtained in the presence of the ligand.

In order to demonstrate that the M-box aptamer is unique in its Mg^{2+}-responsive capabilities as compared to other well characterized riboswitch RNAs, several lines of experimentation have been performed. For example, a SAM-binding riboswitch from *B. subtilis* has been shown to terminate transcription to the exact same degree in Mg^{2+}

concentrations ranging from 1.5 mM to 7 mM. In contrast, the M-box RNA displays significant changes in its degree of transcription termination over the same concentration range (Figure 29). It is important to perform this assay using other RNAs such as the lysine riboswitch in order to definitively demonstrate that the globally compacted structure observed for the RNA in the presence of Mg^{2+} is not the functionally active compacted structure required to produce changes within the expression platform.

In addition, structural probing assays were performed wherein Mg^{2+} was titrated into reactions containing either SAM-, TPP-, or FMN-binding riboswitches in the presence of high levels of monovalent ions. As a result, these three RNA structures did not appear to change conformation (Figure 18). This is in stark contrast to the M-box RNA. This comparison demonstrates that the M-box has a special relationship with Mg^{2+} unlike these other RNAs. It is worth noting that 2.1 M monovalent conditions are not naturally encountered by any of these RNAs inside the cell. Therefore, the difference between the M-box and other riboswitches must be demonstrated under more physiological conditions. An extensive comparison of Mg^{2+} titrations with many characterized riboswitch RNAs should reveal that the transition into the Mg^{2+}-induced compacted state for non-metalloregulatory RNAs occurs at a lower EC_{50} and/or over a greater concentration range as compared to the M-box in order to prevent all riboswitches from being affected by Mg^{2+} depletion.

While the mechanism by which the M-box RNA is set to the correct EC_{50} for Mg^{2+}-sensing as compared to most other naturally occurring structured RNAs has not been demonstrated, speculations can be made. The large number of Mg^{2+} ions coordinated by multiple inner sphere contacts found within the M-box as compared to

any other known RNA of its size surely plays a role. It is estimated that more energy is required to displace a water molecule in order to make an inner sphere contact to Mg^{2+} than is necessary for contacting a hydrated water molecule via outer sphere coordination (Draper et al., 2005). It follows then that the EC_{50} for Mg^{2+}-induced compaction of RNAs forming multiple inner sphere contacts might be slightly higher than that of other molecules. Therefore, the M-box would likely be more sensitive to subtle changes in the levels of intracellular Mg^{2+} as compared to most other structured RNAs. Furthermore, the Mg^{2+}-induced conformational change within the M-box RNA appears to occur over a much narrower range than that of other non-metalloregulatory riboswitch RNAs (Draper, 2008).

It has already been demonstrated that in vivo Mg^{2+} depletion does not result in increased expression of either SAM-riboswitch or *glmS* ribozyme regulated genes (Figure 8e). However, the expression of these genes should be increased if the level of intracellular Mg^{2+} was low enough to disrupt the formation of the compacted riboswitch aptamers in general. Therefore, while these RNA structures are stabilized by Mg^{2+}, the levels of Mg^{2+}-depletion required to elicit an effect within the M-box RNA are not low enough to impact these riboswitches. Similar studies with other unrelated riboswitch classes such as the lysine-binding riboswitch have not yet been performed but will likely reveal the same trend.

Possible connections between Mg^{2+} homeostasis and virulence

The identification of a Mg^{2+}-sensing RNA will open many doors into the study of Mg^{2+} homeostasis. Every single gene regulated by this RNA should, in theory, have a connection to intracellular and extracellular Mg^{2+} levels. It is perfectly logical that so many of these RNAs regulate genes involved in Mg^{2+} transport. However, it is much more difficult to understand why a Mg^{2+} sensor would regulate cell division proteins, PE/PPE proteins, and other proteins of unknown function (Table 1).

