CHARACTERIZATION OF THE S-ADENOSYLMETHIONINE-DEPENDENT REGULATION AND PHYSIOLOGICAL ROLES OF GENES IN THE S BOX SYSTEM

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Brooke A. McDaniel, B.A.

*****
The Ohio State University
2005

Dissertation Committee:
Professor Tina M. Henkin, Adviser
Professor Charles J. Daniels
Professor John N. Reeve
Professor F. Robert Tabita

Approved by

Adviser

Graduate Program in Microbiology
ABSTRACT

The S box system is a transcription termination control system found mostly in Gram-positive bacteria that regulates the expression of many genes involved in sulfur metabolism. Genes regulated by this mechanism are characterized by the presence of a set of highly conserved primary sequence and secondary structural elements in the untranslated leader region upstream of the regulated coding sequence. Synthesis of the full-length transcript is determined by whether the leader folds into the stem-loop of an intrinsic terminator or a competing antiterminator structure; termination is dependent on formation of an anti-antiterminator structure that competes with the antiterminator. It has been demonstrated for several *Bacillus subtilis* S box genes that growth in the presence of methionine results in efficient termination of transcription at the leader region terminator, while terminator readthrough is induced in response to starvation for methionine.

A model for the S box regulatory mechanism was proposed in which methionine availability is monitored by binding of a regulatory factor to the leader RNA, and it was suggested that *S*-adenosylmethionine (SAM) could serve as the metabolic signal. SAM was shown to specifically promote termination at S box leader region terminators in a purified *in vitro* transcription system, in the absence of any additional *trans*-acting factors. S box leader RNA is able to discriminate against SAM-related compounds, which is essential for an appropriate regulatory response in the cell. SAM binding assays
demonstrated that SAM binds specifically and directly to the helix 1-4 region of S box leader RNAs, and that the entire helix 1-4 region is important for SAM binding. Binding of SAM and SAM-directed transcription termination are blocked by leader mutations that cause loss of repression during growth in the presence of methionine. Mutational analysis confirmed the importance of a conserved RNA structural motif found in S box leaders that could facilitate potential tertiary interactions in S box leader RNAs. Analysis of S box leader variants demonstrated that interaction between two covarying regions in S box leader RNA is required for SAM binding, SAM-directed transcription termination, and SAM-induced structural changes. This study provided the first information concerning the tertiary structure of S box leader RNAs.

A trans-acting mutation that leads to a decrease in SAM synthetase activity and derepression of S box gene expression during growth in the presence of methionine was identified in the metK gene, which encodes SAM synthetase. Identification of this mutation, in conjunction with the observation that overproduction of SAM synthetase in vivo results in delayed induction of a yitJ-lacZ transcriptional fusion in response to starvation for methionine, are consistent with our identification of SAM as the molecular effector for S box gene expression.

Roles in methionine biosynthesis have been demonstrated or postulated for a number of S box genes. The presence of the S box regulatory motif in the ykrTS and ykrWXYZ operons and sequence similarity of the gene products encoded in these operons
to proteins of known function suggested that these genes are involved in methionine biosynthesis. Roles in the pathway for recycling of 5’-methylthioadenosine (MTA) to methionine were demonstrated for the products of the $ykrT$, $ykrS$, and $ykrW$ genes, and $ykrT$ was identified as the gene encoding 5-methylthioribose (MTR) kinase.
Dedicated to my family, especially my husband, Corey, and my parents, Bobby and Sharon Murphy
ACKNOWLEDGMENTS

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

I am extremely grateful to Dr. Frank Grundy for his guidance and encouragement. Dr. Grundy played an instrumental role in this project, particularly in the work identifying the molecular effector for transcription termination in the S box system.

I would like to thank my committee members, Dr. Charles Daniels, Dr. John Reeve, and Dr. F. Robert Tabita, for their insight, support and time over the years.

I also wish to thank my colleague and dear friend, Dr. Mary Yousef, for being a source of support on all matters scientific, personal, and spiritual. It is not possible to express how valuable her friendship and encouragement have been, and will continue to be for many years to come. I would like to thank Susan Tigert for her technical assistance, support, and friendship.

Most of all I give thanks to God, without whom this work would not have been possible, and my family, for endless support and encouragement.

This work was supported by grants and fellowships from The Ohio State University, the National Consortium for Graduate Degrees for Minority Students in Engineering and Science, and the National Institutes of Health.
VITA

April 15, 1975.............................................Born - Urbana, Ohio

1997.............................................................B.A. Microbiology and Chemistry, Miami University

1997 – 1998.................................................Coordinator,
Short-Term Research Experience Access for Minority Students, Wright State University

1998 – present.............................................Graduate Research Associate,
The Ohio State University

PUBLICATIONS

Research Publications


FIELDS OF STUDY

Major Field: Microbiology
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>v</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>vi</td>
</tr>
<tr>
<td>Vita</td>
<td>vii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xiii</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>xvi</td>
</tr>
</tbody>
</table>

## Chapters:

1. Regulation of gene expression in bacteria...........................................1
   1.1 Translating ribosomes........................................................................3
   1.2 RNA-binding proteins.........................................................................5
      1.2.1 Positive-acting RNA-binding proteins......................................6
         1.2.1.1 *E. coli* BglG...................................................................6
         1.2.1.2 *B. subtilis* GlpP......................................................7
         1.2.1.3 *Pseudomonas aeruginosa* AmiR......................................9
      1.2.2 Negative-acting RNA-binding proteins...................................11
         1.2.2.1 *B. subtilis* TRAP....................................................11
         1.2.2.2 *B. subtilis* PyrR...................................................14
   1.3 Riboswitches....................................................................................16
      1.3.1 RNA thermosensors......................................................................17
      1.3.2 T box system.............................................................................20
      1.3.3 Cellular metabolites...............................................................23
         1.3.3.1 Thi box system...............................................................23
         1.3.3.2 RFN element..................................................................25
         1.3.3.3 B$_{12}$ element...........................................................25
         1.3.3.4 *pyrG*........................................................................26
1.3.3.5 G box system ................................................................. 28
1.3.3.6 GlcN6P ribozyme ............................................................. 30
1.3.3.7 L box system ................................................................. 31
1.3.3.8 Glycine system ................................................................. 32
1.3.3.9 S box system ................................................................. 33
1.4 Goals of this study ........................................................................ 37

2. Identification of S-adenosylmethionine as the molecular effector for transcription termination in the S box system ................................................................. 39
2.1 Introduction .................................................................................. 39
2.2 Materials and Methods .................................................................. 41
2.2.1 Generation of DNA templates .................................................. 41
2.2.2 In vitro transcription assays ...................................................... 42
2.2.3 SAM binding assays .................................................................. 42
2.3 Results ......................................................................................... 43
2.3.1 In vitro transcription of S box genes .......................................... 43
2.3.2 SAM-dependent transcription termination .................................. 45
2.3.3 Binding of 14C-SAM to B. subtilis yitJ and ykrW leader RNAs .... 49
2.4 Discussion ................................................................................... 50

3. Analysis of cis-acting sequence and structural elements required for S-adenosylmethionine-directed transcription termination ......................................... 55
3.1 Introduction ................................................................................ 55
3.2 Materials and Methods ................................................................ 58
3.2.1 Bacterial strains and growth conditions .................................. 58
3.2.2 In vitro transcription assays .................................................... 59
3.2.3 SAM binding assays .............................................................. 60
3.2.4 β-galactosidase measurements ............................................... 61
3.2.5 Leader RNA structural analysis .............................................. 61
3.3 Results ....................................................................................... 62
3.3.1 Determination of the minimal SAM binding site in ykrW and yoaD leader RNAs .......................................................... 62
3.3.2 yitJ leader mutations affect SAM-directed transcription termination ........................................................................... 67
3.3.3 Role of the kink-turn motif in S box gene regulation in vivo ....... 71
3.3.4 yitJ and ykrW leader covariation mutants ................................ 75
3.3.4.1 In vitro transcription of yitJ leader covariation mutants ........ 79
3.3.4.2 Binding of $^{14}$C-SAM to yitJ leader covariation mutant RNAs ........................................................................... 83
3.3.4.3 RNase A mapping of yitJ and ykrW leader covariation mutant RNAs ........................................................................... 86
3.3.4.4 In vivo expression of yitJ-lacZ fusions for yitJ covariation mutants .............................................................. 94

3.4 Discussion .................................................................................................................................................. 97

4. Characterization of the Bacillus subtilis metK gene and its effect on S box gene expression .......................................................... 105
4.1 Introduction ...................................................................................................................................... 105
4.2 Materials and Methods ............................................................................................................... 107
  4.2.1 Bacterial strains and growth conditions ...................................................................................... 107
  4.2.2 Genetic techniques .................................................................................................................. 108
  4.2.3 β-galactosidase measurements ............................................................................................... 111
  4.2.4 SAM synthetase assay ........................................................................................................... 111
4.3 Results .............................................................................................................................................. 112
  4.3.1 Isolation of S box-depressed mutants of B. subtilis strain IS56B ......................................................... 112
  4.3.2 Determination of the specificity of B. subtilis strain IS56B-SBD1 ................................................... 115
  4.3.3 In vivo expression of yitJ-lacZ in strains containing the SBD1 allele ........................................ 116
  4.3.4 Identification of the SBD1 allele .............................................................................................. 117
  4.3.5 Measurement of SAM synthetase activity in BR151-SBD1 ...................................................... 120
  4.3.6 Effect of metK overexpression on yitJ-lacZ expression in vivo ............................................. 121
4.4 Discussion ...................................................................................................................................... 123