PE/PPE proteins are abundant within the genomes of many *Mycobacterium* spp. and are secreted by type VII secretion systems (Abdallah et al., 2009). Many of these proteins are predicted to be involved in virulence. Connections between bacterial virulence and homeostasis of Mn^{2+}, Fe^{2+}, and Zn^{2+} are well known (Jakubovics and Valentine, 2009). Moreover, the Mg^{2+}-responsive PhoP/Q two-component system regulates the expression of several virulence genes in enteric bacteria (Bijlsma and Groisman, 2005). Therefore, it makes sense that the M-box RNA might be sensing Mg^{2+}-depletion as a signal of entry into the host environment in order to up-regulate virulence factors. This RNA element is predicted to regulate proteins other than Mg^{2+} transporters in a number of pathogens such as *M. tuberculosis* and *L. monocytogenes* (Table 1).

In *L. monocytogenes*, the M-box RNA element appears to regulate the expression an operon containing an MgtA Mg^{2+} transporter and cell division-related proteins. This observation is particularly interesting in light of the cell division defect observed in *B. subtilis* grown under Mg^{2+}-depleted conditions (Figure 7). More importantly, it has been noted that *L. monocytogenes* undergoes differential rates of cell division depending on

their location within the host cell (Birmingham et al., 2008). In the cytoplasm, these bacteria rapidly replicate but when trapped within the phagosome, they undergo significantly slowed rates of cell division. It has also been shown that there is a gradient of Ca^{2+} between the *Listeria*-containing vacuole and the cytoplasm (Shaughnessy et al., 2006). Therefore there is likely to be a gradient of other metal ions such as Mg^{2+} within different compartments of eukaryotic cells. Whether or not these facts relate to the regulation of cell division proteins by a Mg^{2+}-sensing RNA remains to be demonstrated. However, it would be interesting to determine whether or not these bacteria are capable of sensing not only entry into a host cell but also entry into a specific niche within the host cell via a metalloregulatory RNA.

A comparison of transcript abundance of M-box-regulated genes in different compartments of a host cell would be an exciting study. There is debate as to whether or not the *S. enterica*-containing vacuole represents a Mg^{2+}-depleted environment despite the fact that a Mg^{2+}-sensing two-component system up-regulates virulence gene expression within this environment (Alix and Blanc-Potard, 2007). Perhaps the reason for this controversy is the fact that not all compartments within the macrophage contain lowered Mg^{2+} but rather only a subset. This information could be obtained via a comparison of different intracellular bacteria containing M-box regulatory elements fused to reporter genes. Since different intracellular bacteria inhabit different micro-niches within the host cell, Mg^{2+}-limited niches can be determined based on which bacteria exhibit an increase in M-box-controlled gene expression. The information obtained from this experiment could have broad implications because it would solidify the connection between Mg^{2+}-deprivation and virulence gene regulation. Any Mg^{2+}-regulated genes

within a pathogen could be viewed as a potential virulence factor although they would need to be verified as such by further studies. In conclusion, this work could perhaps open doors to the discovery of future antimicrobial drug targets.

BIBLIOGRAPHY

Abdallah, A. M., Verboom, T, Weerdenburg, E. M., Gey van Pittius, N. C., Mahasha, P. W., Jiménez, C., Parra, C. M., Cadieux, N., Brennan, M. J., Appelmelk, B. J., and Bitter, W. (2009). PPE and PE_PGRS proteins of *Mycobacterium marinum* are transported via the type VII secretion system ESX-5. *Mol. Micro.* **73**, 329-340.

Alix, E. and Blanc-Potard, A. B. (2007). MgtC: A key player in intramacrophage survival. *Trend Microbiol.* **15**, 252-256.

Barrick, J.E., Corbino, K.A., Winkler, W.C., Nahvi, A., Mandal, M., Collins, J., Lee, M., Roth, A., Sudarsan, N., Jona, I., et al. (2004). New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc. Natl. Acad. Sci. USA* **101**, 6421–6426.

Barton, L. L., Goulhen, F., Bruschi, M., Woodards, N. A., Plunkett, R. M., Rietmeijer, F. J. M. (2007). The bacterial metallome: composition and stability with specific reference to the anaerobic bacterium *Desulfovibrio desulfuricans* . *Biometals*. **20**, 291-302.

Basu, S. and Strobel, S. A. (1999). Thiophilic metal ion rescue of phosphorothioate interference within the *Tetrahymena* ribozyme P4-P6 domain. *RNA.* **5**, 1399-1407.

Batey, R.T. (2006). Structures of regulatory elements in mRNAs. *Curr. Opin. Struct. Biol.* **16**, 299–306.

Biachoo, N., Wang, T., Ye, R., and Helmann, J.D. (2002). Global analysis of the *Bacillus subtilis* Fur regulon and the iron starvation stimulon. *Mol. Microbiol.* **6**, 1613–1629.

Bijlsma, J. J. E. and Groisman, E. A. (2005). The PhoP/PhoQ system controls the intramacrophage type three secretion system of *Salmonella enterica. Mol. Micro.* **57**, 85-96.

Birmingham, C. L., Canadien, V., Kaniuk, N. A., Steinberg, B. E., Higgins, D. E., and Brumell, J. H. (2008) Listeriolysin O allows *Listeria monocytogenes* replication in macrophage vacuoles. *Nature*. **451**, 350-354.

Bookout, A. L., Cummins, C. L., Mangelsdorf, D. J., Pesola, J. M., and Kramer, M. F. (2006). High-throughput real-time quantitative reverse transcription PCR. *Curr. Protoc. Mol. Biol.* Chapter 15, Unit 15.8.

Brautigam, C. A. and Steitz, T. A. (1998). Structural principles for the inhibition of the 3′-5′ exonuclease activity of *Escherichia coli* DNA polymerase I by phosphorothioates. *J. Mol. Biol.* **277**, 363-377.

Brautigam, C. A., Wakeman, C. A., and Winkler, W. C. (2009). Methods for analysis of ligand-induced RNA conformational changes. *Methods Mol. Biol.* **540**, 77-95.

Carothers, J. M., Davis, J. H., Chou, J. J., and Szostak, J. W. (2006). Solution structure of an informationally complex high-affinity RNA aptamer to GTP. *RNA* **12**, 567–579.

Cate, J.H., Hanna, R.L., and Doudna, J.A. (1997). A magnesium ion core at the heart of a ribozyme domain. *Nat. Struct. Biol.* **4**, 553–558.

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.

Christian, E. L. and Yarus, M. (1992). Analysis of the role of phosphate oxygens in the group I intron from *Tetrahymena*. *J. Mol. Biol.* **228**, 743-758.

Christian, E. L. and Yarus, M. (1993). Metal coordination sites that contribute to structure and catalysis in the group I intron for *Tetrahymena*. *Biochemistry*. **32**, 4475-4480.

Chu, V. B., Bai, Y., Lipfert, J., Herschlag, D., and Doniach, S. (2008). A repulsive field: advances in the electrostatics of the ion atmosphere. *Curr. Opin. Chem. Biol.* **12**, 619-625.

Cochrane, J. C., Lipchock, S. V., and Strobel, S. A. (2007). Structural investigation of the GlmS ribozyme bound to its catalytic cofactor. *Chem. Biol.* **14**, 97-105.

Cochrane, J. C. and Strobel, S. A. (2008). Riboswitch effectors as protein enzyme cofactors. *RNA*. **14**, 993-1002.

Collins, J. A., Irnov, I., Baker, S., and Winkler, W. C. (2007). Mechanism of mRNA destabilization by the *glmS* ribozyme. *Genes Dev.* **21**, 3356-3368.

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.

Cowan, J. A. (1993). Metallobiochemistry of RNA. $Co(NH_3)_6^{3+}$ as a probe for Mg^{2+} (aq) binding sites. *J. Inorg. Biochem.* **49**, 171–175.

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

Daly, M. J., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Zhai, M., Venkateswaran, A., Hess, M., Omelchenko, M. V., Kostandarithes, H. M., Makarova, K. S., Wackett, L. P., Fredrickson, J. K., and Ghosal, D. (2004). Accumulation of Mn(II) in *Deinococcus radiodurans* facilitates gamma-radiation resistance. *Science.* **306**, 1025-1028.