5. Identification of the gene encoding 5-methylthioribose kinase in Bacillus subtilis ................................................................. 129
5.1 Introduction ...................................................................................................................................... 129
5.2 Materials and Methods ..................................................................................................................... 134
  5.2.1 Bacterial strains and growth conditions .................................................................................. 134
  5.2.2 Genetic techniques ................................................................................................................ 135
  5.2.3 Gene inactivation .................................................................................................................... 135
  5.2.4 β-galactosidase measurements and growth experiments ..................................................... 138
5.3 Results .............................................................................................................................................. 138
  5.3.1 Role of ykrTS and ykrWXYZ genes in growth on MTA or MTR .................................................... 138
  5.3.2 Regulation of ykrTS and ykrWXYZ expression .................................................................... 141
5.4 Discussion.............................................................................................................143

6. Summary and future directions..............................................................................148

List of references.......................................................................................................158
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Effect of kink-turn motif mutations on expression of <em>yitJ-lacZ</em> transcriptional fusions</td>
<td>74</td>
</tr>
<tr>
<td>3.2 Expression of <em>yitJ-lacZ</em> transcriptional fusions in response to methionine availability</td>
<td>96</td>
</tr>
<tr>
<td>4.1 Oligonucleotide primers</td>
<td>109</td>
</tr>
<tr>
<td>4.2 Effect of the SBD1 allele on expression of a <em>yitJ-lacZ</em> transcriptional fusion</td>
<td>116</td>
</tr>
<tr>
<td>5.1 Regulation of <em>lacZ</em> fusions in response to methionine and MTA</td>
<td>142</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td><em>E. coli trp</em> operon</td>
<td>5</td>
</tr>
<tr>
<td>1.2</td>
<td><em>E. coli bgl</em> operon</td>
<td>7</td>
</tr>
<tr>
<td>1.3</td>
<td><em>B. subtilis glp</em> regulon</td>
<td>9</td>
</tr>
<tr>
<td>1.4</td>
<td><em>P. aeruginosa amiE</em> gene</td>
<td>11</td>
</tr>
<tr>
<td>1.5</td>
<td><em>B. subtilis trp</em> operon</td>
<td>13</td>
</tr>
<tr>
<td>1.6</td>
<td><em>B. subtilis pyr</em> operon</td>
<td>15</td>
</tr>
<tr>
<td>1.7</td>
<td>Mechanism of gene regulation by an RNA thermosensor</td>
<td>19</td>
</tr>
<tr>
<td>1.8</td>
<td>The T box system</td>
<td>22</td>
</tr>
<tr>
<td>1.9</td>
<td>The Thi box system</td>
<td>24</td>
</tr>
<tr>
<td>1.10</td>
<td><em>B. subtilis pyrG</em> gene</td>
<td>28</td>
</tr>
<tr>
<td>1.11</td>
<td>GlcN6P ribozyme</td>
<td>31</td>
</tr>
<tr>
<td>1.12</td>
<td><em>B. subtilis yitJ</em> leader structural model</td>
<td>35</td>
</tr>
<tr>
<td>1.13</td>
<td>S box antitermination model</td>
<td>37</td>
</tr>
<tr>
<td>2.1</td>
<td><em>In vitro</em> transcription of S box genes</td>
<td>44</td>
</tr>
<tr>
<td>2.2</td>
<td>SAM-dependent transcription termination of <em>B. subtilis</em> S box leaders</td>
<td>46</td>
</tr>
<tr>
<td>2.3</td>
<td>Specificity of SAM-dependent transcription termination of the <em>B. subtilis ykrW</em> leader</td>
<td>48</td>
</tr>
<tr>
<td>2.4</td>
<td>Binding of $^{14}$C-SAM by <em>B. subtilis yitJ</em> and <em>ykrW</em> leader RNAs</td>
<td>50</td>
</tr>
</tbody>
</table>
3.1 Secondary structure models of ykrW and yoaD leader RNAs.................................65
3.2 Binding of $^{14}$C-SAM by ykrW and yoaD leader RNAs........................................66
3.3 B. subtilis yitJ Pst leader variants........................................................................68
3.4 In vitro transcription of B. subtilis yitJ Pst leader variants......................................69
3.5 Binding of $^{14}$C-SAM by yitJ-Pst-2 leader RNA....................................................70
3.6 Kink-turn motif sequence and consensus pattern................................................72
3.7 Secondary structure models of yitJ and ykrW S box leader RNAs..........................78
3.8 Covarying regions in S box leader RNAs...............................................................79
3.9 In vitro transcription of B. subtilis yitJ and ykrW leader variants............................81
3.10 Binding of $^{14}$C-SAM by yitJ and ykrW leader RNAs............................................85
3.11 RNase A mapping of yitJ leader RNAs..................................................................88
3.12 RNase A mapping of ykrW leader RNAs.............................................................90
4.1 Strategy for isolation of S box-specific mutants.....................................................113
4.2 Sequence comparison of bacterial MetK homologs..............................................120
4.3 Effect of overexpression of SAM synthetase on expression of a yitJ-lacZ transcriptional fusion........................................................................................................122
4.4 Location of S80 in the crystal structure of the E. coli SAM synthetase ternary complex.........................................................................................................................126
5.1 Predicted methionine biosynthesis pathways in B. subtilis.................................131
5.2 Region of the B. subtilis chromosome containing the ykrTS and ykrWXYZ operons......................................................................................................................133
5.3 Plasmid used to generate strain BR151-YkrTKO................................................137

5.4 Growth of ykrTS mutants and ykrWXYZ mutants.............................................140

6.1 The proposed model for regulation of S box gene expression in response to SAM.................................................................153
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AdoCbl</td>
<td>adenosylcobalamin</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>bp</td>
<td>base-pair</td>
</tr>
<tr>
<td>CTP</td>
<td>cytidine triphosphate</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>EMS</td>
<td>methanesulfonic acid</td>
</tr>
<tr>
<td>FMN</td>
<td>flavin mononucleotide</td>
</tr>
<tr>
<td>GlcN6P</td>
<td>glucosamine-6-phosphate</td>
</tr>
<tr>
<td>GTP</td>
<td>guanosine triphosphate</td>
</tr>
<tr>
<td>MTA</td>
<td>5'-methylthioadenosine</td>
</tr>
<tr>
<td>MTR</td>
<td>5-methylthioribose</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>nt</td>
<td>nucleotide</td>
</tr>
<tr>
<td>NTP</td>
<td>nucleotide triphosphate</td>
</tr>
<tr>
<td>OD</td>
<td>optical density</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>RNase</td>
<td>ribonuclease</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>RNAP</td>
<td>RNA polymerase</td>
</tr>
<tr>
<td>SAH</td>
<td>S-adenosylhomocysteine</td>
</tr>
<tr>
<td>SAM</td>
<td>S-adenosylmethionine</td>
</tr>
<tr>
<td>TBAB</td>
<td>tryptose blood agar base</td>
</tr>
<tr>
<td>TPP</td>
<td>thiamin pyrophosphate</td>
</tr>
<tr>
<td>TRAP</td>
<td>trp RNA-binding attenuation protein</td>
</tr>
<tr>
<td>Tris-HCl</td>
<td>tris-(hydroxymethyl) aminomethane hydrochloride</td>
</tr>
<tr>
<td>Trp</td>
<td>tryptophan</td>
</tr>
<tr>
<td>UTP</td>
<td>uridine triphosphate</td>
</tr>
<tr>
<td>X-gal</td>
<td>5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside</td>
</tr>
</tbody>
</table>
CHAPTER 1

REGULATION OF GENE EXPRESSION IN BACTERIA

Expression of many genes in bacteria is regulated at the level of premature termination of transcription. Transcription termination in bacteria occurs by two major mechanisms, intrinsic and factor-dependent termination. Intrinsic terminators are composed of a G+C-rich helical region followed by a series of U residues in the nascent transcript. Formation of the helix destabilizes the elongation complex, resulting in pausing of RNA polymerase (RNAP) and subsequent release of the template DNA and nascent transcript (Platt, 1998). Factor-dependent termination requires binding of a protein (designated Rho) to a Rho utilization (rut) site in the nascent transcript which results in transcription termination caused by contact of Rho with paused RNAP (Platt and Richardson, 1992; Platt, 1994). Positioning of a terminator element in the 5’ region of a transcript (the leader region) can be used to regulate the expression of the downstream coding sequence.
The two paradigms for regulation of factor-dependent termination are the bacteriophage λ N-mediated system and the tna operon in *Escherichia coli*. The N protein of phage λ binds to the nascent transcript at N utilization (nut) sites, followed by assembly of a stable ribonucleoprotein complex composed of N protein and a set of host-encoded proteins. This complex associates with elongating RNAP, altering the processivity of RNAP by converting it to a form that is resistant to pausing and that transcribes through downstream terminators in the λ genome (Mason *et al*., 1992; DeVito and Das, 1994; Rees *et al*., 1996). λ N-mediated antitermination results in expression of genes required at later stages of infection. In the case of the tna operon, which encodes enzymes responsible for utilization of tryptophan as a carbon and nitrogen source, antitermination is mediated by direct interference with binding of the Rho protein to the nascent transcript. The leader region of the tna operon encodes a peptide, TnaC, that acts in cis in the presence of excess tryptophan to stall the translating ribosome at the tnaC stop codon. The stalled ribosome blocks access of Rho to the rut site in the transcript, preventing Rho-dependent termination and allowing transcription of the downstream coding regions (Gong *et al*., 2001).

In contrast to factor-dependent transcription termination control systems, regulation of intrinsic termination commonly involves modulation of the RNA structure in the 5’ region of a transcript to control expression of downstream genes in response to regulatory signals sensed by the nascent RNA. Formation of the helix of an intrinsic terminator in the nascent transcript results in premature termination of transcription before the downstream coding region is transcribed; sensing of a regulatory molecule by the transcript determines whether or not the leader region terminator helix is formed.
Gene expression can also be regulated at the level of translation initiation by similar RNA structural switches that control access of the translational machinery to the nascent transcript. The nascent RNA therefore functions as a sensor that gauges a regulatory signal, which can be a translating ribosome, an RNA-binding protein, a temperature change, an RNA molecule, or a cellular metabolite (Narberhaus, 2002; Stülke, 2002; Grundy and Henkin, 2004b). Although a variety of regulatory signals can be sensed, each leader RNA responds specifically to its own signal, which is critical for a sensitive and accurate regulatory response.