Dambach, M. D. and Winkler, W. C. (2009). Expanding roles for metabolite-sensing regulatory RNAs. *Curr. Opin. Microbiol.* **12**, 161-169.

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**, 878-892.

Das, R., Travers, K. J., Bai, Y., and Herschlag, D. (2005a). Determining the Mg^{2+} stoichiometry for folding an RNA metal ion core. *J. Am. Chem. Soc.* **127**, 8272–8273.

Das, R., Laederach, A., Pearlman, S., Herschlag, D., and Altman, R. (2005b). SAFA: Semi-automated footprinting analysis software for high-throughput quantification of nucleic acid footprinting experiments. *RNA.* **11**, 344-354.

de Hoon, M.J., Makita, Y., Nakai, K., and Miyano, S. (2005). Prediction of transcriptional terminators in Bacillus subtilis and related species. *PLoS Comput. Biol.* **1**, e25. 10.1371/journal.pcbi.0010025.

Delano, W. L. (2002). The PyMOL Molecular Graphics System. DeLano Scientific; San Carlos, CA.

DeRose, V.J. (2003). Metal ion binding to catalytic RNA molecules. *Curr. Opin. Struct. Biol.* **13**, 317–324.

Draper, D.E., Grilley, D., and Soto, A.M. (2005). Ions and RNA folding. *Annu. Rev. Biophys. Biomol. Struct.* **34**, 221–243.

Draper, D. E. (2008). RNA folding: Thermodynamics and molecular descriptions of the roles of ions. *Biophysical Journal*. **95**, 5489-5495.

Dupureur, C. M. (2008). Roles of metal ions in nucleases. *Curr. Opin. Chem. Biol.* **12**, 250-255.

Eckstein, F. (1985). Nucleoside phophorothioates. *Ann. Rev. Biochem.* **54**, 367-402.

Eguchi, Y. and Utsumi, R. (2008). Introduction to bacterial signal transduction networks. *Adv. Exp. Med. Biol.* **631**, 1-6.

Ellington, A. D., Chen, X., Robertson, M., and Syrett, A. (2009). Evolutionary origins and directed evolution of RNA. *Int. J. Biochem. Cell Biol.* **41**, 254-265.

Fan, P., Suri, A. K., Fiala, R., Live, D., and Patel, D. J. (1996). Molecular recognition in the FMN-RNA aptamer complex.

Frey, P. A. and Sammons, R. D. (1985). Bond order and charge localization in nucleoside phosphorothiates. *Science*. **228**, 541-545.

Gardner, R.C. (2003). Genes for Mg^{2+} transport. *Curr. Opin. Plant Biol.* **6**, 263–267.

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**, 22347-22351.

Giedroc, D. P. and Arunkumar, A. I. (2007). Metal sensor proteins: nature's metalloregulated allosteric switches. *Dalton Trans.* 3107-3120.

Gish, G. and Eckstein, F. (1988). DNA and RNA sequence determination based on phosphorothioate chemistry. *Science*. **240**, 1520-1522.

Gollnick, P. and Babitzke, P. (2002). Transcription attenuation. *Biochim. Biophys. Acta.* **1577**, 240-250.

Gollnick, P., Babitzke, P., Antson, A., and Yanofsky C. (2005). Complexity in regulation of tryptophan biosynthesis in *Bacillus subtilis*. *Annu. Rev. Genet.* **39**, 47-68.

Gottesman, S., McCullen, C. A., Guillier, M., Vanderpool, C.K., Majdalani, N., Benhammou, J., Thompson, K. M., FitzGerald, P. C., Sowa, N. A., and FitzGerald, D. J. (2006). Small RNA regulators and the bacterial response to stress. *Cold Spring Harb. Symp. Quant. Biol.* **71**, 1-11.

Griffiths, A. D., Potter, B. V. L., and Eperon, I. C. (1987). Stereospecificity of nucleases towards phosphorothioate-substituted RNA: stereochemistry of transcription by T7 polymerase. *Nucl. Acids Res.* **15**, 4145-4162.