Gene regulation by modulation of the RNA structure has several advantages, including rapid response to sudden changes in environmental conditions and simplicity of the regulatory mechanisms. The focus of this chapter will be genetic systems in which gene expression is controlled by modulation of the RNA structure in response to regulatory signals, and in particular systems where sensing of cellular metabolites induces a structural change in the 5’ region of the RNA transcript.

1.1 Translating ribosomes

Expression of many amino acid-related operons in enteric bacteria is regulated by mechanisms that couple transcription and translation of the nascent RNA (Landick et al., 1996). The leader regions of these operons can form an intrinsic terminator helix or a competing antiterminator structure, and include a short peptide coding region that contains codons for the amino acid that affects expression of the operon. The two paradigms for ribosome-mediated transcription termination control are the *E. coli trp* operon and the *S. typhimurium his* operon. The peptide coding region in the nascent
transcript of the *E. coli trp* operon contains two tandem tryptophan codons. When cells are lacking in charged tRNA\(^{\text{Trp}}\), the translating ribosome stalls at one of these tryptophan codons, which allows formation of the downstream antiterminator structure and prevents termination of transcription (Landick *et al.*, 1996; Fig. 1.1). However, when cells possess adequate levels of charged tRNA\(^{\text{Trp}}\), the ribosome completes translation of the leader peptide, allowing the terminator to form and preventing expression of the downstream coding region. A number of amino acid biosynthesis and aminoacyl-tRNA synthetase operons, predominantly in Gram-negative bacteria, are regulated by a similar mechanism (Landick *et al.*, 1996).
Figure 1.1. *E. coli* *trp* operon. The ribosome efficiently translates the leader peptide in the presence of adequate levels of charged tRNA$_{Trp}$, allowing formation of the terminator helix and premature termination of transcription. When cells are deficient in charged tRNA$_{Trp}$, the translating ribosome stalls at one of the two tandem tryptophan codons in the nascent transcript, resulting in formation of the antiterminator structure and expression of the downstream coding regions. T, terminator; AT, antiterminator. Modified from Henkin and Yanofsky (2002).

1.2 RNA-binding proteins

Binding of an RNA-binding protein to a leader RNA can also control formation of competing structures in the RNA that promote transcription termination or antitermination. The activity of the RNA-binding protein can in turn be controlled by direct interaction with an effector molecule, by protein-protein interactions with a second protein, or by protein phosphorylation (Stülke, 2002). RNA-binding proteins can act positively, promoting transcript elongation, or negatively, triggering premature termination of transcription.
1.2.1 Positive-acting RNA-binding proteins

1.2.1.1 *E. coli* BglG

The *bgl* operon of *E. coli* contains three genes, *bglG*, *bglF*, and *bglB*, that are involved in the utilization of aromatic β-glucoside sugars. Regulation of this operon is controlled by the RNA-binding protein BglG (Amster-Choder and Wright, 1992). The leader region of the *bgl* operon contains alternative terminator and antiterminator structures. Formation of the antiterminator is dependent on binding of BglG to the 5′ region of the RNA transcript, which stabilizes the antiterminator structure (Amster-Choder and Wright, 1993). The concentration of the sugar substrate is measured by the BglF protein, which transports the sugar into the cell and also controls phosphorylation of BglG. In the absence of the sugar substrate, BglF catalyzes phosphorylation of BglG, which prevents BglG from binding the RNA and allows formation of the more stable terminator helix. During growth in the presence of the sugar substrate, BglG catalyzes dephosphorylation of BglF, which allows BglG to dimerize and bind to the leader RNA, stabilizing the antiterminator structure and allowing expression of the downstream genes (Amster-Choder and Wright, 1997; Chen *et al*., 2000; Fig. 1.2). It has been demonstrated that BglG also interacts with the β’ subunit of RNAP (Nussbaum-Schochat and Amster-Choder, 1999), which may also play a role in antitermination. Several other BglG-like antiterminator proteins, including *B. subtilis* SacY (Tortosa *et al*., 1997), LicT (Schnetz *et al*., 1996) and GlcT (Bachem and Stülke, 1998), and *Lactobacillus casei* LacT (Alpert and Siebers, 1997), participate in similar mechanisms for regulation of sugar utilization operons.
Figure 1.2. *E. coli bgl* operon. In the absence of the sugar substrate, BglF phosphorylates BglG, which prevents BglG dimerization. Phosphorylated BglG monomers cannot bind to the leader RNA, therefore the more stable terminator helix forms. In the presence of the sugar substrate, BglG is dephosphorylated by BglF, which allows BglG to dimerize and bind to the antiterminator structure. Binding of BglG stabilizes the antiterminator structure, which prevents formation of the terminator helix, thus preventing premature termination of transcription. T, terminator; AT, antiterminator. Modified from Dole et al. (2002).

1.2.1.2 *B. subtilis* GlpP

The *B. subtilis* *glp* regulon contains four operons, *glpP, glpFK, glpD* and *glpTQ*, involved in the utilization of glycerol and glycerol-3-phosphate (G3P; Holmberg et al., 1990; Beijer et al., 1993; Nilsson et al., 1994). Expression of *glpD* is controlled at the level of transcription termination in response to G3P (Holmberg and Rutberg, 1992). GlpP is an RNA-binding protein that induces expression of the *glp* operons when cells
are grown in the presence of glycerol, which is converted into G3P (Glatz et al., 1996). GlpP activity was proposed to be controlled by binding of G3P, in contrast to the BglG family of antiterminator proteins, which are regulated by reversible phosphorylation (Rutberg, 1997). The leader regions of the *glpFK, glpD* and *glpTQ* operons contain sequences that can form an intrinsic terminator helix (Rutberg, 1997), and interaction between GlpP and *glpD* leader RNA has been demonstrated both *in vivo* and *in vitro* (Glatz et al., 1998a,b). It was proposed that activated GlpP binds to and stabilizes a competing antiterminator structure in the *glpD* leader, preventing transcription termination (Glatz et al., 1998b; Fig. 1.3). Conservation of sequence in the lower regions of the terminator stems in the *glpFK, glpTQ* and *glpD* leader RNAs suggests that the *glpFK* and *glpTQ* operons could be regulated by a similar mechanism (Rutberg, 1997).
Figure 1.3. *B. subtilis glp regulon*. GlpP does not bind to the leader RNA in the absence of glycerol-3-phosphate, which results in formation of the terminator helix and premature termination of transcription. When glycerol-3-phosphate is present, it activates GlpP to bind to and stabilize the competing antiterminator structure, allowing expression of the downstream coding regions. T, terminator; AT, antiterminator; G3P, glycerol-3-phosphate.

1.2.1.3 *Pseudomonas aeruginosa* AmiR

The *P. aeruginosa amiE* gene is involved in utilization of short-chain aliphatic amides such as acetamide, and its expression is positively regulated by the AmiR RNA-binding protein (Cousens *et al.*, 1987). Expression of *amiE* is mediated by transcription antitermination during growth in the presence of acetamide and other short-chain aliphatic amides (Wilson and Drew, 1995). A second regulatory protein, AmiC, monitors the level of acetamide in the cell. In the absence of acetamide, AmiC binds to and inactivates AmiR, resulting in transcription termination (Wilson *et al.*, 1993).
presence of acetamide, AmiC binds acetamide and fails to bind AmiR, allowing AmiR to bind to its target site in the 5’ region of the RNA transcript and promote antitermination (O’Hara et al., 1999; Fig. 1.4). Direct interaction between AmiR and amiE leader RNA was demonstrated in vitro (O’Hara et al., 1999). However, the AmiR-binding site does not resemble an antiterminator structure, in contrast to the binding sites for E. coli BglG and B. subtilis GlpP. It appears that AmiR promotes antitermination by directly interfering with formation of the terminator structure (Wilson et al., 1996), but the molecular mechanism has yet to be determined.
Figure 1.4. *P. aeruginosa amiE* gene. AmiC binds to and inactivates AmiR in the absence of acetamide, allowing formation of the terminator helix and premature termination of transcription. In the presence of acetamide, AmiC binds acetamide and fails to bind AmiR, freeing AmiR to bind to its target site in the nascent transcript and promote antitermination. T, terminator.

1.2.2 Negative-acting RNA-binding proteins

1.2.2.1 *B. subtilis* TRAP

Expression of the *B. subtilis trp* operon, which is involved in tryptophan biosynthesis and transport, is regulated by the *trp* RNA-binding attenuation protein (TRAP; Gollnick, 1994; Babitzke, 1997). The leader region of the *trp* operon transcript can fold into mutually exclusive terminator and antiterminator structures (Shimotsu *et al.*, 1986). When cells are grown in the presence of tryptophan, TRAP is activated by 11
tryptophan molecules and binds to a series of 11 closely spaced (G/U)AG repeats in the leader region that overlap sequence required for formation of the antiterminator. Binding of TRAP prevents formation of the antiterminator structure, allowing formation of the terminator helix (Babitzke et al., 1994; Antson et al., 1995; Fig. 1.5). Under conditions where tryptophan is limiting, TRAP does not bind to the RNA, which allows the more stable antiterminator to form, leading to expression of the downstream coding sequences. The crystal structure of the TRAP protein revealed that it is composed of 11 identical subunits (Antson et al., 1995; Chen et al., 1999), which led to a model in which the 11 (G/U)AG repeats in the bound RNA wrap around the TRAP protein. The crystal structure of the *B. stearothermophilus* TRAP protein-leader RNA complex provided clear support for this model (Antson et al., 1999).
A second protein, the anti-TRAP protein (AT), is also involved in regulation of expression of the *trp* operon. AT binds to tryptophan-activated TRAP and prevents it from binding the *trp* leader RNA (Valbuzzi and Yanofsky, 2001; Valbuzzi et al., 2002). Expression of the AT protein is induced by a decrease in the charging ratio of tRNA$^{Trp}$ through the T box mechanism (Sarsero et al., 2000; see Section 1.4.2). These results
indicate that the both free tryptophan and uncharged tRNA\textsuperscript{Trp} are sensed in controlling tryptophan metabolism in \textit{B. subtilis} (Lee \textit{et al.}, 1996; Sarsero \textit{et al.}, 2000).

Tryptophan-activated TRAP also binds to the leader region of \textit{trpG}, the only gene involved in tryptophan biosynthesis that is not included in the \textit{trp} operon (Otridge and Gollnick, 1993). TRAP binds to a series of (G/U)AG repeats in the \textit{trpG} leader region that overlap the \textit{trpG} ribosome-binding site, inhibiting translation initiation (Yang \textit{et al.}, 1995). Tryptophan-activated TRAP also regulates expression of \textit{trpE} (the first gene in the \textit{trp} operon) at the level of translation initiation by promoting formation of a structure that sequesters the \textit{trpE} Shine-Dalgarno sequence (Du and Babitzke, 1998).

\textbf{1.2.2.2 \textit{B. subtilis} PyrR}

The \textit{pyr} operon in \textit{B. subtilis} includes genes that are involved in pyrimidine biosynthesis, and expression of this operon is repressed at the level of transcription termination during growth in the presence of pyrimidines (Quinn \textit{et al.}, 1991). Binding of the PyrR regulatory protein to the leader region of the \textit{pyr} operon promotes transcription termination by stabilizing an anti-antiterminator structure, which competes with formation of the antiterminator structure and allows formation of the terminator helix, resulting in premature termination of transcription (Turner \textit{et al.}, 1994; Lu \textit{et al.}, 1996; Switzer \textit{et al.}, 1999). PyrR cannot bind to its target sequence in the absence of pyrimidines, allowing formation of the antiterminator structure and expression of the downstream coding regions (Fig. 1.6). In contrast to the single-stranded RNA binding site for TRAP, the sequence to which PyrR consists of a stem-loop structure with a purine-rich internal bulge (Tomchick \textit{et al.}, 1998). Similar PyrR-dependent regulatory
mechanisms have been proposed to operate in other Gram-positive bacteria, including *Bacillus caldolyticus, Enterococcus faecalis, Lactobacillus plantarum, and L. lactis* (Stülke, 2002), as well as a Gram-negative thermophile (Van de Casteele et al., 1997).

**Figure 1.6. B. subtilis pyr operon.** In the absence of pyrimidines, PyrR cannot bind to its target sequence, allowing formation of the antiterminator structure and expression of the downstream coding regions. In the presence of pyrimidines, PyrR promotes transcription termination by binding to and stabilizing the anti-antiterminator structure. The anti-antiterminator structure allows formation of the terminator helix by competing with the antiterminator structure, causing premature termination of transcription. T, terminator; AT, antiterminator, AAT, anti-antiterminator. Modified from Switzer et al. (1999).
1.3 Riboswitches

Recent studies have identified a novel class of gene regulation systems (termed “riboswitches”) in which the nascent transcript directly senses a regulatory signal (which can be a temperature change, a small RNA, or a cellular metabolite) in the absence of any accessory proteins (Mironov et al., 2002; Lai, 2003; Grundy and Henkin, 2004b; Mandal and Breaker, 2004b). An RNA structural switch occurs in the leader region in response to sensing of the regulatory signal to control expression of the downstream coding regions. In the case of cellular metabolites, riboswitches often represent a mechanism for feedback control of expression of biosynthetic pathways. These systems exhibit a high degree of specificity and affinity for their particular effector molecules, which is essential to prevent recognition of related biosynthetic intermediates by the regulatory RNA.

The RNA structural switch that is triggered by sensing of the effector molecule can regulate gene expression by premature termination of transcription, inhibition of translation initiation, or regulation of RNA processing (Lai, 2003; Grundy and Henkin, 2004b; Soukup and Soukup, 2004). In genes for which expression is regulated at the level of transcription termination, the RNA structural switch controls formation of the intrinsic terminator helix in the nascent transcript. Formation of the terminator helix is often prevented by formation of a more stable competing antiterminator structure in the absence of the effector. Binding of the effector molecule can stabilize a third competing structure, the anti-antiterminator, which competes with the antiterminator and permits formation of the terminator (Lai, 2003; Grundy and Henkin, 2004b). The RNA structural switch induced by metabolite binding alters the accessibility of the Shine-Dalgarno.
sequence to the ribosome in genes for which expression is regulated at the level of translation initiation (Lai, 2003; Soukup and Soukup, 2004; Grundy and Henkin, 2004b).

It is of interest to note that these RNA sensors primarily regulate gene expression at the level of premature termination of transcription in Gram-positive bacteria, while in Gram-negative bacteria the predominant mechanism for gene regulation is inhibition of translation initiation (Lai, 2003; Nudler and Mironov, 2004). It appears likely that this correlation reflects the fact that more genes in Gram-positive bacteria are arranged in operons, and it is simpler and more economical to prematurely terminate polycistronic RNA messages rather than exclusively utilizing a post-transcriptional regulation mechanism (Lai, 2003; Nudler and Mironov, 2004). This tendency could also be due to independent evolution of riboswitch-regulated genes in Gram-positive versus Gram-negative bacteria (Vitreschak et al., 2004). There are exceptions to this generalization, which have been proposed to be due to horizontal gene transfer between bacterial species (Vitreschak et al., 2002; Nudler and Mironov, 2004).

1.3.1 RNA thermosensors

Fluctuations in temperature can be sensed directly by the 5’-untranslated region of RNAs, which leads to changes in the leader RNA structure that affect translation initiation. RNA is capable of acting as a thermosensor because the stability of RNA secondary structure is highly susceptible to changes in temperature. In some cases the 5’ region of the RNA transcript contains secondary structure that sequesters the Shine-Dalgarno sequence from the ribosome at lower temperatures, but at higher temperatures the RNA secondary structure is melted and the Shine-Dalgarno sequence is available for
translation initiation (Fig. 1.7). The best-studied example of temperature-dependent activation of gene expression is the *E. coli rpoH* thermostensor. The leader region of the *E. coli rpoH* transcript (which encodes σ^{32}, the heat shock sigma factor) contains sequence capable of forming secondary structure (Nagai et al., 1991a,b) that sequesters the Shine-Dalgarno sequence and the AUG start codon (Morita et al., 1999a,b). This secondary structure inhibits ribosome binding at lower temperatures, but not at higher temperatures when the RNA secondary structure is destabilized (Morita et al., 1999b). Mutational analysis, chemical and enzymatic probing, and *in vitro* translation assays support the model that the leader region of *rpoH* transcripts directly senses temperature changes and regulates translation by controlling access of the ribosome to the translation initiation signals (Morita et al., 1999a,b).
Figure 1.7. Mechanism of gene regulation by an RNA thermosensor. At low temperatures, secondary structure in the nascent transcript sequesters the Shine-Dalgarno sequence and the start codon from the ribosome, inhibiting translation initiation. At higher temperatures the RNA secondary structure is melted, allowing the ribosome to access the Shine-Dalgarno sequence and the start codon for translation initiation. SD, Shine-Dalgarno sequence. Modified from Lai (2003).

Evidence also exists for temperature-dependent activation of expression of virulence genes in bacterial pathogens. The prfA gene, which encodes the key transcriptional activator of Listeria monocytogenes virulence genes, is transcribed at 30°C and 37°C, but is translated only at 37°C. The prfA leader RNA can form secondary structure that sequesters the Shine-Dalgarno sequence. Mutations predicted to destabilize the RNA secondary structure abolish thermoregulation by the prfA leader RNA in vivo and in vitro (Johansson et al., 2002). Controlling the expression of virulence genes in
pathogens in response to temperature is economical, as these genes are expressed only at the elevated temperatures encountered upon entry into the host.

1.3.2 T box system

The T box system is the best characterized system in which the regulatory signal sensed by the nascent transcript is an RNA molecule. Expression of many genes involved in amino acid metabolism in Gram-positive bacteria, including aminoacyl-tRNA synthetase genes, amino acid biosynthesis genes, and amino acid transporter genes, is regulated at the level of transcription termination by the T box system (Grundy and Henkin, 1993, 1994, 2003, 2004a). Genes regulated by the T box system exhibit a set of highly conserved primary sequence and secondary structural elements in their leader RNA regions. Expression of the downstream coding region is regulated by formation of competing terminator and antiterminator structures in the nascent transcript. The largest conserved primary sequence element is the T box, a 14-nucleotide element that forms the 5’ strand of the antiterminator. Expression of each T box-regulated gene is induced in response to limitation for its cognate amino acid, which is sensed by monitoring the charging ratio of the cognate tRNA. Specific interactions between the cognate uncharged tRNA and the 5’ region of the RNA transcript stabilize the antiterminator structure and promote antitermination (Grundy and Henkin 1993, 1994; Grundy et al., 2002b; Fig. 1.8).

The specificity of the leader RNA-tRNA interaction is determined by pairing of the anticodon of the tRNA with a codon (designated the specifier sequence) in the leader RNA, as well as pairing of the acceptor end of the tRNA with residues located in a bulge in the antiterminator structure. Interaction of the acceptor end of the tRNA with the
antiterminator stabilizes the antiterminator and prevents formation of the intrinsic terminator helix, and this interaction is blocked by charging of the tRNA (Grundy and Henkin, 1993, 1994; Yousef et al., 2003; Fig. 1.8). The T box, adenine (see Section 1.4.3.5), and glycine (see Section 1.4.3.8) riboswitches are the only RNA sensors identified to date that induce gene expression in response to binding of their respective effector molecules. In each case, the positive response triggered by the effector molecule correlates to the function of the regulated gene(s). An increase in the T box effector molecule (uncharged tRNA) signals the need for increasing the level of T box gene products, to facilitate synthesis and transport of amino acids and tRNA charging.
Figure 1.8. **The T box system.** When the charging ratio of the cognate tRNA is high, the terminator helix forms and premature termination of transcription occurs. When the charging ratio of the cognate tRNA is low, uncharged tRNA interacts with the specifier sequence and the antiterminator structure, stabilizing the antiterminator structure, which prevents formation of the terminator helix and allows synthesis of the full-length transcript. T, terminator; AT, antiterminator; AA, amino acid. Modified from Grundy and Henkin (2004b).