Griffiths-Jones S., Moxon S., Marshall M., Khanna A., Eddy S.R., and Bateman A. (2005). Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res*. **33**, D121-4.

Gusarov, I., and Nudler, E. (1999). The mechanism of intrinsic transcription termination. *Mol. Cell* **3**, 495–504.

Hannon, G. J., Rivas, F.V., Murchison, E.P., and Steitz, J. A. (2006). The expanding universe of noncoding RNAs. *Cold Spring Harbor Symp. Quant. Biol.* **71**, 551-564.

Hinrichs, W., Steifa, M., Saenger, W., and Eckstein, F. (1987). Absolute configuration of *R*p-uridine 3´,5´-cyclic phosphorothioate. *Nucl. Acids Res*. **15**, 4945-4955.

Hmiel, S.P., Snavely, M.D., Miller, C.G., and Maguire, M.E. (1986). Magnesium transport in *Salmonella typhimurium*: Characterization of magnesium influx and cloning of a transport gene. *J. Bacteriol.* **168**, 1444–1450.

Horton, T. E., Maderia, M., and DeRose, V. J. (2000). Impact of phosphorothioate substitutions on the thermodynamic stability of a RNA GAAA tetraloop: an unexpected stabilization. *Biochemistry.* **39**, 8201-8207.

Hougland, J.L., Kravchuk, A.V., Herschlag, D., and Piccirilli, J.A. (2005). Functional identification of catalytic metal ion binding sites within RNA. *PLoS Biol.* **3**, e277. 10.1371/journal.pbio.0030277.

Jakubovics, N. S. and Valentine, R. A. (2009). A new direction for manganese homeostasis in bacteria: identification of a novel efflux system in *Streptococcus pneumoniae. Mol. Micro*. **72**, 1-4.

Jansen, J. A., McCarthy, T. J., Soukup, G. A., and Soukup, J. K. (2006). Backbone and nucleobase contacts to glucosamine-6-phosphate in the *glmS* ribozyme. *Nat. Struct. Mol. Biol.* **13**, 517-523.

Jarmer, H., Berka, R., Knudsen, S., and Saxild, H.H. (2002). Transcriptome analysis documents induced competence of *Bacillus subtilis* during nitrogen limiting conditions. *FEMS Microbiol. Lett.* **206**, 197– 200.

Jiang, F., Kumar, R. A., Jones, R. A., and Patel, D. J. (1996). Structural basis of RNA folding and recognition in an AMP-RNA aptamer complex. *Nature.* **382**, 183-186.

Kato, A. and Groisman, E. A. (2008). The PhoQ/PhoP regulatory network of *Salmonella enteric*. *Adv. Exp. Med. Biol.* **631**, 7-21.

Kehres, D.G., and Maguire, M.E. (2002). Structure, properties and regulation of magnesium transport proteins. *Biometals* **15**, 261–270.

Kieft, J. S., and Batey, R. T. (2004). A general method for rapid and nondenaturing purification of RNAs. *RNA* **10**, 988-995.

Klein, D.J., Moore, P.B., and Steitz, T.A. (2004). The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA* **10**, 1366–1379.

Landick, R., and Yanofsky, C. (1987). Transcription attenuation. In *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology, F.C. Neidhardt, ed. (Washington, D.C.: American Society for Microbiology), pp. 1276–1301.

Latham, J.A., and Cech, T.R. (1989). Defining the inside and outside of a catalytic RNA molecule. *Science* **245**, 276–282.

Laue, T.M., Shah, B.D., Ridgeway, T.M., and Pelletier, L.S. (1992). Computer-aided interpretation of analytical sedimentation data for proteins. In Analytical Ultracentrifugation in Biochemistry and Polymer Science, S. Harding, A. Rowe, and J. Horton, eds. (Royal Society of Chemistry), pp. 90-125.

Lee, J.C., Gutell, R.R., and Russell, R. (2006). The UAA/GAN internal loop motif: a new RNA structural element that forms a cross-strand AAA stack and long-range tertiary interactions. *J. Mol. Biol.* **360**, 978–988.