Mutational analysis determined that the conserved elements in the 5’ region of the RNA transcript are required for efficient antitermination (Rollins et al., 1997). Addition of tRNA$^{\text{Gly}}$ to a purified *in vitro* transcription system triggers antitermination of the *B. subtilis glyQS* gene in the absence of any additional factors (Grundy et al., 2002b), indicating that the tRNA can interact directly with the leader RNA. It also appears that regulation of several T box genes occurs at the level of translation initiation, as formation
of the terminator helix could sequester the Shine-Dalgarno sequence (Grundy FJ, Henkin TM, unpublished). In contrast to the *E. coli* trp operon, which also senses tRNA charging ratios (see Section 1.1), the mechanism by which expression of the T box system is regulated is not ribosome-mediated.

1.3.3 Cellular metabolites

1.3.3.1 Thi box system

A group of riboswitches has been identified that regulate the expression of vitamin-related genes. A conserved RNA structural element (the THI box element) was identified upstream of genes involved in biosynthesis and transport of thiamin (vitamin B₁; Miranda-Rios *et al.*, 2001; Rodionov *et al.*, 2002). Regulation of the *Rhizobium etli* thiCOGE operon occurs at the level of premature termination of transcription *in vivo*, and a structural change in the leader region induced by thiamin blocks ribosome binding to the Shine-Dalgarno sequence, allowing the terminator helix (which overlaps the Shine-Dalgarno sequence) to form (Miranda-Rios *et al.*, 2001). Structural analysis of *E. coli* thiC leader RNA showed conformational changes in the RNA in response to thiamine pyrophosphate (TPP, a cofactor of key enzymes in carbohydrate metabolism), and it was proposed that regulation occurs at the levels of transcription termination and translation initiation (Winkler *et al.*, 2002b). Regulation of the *E. coli* thiM gene appears to occur at the level of translation initiation, by TPP-dependent sequestration of the Shine-Dalgarno sequence independent of transcription termination (Winkler *et al.*, 2002b; Fig. 1.9). TPP-dependent transcription termination in the leader region was demonstrated for the *B. subtilis* tenA gene *in vitro*, and mutations in the THI box completely abolish the
TPP-dependent response, indicating that interaction of TPP with the THI box element controls regulation (Mironov et al., 2002). As observed for other riboswitches, regulation of thiamine-related genes is primarily at the level of transcription termination in Gram-positive organisms and at the level of translation initiation in Gram-negative organisms (Rodionov et al., 2002; Mironov et al., 2002).

**Figure 1.9. The Thi box system.** In the absence of thiamine pyrophosphate, the Shine-Dalgarno sequence is free to interact with the ribosome, and translation occurs. In the presence of thiamine pyrophosphate, the Shine-Dalgarno sequence is sequestered from the ribosome, which inhibits translation initiation. SD, Shine-Dalgarno sequence; TPP, thiamine pyrophosphate. Modified from Soukup and Soukup (2004). In Gram-positive organisms, regulation of thiamine-related genes is primarily at the level of transcription termination, and binding of TPP induces termination by stabilizing the anti-antiterminator structure (Mironov et al., 2002).
1.3.3.2 RFN element

The RFN element consists of conserved RNA structure found upstream of genes involved in riboflavin biosynthesis, and it was proposed that regulation involves direct binding of flavin mononucleotide (FMN) to the RFN element (Gelfand et al., 1999). Comparative sequence analysis suggested that the mechanism for regulation by the RFN element was premature transcription termination or inhibition of translation initiation (Vitreschak et al., 2002). Expression of the *B. subtilis* rib operon is regulated at the level of transcription termination (Mironov et al., 2002; Winkler et al., 2002a), while regulation of the *B. subtilis ypaA* gene is at the level of translation initiation (Winkler et al., 2002a). FMN binds directly to the RFN element in the *rib* operon *in vitro*, resulting in FMN-dependent transcription termination, and mutations in conserved sequences in the RFN element eliminate FMN-dependent transcription termination (Mironov et al., 2002). Structural analysis of *rib* leader RNA demonstrated conformational changes in the RNA in response to FMN binding (Mironov et al., 2002; Winkler et al., 2002a). In Gram-negative organisms regulation of riboflavin-related genes appears to occur at the level of translation initiation (Vitreschak et al., 2002).

1.3.3.3 B12 element

Comparative sequence analysis was used to identify a conserved sequence, designated the B12 element, in the leader regions of genes involved in biosynthesis and transport of vitamin B12 (cobalamin; Vitreschak et al., 2003; Rodionov et al., 2003a). Expression of these genes in *Salmonella typhimurium* (Ravnum and Andersson, 1997) and *E. coli* (Nou and Kadner, 1998) is repressed at the level of translation initiation.
during growth in the presence of adenosylcobalamin (AdoCbl), mediated by secondary structure in the 5’ region of the RNA transcript that sequesters the Shine-Dalgarno sequence. AdoCbl inhibits ribosome binding to *E. coli* and *Salmonella typhimurium btuB* leader RNAs and induces structural changes in the RNA (Nou and Kadner, 2000). Direct binding of AdoCbl to *E. coli btuB* leader RNA and AdoCbl-dependent structural changes in the leader RNA that result in sequestration of the Shine-Dalgarno sequence and inhibition of ribosome binding were also demonstrated (Nahvi *et al*., 2002). It appears that regulation of cobalamin-related genes mediated by AdoCbl binding to the B₁₂ element occurs at the level of transcription termination in Gram-positive organisms (Vitreschak *et al*., 2003).

1.3.3.4 pyrG

A unique type of riboswitch regulates the expression of pyrimidine biosynthesis genes in Gram-positive organisms. Expression of the *B. subtilis* and *Lactococcus lactis* *pyrG* genes, which encode CTP synthetase, is subject to CTP-dependent regulation mediated by an intrinsic transcription terminator in the *pyrG* leader region (Meng and Switzer, 2002; Jørgensen *et al*., 2003). Mutational analysis demonstrated that the first four nucleotides at the 5’ end of the *pyrG* transcript (GGGC), the first six nucleotides of the 5’ strand of the terminator stem (GCUCCC), and the last six nucleotides of the 3’ strand of the terminator stem (GGGAGC), are required for antitermination induced by starvation for CTP. It was initially proposed that when intracellular CTP levels are low, a regulatory protein binds to these sequence elements and prevents premature termination of transcription (Meng and Switzer, 2002). A recent study revealed that regulation of
$pyrG$ expression in *B. subtilis* instead occurs by a novel control mechanism that utilizes reiterative transcription (Meng *et al.*, 2004). It appears that pyrimidine limitation results in a pause before insertion of the C residue at +4 in the $pyrG$ transcript, which facilitates reiterative addition of G residues, which allows base pairing between the poly(G) end of the transcript and a run of nine pyrimidines (CUCCCUUUC) in the 5’ strand of the terminator hairpin. This pairing sequesters residues required for formation of the terminator hairpin, which allows transcription of the downstream coding region (Meng *et al.*, 2004; Fig. 1.10). Regulation of $pyrG$ expression in other Gram-positive bacteria is likely to occur by an analagous mechanism (Meng *et al.*, 2004). The $pyrG$ riboswitch is unique in that the mechanism of regulation does not involve direct binding of the effector molecule (CTP) by the transcript, and it also unique in that the antiterminator sequence is not encoded in the $pyrG$ DNA but is instead generated by reiterative transcription in response to CTP limitation.
When CTP levels are high, normal transcription occurs, which results in formation of the terminator helix and premature termination of transcription. Low levels of CTP induce a transcriptional pause at position +4, and reiterative transcription occurs, which results in the addition of extra G residues. The extra G residues form the antiterminator structure by base pairing with residues required for formation of the terminator helix, allowing expression of the downstream coding region. T, terminator; AT, antiterminator. Modified from Meng et al. (2004).

1.3.3.5 G box system

A conserved RNA element designated the G box was identified in the 5’ untranslated region of genes involved in guanine biosynthesis and transport, primarily in low G+C Gram-positive organisms (Johansen et al., 2003; Mandal et al., 2003). Guanine binds to the G box in B. subtilis xpt-pbuX RNA in vitro and induces a structural change in the 5’ region of the RNA transcript, in agreement with the model that binding of guanine
promotes formation of a terminator helix. The G box functions as a guanine-dependent anti-antiterminator that promotes transcription termination in vivo (Johansen et al., 2003). The crystal structure of the guanine-binding domain from the B. subtilis spt-pbuX operon bound to hypoxanthine (a metabolite in the bacterial purine salvage pathway) revealed that the RNA forms a purine-binding pocket that is anchored by interaction of phylogenetically conserved nucleotides in the terminal loops. This interaction occurs in the presence and the absence of the effector, and is required for and stabilized by effector binding (Batey et al., 2004). The crystal structure of the B. subtilis guanine-sensing riboswitch complexed with guanine has also been solved. The guanine-bound complex differs from the hypoxanthine-bound complex only in the presence of two hydrogen bonds associated with 2-amino group recognition, as hypoxanthine lacks the 2-amino group present in guanine (Serganov et al., 2004).

Several G box variants that contain a C to U change in the sequence have been identified upstream of genes involved in adenine metabolism (Johansen et al., 2003; Mandal and Breaker, 2004a). The specificity of these RNA elements is altered such that they bind adenine instead of guanine. It appears that the single C vs. U variation in the purine-binding domain is sufficient to determine specificity of the riboswitch for guanine or adenine. Binding of adenine to the B. subtilis ydhL gene induces a structural rearrangement that prevents formation of the intrinsic transcription terminator helix and activates gene expression (Mandal and Breaker, 2004a). Activation of ydhL expression by binding of adenine is consistent with the proposed function of the ydhL gene product as a purine efflux pump (Johansen et al., 2003; Nygaard and Saxild, 2005).
The crystal structure of the *Vibrio vulnificus* adenine-sensing riboswitch bound to adenine has been solved (Serganov et al., 2004), and comparison of this crystal structure to that of the *B. subtilis* guanine-sensing riboswitch bound to guanine revealed nearly identical purine-binding pockets. As predicted by Mandal and Breaker (2004a), the pyrimidine residue that differs between the guanine- and adenine-sensitive riboswitch RNAs is the key specificity determinant for recognition of guanine vs. recognition of adenine. The crystal structures of the purine-sensing riboswitches bound to their respective ligands revealed Watson-Crick base-pairing between the purines and the specificity-determining pyrimidine residue in the binding domain (Serganov et al., 2004); this interaction was confirmed by NMR spectroscopy (Noeske et al., 2005).

1.3.3.6 GlcN6P ribozyme

A novel mechanism was recently identified for riboswitch control of expression of the *glmS* gene, which encodes the enzyme (glutamine-fructose-6-phosphate amidotransferase) responsible for synthesis of glucosamine-6-phosphate (GlcN6P) (Winkler et al., 2004; Fig. 1.11). GlcN6P activates a self-cleaving ribozyme that is present in the 5’ untranslated region of the *glmS* RNA in *B. subtilis* and many other Gram-positive organisms (Winkler et al., 2004; Barrick et al., 2004). Mutations that deactivate the ribozyme result in derepression of gene expression. The details of the mechanism of repression of gene expression by the ribozyme are not yet clear, but it is possible that it functions as a true allosteric ribozyme and induces structural alterations in the cleaved RNA in response to binding of GlcN6P. It is possible that the non-coding
RNA released as a 5’ cleavage fragment could act as a signal of adequate GlcN6P levels and influence glmS gene expression.

Figure 1.11. GlcN6P ribozyme. GlcN6P activates a self-cleaving ribozyme that is present in the leader region of glmS RNA, which prevents expression of the downstream coding region. GNP, glucosamine-6-phosphate. Modified from Soukup and Soukup (2004).

1.3.3.7 L box system

A group of riboswitches that regulate amino acid metabolic pathways, including the L box system, the glycine system, and the S box system, have also been identified. Expression of the B. subtilis lysC gene (which encodes aspartokinase II, the first enzyme in synthesis of lysine from aspartate) is repressed at the level of premature termination of
transcription during growth in the presence of lysine in vivo (Lu et al., 1992; Kochar and Paulus, 1996). Lysine-dependent transcription termination was demonstrated in vitro for B. subtilis lysC (Grundy et al., 2003; Sudarsan et al., 2003). Conserved primary sequence and structural elements, designated the L box, were identified in the leader region of lysC and several other lysine biosynthesis genes in Gram-positive and Gram-negative bacteria (Grundy et al., 2003). The model for lysine-dependent transcription termination is that binding of lysine promotes termination by stabilizing the proposed anti-antiterminator element, which prevents formation of the competing antiterminator structure, and a lysine-induced structural rearrangement in the lysC RNA consistent with the model was demonstrated (Grundy et al., 2003; Sudarsan et al., 2003). Additional comparative sequence analysis identified several other lysine-related genes in a variety of organisms that contain the L box element (Rodionov et al., 2003b). While expression of L box genes in Gram-positive organisms is regulated primarily at the level of transcription termination, it appears that expression of lysine-related genes from Gram-negative organisms is regulated at the level of translation initiation, by sequestration of the Shine-Dalgarno element (Grundy et al., 2003).

1.3.3.8 Glycine system

The most recently identified riboswitch controls expression of genes involved in glycine metabolism. The gcvT operon in B. subtilis encodes proteins involved in cleavage of glycine for use as a carbon and energy source (Kikuchi, 1973; Douce et al., 2001). Expression of this operon must be tightly controlled, as glycine is also essential for protein synthesis. An RNA motif containing two glycine-binding RNA elements was
identified upstream of this operon in *B. subtilis* and similar motifs were identified in other bacterial species, including *Vibrio cholerae*. These elements are arranged as tandem domains that are separated by a linker region that also exhibits some conservation of length and sequence, suggesting functional coupling of the metabolite-binding domains (Mandal *et al.*, 2004). Structural changes induced by glycine binding were detected in each *V. cholerae* RNA element independent of the other element, indicating that each of these elements acts as a separate glycine-binding domain. Binding of glycine was shown to be cooperative, as binding to one site improves binding affinity at the second site from ~100 to ~1,000 fold, which allows the bacteria to respond to a less significant increase in glycine concentration (Mandal *et al.*, 2004). Glycine induces expression of the *B. subtilis gcvT* operon both *in vivo* and *in vitro*, demonstrating that the default state of the glycine-dependent riboswitch is off, and that glycine binding is required to activate gene expression (Mandal *et al.*, 2004). Activation of *gcvT* operon gene expression by glycine is consistent with the roles of the *gcvT* operon gene products in cleavage of glycine for use as a carbon and energy source.

1.3.3.9 S box system

The S box transcription termination control system regulates the expression of many genes involved in sulfur metabolism, primarily in low G+C Gram-positive bacteria; there are 11 S box transcriptional units in *Bacillus subtilis* (Grundy and Henkin, 1998, 2002, 2003). Genes in the S box regulon are characterized by the presence of a set of highly conserved primary sequence and secondary structural elements in the untranslated leader region upstream of the regulated coding sequence. The secondary structure model
of the S box leader region was initially based on the arrangement of these elements and sequence covariation within helical domains (Grundy and Henkin, 1998; Fig. 1.12). Analysis of additional S box leader RNAs provided further confirmation of the model (Grundy and Henkin, 1998; Grundy and Henkin, 2003; Winkler et al., 2003; Rodionov et al., 2004).
Figure 1.12. *B. subtilis* yitJ leader structural model. Numbering is relative to the predicted transcription start-site (+1). The sequence is shown in the terminator conformation; red and blue residues illustrate the alternate pairing required for formation of the antiterminator, shown above the terminator. Asterisks indicate regions of covariance. Helices 1-5 are identified by boxed numbers; T, terminator; AT, antiterminator, AAT, anti-antiterminator.
A model for S box transcription termination control was proposed in which synthesis of the full-length transcript is determined by whether the leader folds into the stem-loop of an intrinsic terminator (when cells are grown in the presence of methionine) or a competing antiterminator structure (when methionine is limiting). Termination is dependent on formation and stabilization of an anti-antiterminator structure that competes with the antiterminator, as the antiterminator is more stable than the terminator (Grundy and Henkin, 1998, 2003; Fig. 1.13). It has been demonstrated for several *B. subtilis* S box genes (including *yitJ, ykrT, ykrW, yjcI, yusC, yxjG, and yxjH*) that growth in the presence of methionine results in efficient termination of transcription, while terminator readthrough is induced in response to starvation for methionine (Grundy and Henkin, 1998; Murphy *et al.*, 2002; Auger *et al.*, 2002; Hullo *et al.*, 2004). S-adenosylmethionine (SAM) is the molecular effector sensed by S box leader RNAs in the absence of any *trans*-acting factors (McDaniel *et al.*, 2003, Winkler *et al.*, 2003; Epshtein *et al.*, 2003; Chapter 2), identifying the S box system as a member of the subset of riboswitches that directly bind cellular metabolites. Few S box-regulated genes have been identified in Gram-negative organisms and expression of these genes appears to be regulated at the level of translation initiation, consistent with the pattern for other riboswitches (Grundy FJ, McDaniel BA, Henkin TM, unpublished).
Figure 1.13. S box antitermination model. When methionine levels are low, the anti-antiterminator structure is destabilized, allowing formation of the antiterminator structure and expression of the downstream coding regions. When methionine levels are high, the anti-antiterminator structure is stabilized, preventing formation of the antiterminator and allowing formation of the terminator and premature termination of transcription. The filled circle represents the effector molecule. T, terminator; AT, antiterminator, AAT, anti-antiterminator. Modified from Grundy and Henkin (2003).

1.4 Goals of this study

At the beginning of this project few details of the molecular mechanism of transcription termination control of S box gene expression were known. Grundy and Henkin (1998) had established that expression of genes in the S box regulon is coordinately regulated at the level of premature termination of transcription in response to methionine availability. Initially 18 transcriptional units were identified with leader
regions capable of forming secondary structures similar to those predicted for *B. subtilis yitJ*. A total of 26 residues were identified in the leader regions of these transcriptional units that are 100% conserved, while an additional 29 residues are greater than 90% conserved. The majority of these conserved residues are located in helices 1-4 of the leader RNA secondary structure (helix 1-4 region; Fig. 1.12). Initial mutational analysis of the *B. subtilis yitJ* leader indicated that disruption of conserved primary sequence (other than sequence required for formation of the antiterminator) and secondary structure elements in the helix 1-4 region results in a loss of repression of S box gene expression during growth in the presence of methionine. This led to a model in which the helix 1-4 region could act as the binding site for a negative regulatory factor that stabilizes the anti-antiterminator structure (helix 1) under repressing conditions for S box gene expression.