Leontis, N.B. and Westhof, E. (2001). Geometric nomenclature and classification of RNA base pairs. *RNA.* **7**, 499-512.

Li, Y., and Breaker, R.R. (1999). Kinetics of RNA degradation by specific base catalysis of transesterification involving the 20-hydroxyl group. *J. Am. Chem. Soc.* **121**, 5364–5372.

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 cholera*. *J. Mol. Biol.* **365**, 1393-1406.

Macdiarmid, C.W., and Gardner, R.C. (1998). Overexpression of the Saccharomyces cerevisiae magnesium transport system confers resistance to aluminum ions. *J. Biol. Chem.* **273**, 1727–1732.

Maguire, M.E. (2006). Magnesium transporters: properties, regulation and structure. *Front. Biosci.* **111**, 3149–3163.

Mandal, M., Lee, M., Barrick, J. E., Weinberg, Z., Emilsson, G. M., Ruzzo, W. L., and Breaker, R. R. (2004). A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science*. **306**, 275-279.

Merino, E.J., Wilkinson, K.A., Coughlan, J.L., and Weeks, K.M. (2005). RNA structure analysis at single nucleotide resolution by selective 2′-hydroxyl acylation and primer extension (SHAPE). *J. Am. Chem. Soc.* **127**, 4223-4231.

Mironov, A.S., Gusarov, I. Rafikov, R., Lopez, L.E., Shatalin, K., Kreneva, R.A., Perumov, D.A., and Nudler, E. (2002). Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* **111**, 747-756.

Montange, R. K. and Batey, R. T. (2006). Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. *Nature*. **441**, 1172-1175.

Montange, R. K. and Batey, R. T. (2008). Riboswitches: emerging themes in RNA structure and function. *Annu. Rev. Biophys*. **37**, 117-133.

Moore, C.M., and Helmann, J.D. (2005). Metal ion homeostasis in *Bacillus subtilis*. *Curr. Opin. Microbiol.* **8**, 188–195.

Nissen, P., Ippolito, J.A., Ban, N., Moore, P.B., and Steitz, T.A. (2001). RNA tertiary interactions to the large ribosomal subunit: the A-minor motif. *Proc. Natl. Acad. Sci. USA* **98**, 4899-4903.

Outten, C.E., and O'Halloran, T.V. (2001). Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* **292**, 2488–2492.

Pennella, M.A., and Giedroc, D.P. (2005). Structural determinants of metal selectivity in prokaryotic metal-responsive transcriptional regulators. *Biometals* **18**, 413–428.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004). UCSF Chimera – A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 605-612.

Que, Q. and Helmann, J. D. (2000). Manganese homeostasis in *Bacillus subtilis* is regulated by MntR, a bifunctional regulator related to the diphtheria toxin repressor family of proteins. *Mol. Micro.* **35**, 1454-1468.

Ramesh, A. and Winkler, W. C. (2009). Magnesium-sensing riboswitches in bacteria. *RNA Biology*, in press.

Roth, A. and Breaker, R. R. (2009). The structural and functional diversity of metabolite-binding riboswitches. *Annu. Rev. Biochem.* **78**, 305-334.

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**, 308-317.

Ruffner, D. E. and Uhlenbeck, O. C. (1990). Thiophosphate interference experiments locate phosphates important for the hammerhead RNA self cleavage reaction. *Nucleic Acids Res.* **18**, 6025-6029.

Ryder, S. P. and Strobel, S. A. (1999). Nucleotide analog interference mapping. *METHODS: A Companion to Methods in Enzymology*. **18**, 38-50.

Schuck, P. (2000). Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophysical Journal.* **78**, 1606-1619.

Schwalbe, H., Buck, J., Furtig, B., Noeske, J., and Wohnert, J. (2007). Structures of RNA switches: insight into molecular recognition and tertiary structure. *Angew. Chem. Int. Ed. Engl.* **46**, 1212-1219.

Selmer, M., Dunham, C.M., Murphy, F.V., 4th, Weixlbaumer, A., Petry, S., Kelley, A.C., Weir, J.R., and Ramakrishnan, V. (2006). Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**, 1935–1942.