The focus of research described in the following chapters is elucidation of the molecular mechanism of transcription termination control governing the S box system. A variety of genetic and biochemical approaches were utilized to determine if any *trans*-acting factors play a role in S box gene regulation, and to examine the importance of conserved *cis*-acting sequence and structural features in the S box leader RNA. At the time this work began, none of the effector binding RNAs had been shown to function *in vitro*, and none of these systems had been demonstrated to operate by a direct interaction between the effector molecule and the leader RNA. A secondary goal of this project was to take advantage of the information provided by identification of the S box RNA pattern to investigate the physiological roles of *B. subtilis* S box-regulated genes of unknown function.
CHAPTER 2

IDENTIFICATION OF S-ADENOSYLMETHIONINE AS THE MOLECULAR EFFECTOR FOR TRANSCRIPTION TERMINATION IN THE S BOX SYSTEM

2.1 Introduction

The high level of primary sequence conservation in the helix 1-4 region of S box leader RNAs and initial genetic analysis of the \textit{B. subtilis yitJ} leader led to a model for S box transcription termination control (Grundy and Henkin, 1998; Chapter 1). Mutations that disrupt the leader region terminator helix result in high-level, constitutive expression, in agreement with data demonstrating that regulation is at the level of premature termination of transcription (Grundy and Henkin, 1998). Disruption of the antiterminator structure results in very low level expression, even when cells are starved for methionine, indicating that the terminator is always functional when the antiterminator is prevented from forming (Grundy and Henkin, 1998). Mutations that prevent formation of the anti-antiterminator structure but leave the antiterminator structure intact result in constitutive expression, indicating that the antiterminator competes very efficiently with the terminator helix, and that the anti-antiterminator must be stabilized to prevent formation
of the antiterminator and allow formation of the terminator helix (Grundy and Henkin, 1998; Chapter 3). Mutations in conserved primary sequence and secondary structural elements in the helix 1-4 region other than the anti-antiterminator structure itself also result in a loss of repression during growth in the presence of methionine, indicating that this region plays a role in methionine sensing (Grundy and Henkin, 1998; Winkler et al., 2001; Chapter 3). Based on this mutational analysis, a model for the S box regulatory mechanism was proposed in which methionine availability is monitored by binding of a regulatory factor to the helix 1-4 region of the leader RNA (Grundy and Henkin, 1998, 2003; Fig. 1.13).

It was proposed that SAM could serve as the metabolic signal in regulation of S box gene expression in B. subtilis (Grundy and Henkin, 1998, 2002, 2003), based on the observations that a 3- to 4-fold reduction in intracellular SAM pools (due to a decrease in SAM synthetase activity) leads to a 7-fold or greater increase in intracellular methionine levels (Wabiko et al., 1988), that overexpression of SAM synthetase results in methionine auxotrophy (Yocum et al., 1996), and that in E. coli SAM acts as a corepressor for control of methionine biosynthesis genes at the level of transcription initiation (Greene, 1996). We hypothesized that SAM could promote premature termination of transcription of S box genes by interacting directly with the S box leader RNA in the absence of any additional regulatory factors. This hypothesis was tested using halted complex in vitro transcription assays and SAM binding assays.
2.2 Materials and Methods

2.2.1 Generation of DNA templates

*B. subtilis* strain BR151MA (*lys-*3 *trpC*2) was the source of chromosomal DNA for amplification by PCR. Oligonucleotide primers were purchased from Integrated DNA Technologies (Coralville, IA). PCR products were purified with a Qiagen PCR cleanup kit (Qiagen, Chatsworth, CA).

DNA templates for *in vitro* transcription by *B. subtilis* or *E. coli* RNAP were generated by PCR using oligonucleotide primers that contained the *B. subtilis glyQS* promoter sequence (Grundy *et al.*, 2002b) and hybridized within the leader region of the target gene to generate a transcription start site ~20 nt upstream of the start of helix 1 (Fig. 1.1). The promoter sequences were designed to allow initiation with a dinucleotide corresponding to the +1/+2 positions of the transcript and a halt early in the transcript by omission of a single nucleotide; the position of the halt ranged from +15 to +25 relative to the start-site of transcription for different S box genes. The PCR fragments were ~400 bp in length and included 50-150 bp downstream from the transcription terminator to allow resolution of terminated and read-through products (McDaniel *et al.*, 2003).

DNA templates for T7 RNAP transcription were generated by PCR using a primer containing a T7 RNAP promoter initiating with tandem G residues fused to position +14 for the *yitJ* leader RNAs and position +2 for the *ykrW* leader RNAs. The endpoint of the PCR products corresponded to the position just 5’ to the start of the terminator helix (McDaniel *et al.*, 2003).
2.2.2  *In vitro* transcription assays

Single-round transcription reactions were carried out as described (Grundy *et al*., 2002b; McDaniel *et al*., 2003) with template DNA (10 nM), His-tagged *B. subtilis* RNAP (6 nM) purified as described by Qi and Hulett (1998), or *E. coli* RNAP (10 nM) purified as described by Hager *et al*., (1990) and provided by Dr. Irina Artsimovitch (The Ohio State University). The initiation reactions contained 1X transcription buffer (Grundy *et al*., 2002a), the appropriate dinucleotide (150 µM; Sigma), UTP (0.75 µM), [α^32P]UTP (800 Ci/mmol; 1 Ci = 37 GBq) at 0.25 µM, and the remaining two NTPs required for elongation to the halt position at 2.5 µM. The initiation reactions were incubated at 37°C for 15 min, then placed on ice. Heparin (20 µg/ml; Sigma) was added to block reinitiation, and elongation was triggered by the addition of NTPs to 10 µM final concentration in the presence of other reagents as indicated. Elongation reactions were incubated at 37°C for an additional 15 min and were terminated by extraction with phenol. Transcription products were resolved by denaturing polyacrylamide gel electrophoresis (PAGE) and visualized by PhosphorImager (Molecular Dynamics) analysis. Percent termination (%T) was calculated from the percentage of product in the terminated band relative to the total of the terminated and read-through products. The results obtained had a margin of error of ±5%.
2.2.3 SAM binding assays

SAM binding assays were carried out as described in McDaniel et al. (2003). T7 RNAP transcription was carried out using a MEGAshortscript T7 High Yield Transcription kit (Ambion). RNA (8 µM) was heated to 65°C for 5 min in 1X transcription buffer containing MgCl₂ (10 mM), then slow-cooled to 40°C before addition of [methyl-¹⁴C]-SAM (ICN; 52 mCi/mmol [1.92 GBq/mmol]; 8 µM final concentration) and incubation at room temperature for 5 min. Nonradioactive SAM or S-adenosylhomocysteine (SAH) was added as a competitor compound at 400 µM where indicated. Samples were filtered through a Nanosep 10K Omega filter microconcentrator (Pall) and washed 5 times with 150 µl 1X transcription buffer. Material retained by the filter was mixed with Packard Bioscience Ultima Gold scintillation fluid and counted in a Packard Tri-Carb 2100TR liquid scintillation counter. The results obtained had a margin of error of ±1%.

2.3 Results

2.3.1 In vitro transcription of S box genes

Purified in vitro transcription assays were performed using B. subtilis S box leader region DNA templates. Initial attempts at transcription of the B. subtilis yitJ and ykrW genes by B. subtilis RNAP resulted in weak transcription, due to the lack of similarity of the sequence in the -10 regions of these promoters to the consensus for B. subtilis Eσ²⁸ RNAP. Transcription of the B. subtilis ykrW gene with E. coli RNAP, which
is less stringent in promoter recognition, resulted in efficient transcription and terminator read-through (McDaniel et al., 2003; Fig. 2.1, lane 1).

**Figure 2.1. In vitro transcription of S box genes.** DNA templates were transcribed with *E. coli* (lanes 1-6) or *B. subtilis* (lanes 7-13) RNAP. Lanes 1 and 2, ykr*W* DNA; lanes 3, 4, 7, and 8, Pgly-ykrW DNA; lanes 5 and 6, Ptyr-yitJ DNA; lanes 9 and 10, Pgly-yitJ DNA; lanes 11-13, glyQS DNA. SAM was added at 150 µM where indicated (+). tRNA_Gly was added at 70 nM where indicated (+). The Pgly-yitJ constructs contain additional sequences downstream of the terminator, resulting in a larger read-through product than is observed with Ptyr-yitJ DNA. Open circles, terminated transcript; filled circles, read-through transcript. Percent termination (%T) is shown at the bottom of each lane.

It was previously demonstrated that replacement of the yitJ promoter has no effect on the response to starvation for methionine *in vivo* (Grundy and Henkin, 1998), so we
generated constructs in which the promoters of the S box genes were replaced with the strong *B. subtilis glyQS* promoter to allow transcription of these genes with *B. subtilis* RNAP. The region immediately downstream of the transcription start site was modified (where necessary) to allow a halt early in the transcription reaction by omission of a single nucleotide from the initiation reaction. Transcription of the P<sub>gly-ykrW</sub> construct with either *E. coli* or *B. subtilis* RNAP resulted in efficient terminator read-through (McDaniel *et al.*, 2003; Fig. 2.1, lanes 3 and 7). Transcription of the *B. subtilis yitJ* leader under control of the *B. subtilis tyrS* promoter (P<sub>tyr-yitJ</sub>) with *E. coli* RNAP also resulted in efficient terminator read-through (Fig. 2.1, lane 5), as did transcription of the *B. subtilis yitJ* leader under control of the *glyQS* promoter (P<sub>gly-yitJ</sub>) with *B. subtilis* RNAP (Fig. 2.1, lane 9), although slightly more terminated product was observed with *B. subtilis* RNAP.