Serganov, A. (2009). The long and the short of riboswitches. *Curr. Opin. Struct. Biol.* **19**, 251-259.

Serganov, A., Huang, L., and Patel, D. J. (2009). Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch. *Nature*. **458**, 233-237.

Serganov, A., Polonskaia, A., Phan, A.T., Breaker, R.R., and Patel, D.J. (2006). Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature* **441**, 1167–1171.

Serganov, A., Yuan, Y. R., Pikovskaya, O., 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**, 1729-1741.

Shaughnessy, L. M., Hoppe, A. D., Christensen, K. A., and Swanson, J. A. (2006). Membrane perforations inhibit lysosome fusion by altering pH and calcium in *Listeria monocytogenes* vacuoles. *Cell. Micro.* **8**, 781-792.

Sigel, R.K., and Pyle, A.M. (2007). Alternative roles for metal ions in enzyme catalysis and the implications for ribozyme chemistry. *Chem. Rev.* **107**, 97–113.

Smith, R.L., and Maguire, M.E. (1995). Distribution of the CorA Mg^{2+} transport system in Gram-negative bacteria. *J. Bacteriol.* **177**, 1638–1640.

Smith, R.L., and Maguire, M.E. (1998). Microbial magnesium transport: unusual transporters searching for identity. *Mol. Microbiol.* **28**, 217– 226.

Soukup, G.A., and Breaker, R.R. (1999). Relationship between internucleotide linkage geometry and the stability of RNA. *RNA* **5**, 1308–1325.

Spinelli, V. S., Pontel, L. B., Garcia Vescovi, E., and Soncini, F. C. (2008). Regulation of magnesium homeostasis in *Salmonella*: Mg^{2+} targets the *mgtA* transcript for degradation by RNase E. *FEMS Micro. Letters*. **280**, 226-234.

Spitale, R. C., Torelli, A. T., Krucinska, J., Bandarian, V., and Wedekind, J. E. (2009). The structural basis for recognition of the PreQ0 metabolite by an unusually small riboswitch aptamer domain. *J. Biol. Chem.* **284**, 11012-11016.

Stahley, M. R., Adams, P. L., Wang, J., and Strobel, S. A. (2007). Structural metals in the group I intron: a ribozyme with a multiple metal ion core. *J. Mol. Biol.* **372**, 89-102.

Stoddard, C. D., Gilbert, S. D., and Batey, R. T. (2008). Ligand-dependent folding of the three-way junction in the purine riboswitch. *RNA*. **14**, 675-684.

Stoltenberg, R., Reinemann, C., and Strehlitz, B. (2007). SELEX - A (r)evolutionary method to generate high-affinity nucleic acid ligands. *Biomol. Eng.* **24**, 381-403.

Storz, G., Opdyke, J. A., and Wassarman, K. M. (2006). Regulating bacterial transcription with small RNAs. *Cold Spring Harb. Symp. Quant. Biol.* **71**, 269-273.

Strobel, S. A. and Shetty, K. (1997). Defining the chemical groups essential for *Tetrahymena* group I intron function by nucleotide analog interference mapping. *Proc. Natl. Acad. Sci. USA*. **94**, 2903-2908.

Sudarsan, N., Wickiser, K., Nakamura, S., Ebert, M. S., and Breaker, R. R. (2003). An mRNA structure in bacteria that controls gene expression by binding lysine. *Genes & Dev.* **17**, 2688-2697.

Sussman, D., Nix, J. C., and Wilson, C. (2000). The structural basis for molecular recognition by the vitamin B12 RNA aptamer. *Nat. Struct. Biol.* **7**, 53-57.

Takamoto K, He Q, Morris S, Chance M.R., and Brenowitz M. (2002). Monovalent cations mediate formation of native tertiary structure of the Tetrahymena thermophila ribozyme. *Nat Struct Biol.* **12**, 928-33.

Takamoto, K., Das, R., He, Q., Doniach, S., Brenowitz, M., Herschlag, D., and Chance, M.R. (2004). Principles of RNA compaction: insights from the equilibrium folding pathway of the P4–P6 RNA domain in monovalent cations. *J. Mol. Biol.* **343**, 1195–1206.

Toledo-Arana, A., Dussurat, O., Nikitas, G., Sesto, N., Guet-Revillet, H., Balestrino, D., Loh, E., Gripenland, J., Tiensuu, T., Vaitkevicius, K., Barthelemy, M., Vergassola, M., Nahori, M. A., Soubigou, G., Regnault, B., Coppee, J. Y., Lecuit, M., Johansson, J., and Cossart, P. (2009). The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature*. **459**, 950-956.

Vicens, Q., and Cech, T.R. (2006). Atomic level architecture of group I introns revealed. *Trends Biochem. Sci.* **31**, 41–51.

Wachter, A., Tunc-Ozdemir, M., Grove, B. C., Green, P. J., Shintani, D. K., and Breaker, R. R. (2007). Riboswitch control of gene expression in plants by splicing and alternative 3′ end processing of mRNAs. *Plant Cell.* **19**, 3437-3450.

Wakeman, C. A., Winkler, W. C., & Dann, C. E. (2007). Structural features of metabolite-sensing riboswitches. *TiBS*. **32**, 415-424.

Wakeman, C. A. and Winkler, W. C. (2009a). Analysis of the RNA backbone: Structural analysis of riboswitches by in-line probing and selective 2′-hydroxyl acylation and primer extension. *Methods Mol. Biol.* **540**, 173-191.

Wakeman, C. A. and Winkler, W. C. (2009b). Structural probing techniques on natural aptamers. *Methods Mol. Biol.* **535**, 115-133.

Wakeman, C. A., Ramesh, A., and Winkler, W. C. (2009). Multiple Metal-Binding Cores Are Required for Metalloregulation by M-box Riboswitch RNAs. *J. Mol. Biol.* **392**, 723-735.

Waldsich, C. (2008). Dissecting RNA folding by nucleotide analog interference mapping (NAIM). *Nature Protocols.* **3**, 811-823.

Waring, R. B. (1989). Identification of phosphate groups important to self-splicing of the *Tetrahymena* rRNA intron as determined by phosphorothioate substitution. *Nucleic Acids Res.* **17**, 10281-10293.

Weinstein, L. B., Jones, B. C. N. M., Cosstick, R., & Cech, T. R. (1997). A second catalytic metal ion in a group I ribozyme. *Nature*. **388**, 805-808.

Welz, R. and Breaker R. R. (2007). Ligand binding and gene control characteristics of tandem riboswitches in *Bacillus anthracis*. *RNA.* **13**, 573-582.

Winkler, W.C. (2005a). Riboswitches and the role of noncoding RNA in bacterial metabolic control. *Curr. Opin. Chem. Biol.* **9**, 594–602.

Winkler, W.C. (2005b). Metabolic monitoring by bacterial mRNAs. *Arch. Microbiol.* **183**, 151–159.

Winkler, W.C., and Breaker, R.R. (2005). Regulation of bacterial gene expression by riboswitches. *Annu. Rev. Microbiol.* **59**, 487–517.

Winkler, W.C., Nahvi, A., and Breaker, R.R. (2002). Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**, 952–956.

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. U.S.A.* **99**, 15908-15913.

Winkler, W.C., Nahvi, A., Roth, A., Collins, J.A., and Breaker, R.R. (2004). Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428**, 281-286.

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**, 701-707.

Woodson, S.A. (2005). Metal ions and RNA folding: a highly charged topic with a dynamic future. *Curr. Opin. Chem. Biol.* **9**, 104–109.

Yanofsky, C. (2004). The different roles of tryptophan transfer RNA in regulating *trp* operon expression in *E. coli* versus *B. subtilis*. *Trends Genet.* **20**, 367-374.

Yanofsky, C. (2007). RNA-based regulation of genes of tryptophan synthesis and degradation, in bacteria. *RNA*. **13**, 1141-1154.

Zuker, M. (2003). MFOLD web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406-3415.