### 2.3.2 SAM-dependent transcription termination

*In vitro* transcription assays were performed to determine if methionine, SAM, or related compounds could stimulate premature termination of transcription of S box gene expression. SAM-dependent transcription termination was observed for the P<sub>gly-ykrW</sub> construct transcribed with either *B. subtilis* or *E. coli* RNAP (McDaniel *et al.*, 2003; Fig. 2.1, lanes 4 and 8). Similar results were obtained for *ykrW* with its own promoter (Fig. 2.1, lane 2) and *yitJ* with the *tyrS* promoter (Fig. 2.1, lane 6), both transcribed by *E. coli* polymerase, as well as *yitJ* with the *glyQS* promoter transcribed by *B. subtilis* polymerase (Fig. 2.1, lane 10), confirming that the response to SAM is promoter-independent (McDaniel *et al.*, 2003).
Addition of SAM failed to promote transcription termination of the *glyQS* leader region terminator (McDaniel *et al.*, 2003; Fig. 2.1, lanes 11-13), which is inhibited by tRNA$_{	ext{Gly}}$ via the T box transcription termination control system (Grundy *et al.*, 2002). This result indicates that SAM does not generally promote transcription termination, but acts specifically on S box leader region terminators. The generality of the SAM response for the S box system was determined by analysis of 8 additional S box leaders in the presence and absence of SAM. All of the S box leaders examined exhibited SAM-dependent transcription termination *in vitro* in the absence of any additional factors (McDaniel *et al.*, 2003; Fig. 2.2, lanes 1-16), although some variability was observed in the efficiency of termination in the absence of SAM and in the degree of the response to SAM (McDaniel *et al.*, 2003).
Compounds related to SAM were tested for their ability to promote transcription termination of an S box leader \textit{in vitro}. Methionine, homocysteine, adenosine, SAH, and methylthioadenosine all failed to stimulate transcription termination of the \textit{B. subtilis} \textit{ykrW} leader (McDaniel \textit{et al}., 2003; Fig. 2.3), indicating that S box gene expression is specifically repressed by SAM. These compounds were also unable to inhibit transcription termination in the presence of a non-saturating concentration of SAM (7.5 µM) when added at a 200-fold molar excess (except homocysteine, which was added at a 67-fold molar excess) compared to SAM. Sinefungin, an analog of SAM produced by \textit{Streptomyces} \textit{sp.} in which ornithine is substituted for methionine (Boeck \textit{et al}., 1973), acts as a competitive inhibitor of many methyltransferase reactions (Schluckebier \textit{et al}., 1973).
yet it failed to promote transcription termination of ykrW (McDaniel et al., 2003; Fig. 2.3, lanes 15 and 17). Addition of sinefungin at a 200-fold molar excess resulted in a 4-fold increase in read-through in the presence of SAM (McDaniel et al., 2003; Fig. 2.3, lane 18), which suggests that sinefungin interacts with the ykrW leader with a much lower affinity than SAM. These results as a whole demonstrate that SAM recognition by the leader RNA is highly specific.

**Figure 2.3. Specificity of SAM-dependent transcription termination of the B. subtilis ykrW leader.** The ability of SAM-related compounds to stimulate termination of transcription by E. coli RNAP, and to block SAM-dependent termination, was tested. MET, methionine; ADO, adenosine; HCY, homocysteine; MTA, methylthioadenosine; SF, sinefungin. All compounds were tested at 1.5 mM except homocysteine (500 µM) and sinefungin, which was tested at 500 µM (low, lanes 15 and 16) and 1.5 mM (high, lanes 17 and 18). SAM was included at 7.5 µM where indicated (+). Open circles, terminated transcript; filled circles, read-through transcript. Percent termination (%T) is shown at the bottom of each lane.
2.3.3 Binding of $^{14}$C-SAM to \textit{B. subtilis} \textit{yitJ} and \textit{ykrW} leader RNAs

Binding of SAM to \textit{B. subtilis} \textit{yitJ} and \textit{ykrW} S box leader RNAs containing the helix 1-4 region was tested by using size-exclusion filtration. $^{14}$C-SAM was added to the leader RNA before the sample was completely cooled, to mimic \textit{in vivo} conditions by allowing folding of the leader RNA in the presence of SAM. $^{14}$C-SAM was retained by the size-exclusion filter in the presence of both \textit{yitJ} and \textit{ykrW} leader RNA (McDaniel \textit{et al.}, 2003; Fig, 2.4), indicating that SAM binds directly to the helix 1-4 region of these RNAs. An equimolar amount of $^{14}$C-SAM was added to the leader RNA, indicating that a significant molar excess of $^{14}$C-SAM is not required for SAM binding in this assay. Addition of a 50-fold molar excess of nonradioactive SAM inhibited $^{14}$C-SAM binding, whereas addition of nonradioactive SAH was unable to prevent $^{14}$C-SAM binding (McDaniel \textit{et al.}, 2003; Fig, 2.4), further demonstrating the specificity of the SAM-leader RNA interaction.
Figure 2.4. Binding of $^{14}$C-SAM by *B. subtilis* yitJ and ykrW leader RNAs. $^{14}$C-SAM (8 µM) and T7 RNAP-transcribed leader RNA (8 µM) were incubated together in 1X transcription buffer containing MgCl$_2$ (10 mM) in the presence or absence of an excess of nonradioactive competitor compound (SAM [filled bars] or SAH [hatched bars] at 400 µM). SAM binding is expressed as the percentage of $^{14}$C-SAM retained after filtration through a Nanosep 10K filter and 5 washes with 150 µl 1X transcription buffer, relative to the amount of $^{14}$C-SAM added to the binding reaction. SAM retention in the absence of leader RNA was <0.1%.

2.4 Discussion

The observation that expression of an S box transcriptional fusion in a methionine auxotroph is repressed during growth in the presence of methionine, but induced when the cells are starved for methionine, was initially the only clue to the physiological signal involved in regulation of expression of S box genes (Grundy and Henkin, 1998). However, it was not known whether methionine itself, or some other physiological signal affected by the level of methionine, is the molecular effector. The observation that a reduction in intracellular SAM pools due to decreased SAM synthetase activity leads to
an increase in intracellular methionine levels in *B. subtilis* (Wabiko *et al*., 1988), and overexpression of SAM synthetase results in methionine auxotrophy (Yocum *et al*., 1996), suggested that SAM could be the effector. Expression of methionine biosynthesis genes in *E. coli* is regulated at the level of transcription initiation, in part by intracellular SAM levels, which decrease when cells are starved for methionine (Greene, 1996). This study demonstrated that SAM is sensed by S box leader RNAs, in the absence of any *trans*-acting factors, to regulate gene expression at the level of premature termination of transcription. S box leader region terminators exhibited low activity in purified *in vitro* transcription assays with either *B. subtilis* or *E. coli* RNAP, and termination was promoted by addition of SAM at a physiologically relevant concentration. These results are in agreement with the observation that read-through is the default state of S box gene expression *in vivo*, as demonstrated for *B. subtilis* yitJ, ykrT, ykrW, yjcI, yusC, yxjG, and yxjH (Grundy and Henkin, 1998; Murphy *et al*., 2002; Auger *et al*., 2002; Hullo *et al*., 2004). Our findings were supported by subsequent studies (Epshtein *et al*., 2003; Winkler *et al*., 2003).

The SAM binding assays demonstrated that SAM binds specifically and directly to the helix 1-4 region of S box leader RNAs. The S box system is thus a member of the novel class of gene regulation systems in which expression is regulated via direct sensing of an effector molecule by the nascent transcript. These systems exhibit a remarkable ability to discriminate against compounds related to their particular effector molecule, as demonstrated for the L box system (Grundy *et al*., 2003; Sudarsan *et al*., 2003), the G box system (Johansen *et al*., 2003; Mandal *et al*., 2003), the RFN element (Mironov *et al*., 2002; Winkler *et al*., 2002a), the B12 element (Nahvi *et al*., 2002), the Thi box system
(Winkler et al., 2002b), and the glycine system (Mandal et al., 2004). We demonstrated that the S box system also effectively discriminates against closely related compounds, including SAH, which differs from SAM by a single methyl group and the associated positive charge, and sinefungin, an analog of SAM in which ornithine is substituted for methionine. Compounds related to SAM were unable to act as competitive inhibitors of SAM binding or SAM-directed transcription termination, although sinefungin added at a 200-fold molar excess to SAM resulted in a slight (4-fold) increase in read-through in the presence of SAM. Highly specific recognition of the effector molecule is required for discrimination from related compounds, which is essential for an appropriate regulatory response in vivo.

Genetic systems that sense cellular metabolites must also respond to physiologically relevant concentrations of the effector molecule. In related experiments performed by Frank Grundy, it was demonstrated that 1.6 µM SAM gave 60% termination of ykrW in vitro, while 100 µM SAM was a saturating level, resulting in >95% termination (McDaniel et al., 2003). Wabiko et al. (1988) reported that SAM pools in B. subtilis are 400 µM for cells grown in the presence of methionine, and 80 µM for methionine prototrophic cells grown in the absence of methionine. S box genes are expressed at a low level in methionine prototrophic cells grown in the absence of methionine, but are not fully repressed (Grundy and Henkin, 1998; McDaniel et al., 2003; Chapter 4). The SAM concentration required to promote transcription termination of S box gene expression in vitro is therefore within the physiologically relevant range for B. subtilis (McDaniel et al., 2003).
A previous study indicated that expression of the B. subtilis cysH operon, which contains an S box leader element, is regulated primarily at the level of transcription initiation in response to O-acetyl-serine, a direct precursor of cysteine, instead of at the level of transcription termination in response to methionine (Mansilla et al., 2000). However, our in vitro transcription assay results indicate that the cysH leader is capable of a response to SAM. The cysH leader exhibited a weaker response to SAM than the other S box leaders, which is consistent with the roles of genes in the cysH operon in cysteine biosynthesis instead of the primary methionine biosynthetic pathway or alternate routes to biosynthesis of methionine. The yoaD leader, which exhibited the weakest response to SAM, differs from all other S box leaders identified to date in that it has a tetraloop in place of the larger helix at the top of helix 3 in all other S box leaders, and an extra stem loop between the base of helix 1 and the terminator (Grundy and Henkin, 1998; Chapter 3). The function of the yoaD gene is unknown, although it is related to serA, which is involved in serine biosynthesis (Grundy and Henkin, 1998; Grundy and Henkin, 2002). Similarity to serA suggests that yoaD might also respond primarily to a regulatory signal other than SAM. Other S box leaders have been identified that have a cruciform structure at the top of helix 3, including ykrT, expression of which responds normally to starvation for methionine (Grundy and Henkin, 1998; Murphy et al., 2002; Chapter 5). It therefore appears that significant variability is tolerated in helix 3.

As a whole, our findings support a model in which SAM binds directly to the anti-antiterminator, stabilizing it, resulting in premature termination of transcription at the leader region terminator (McDaniel et al., 2003; Chapter 6). The anti-antiterminator is destabilized when SAM is not available for binding, allowing formation of the
antiterminator, which results in readthrough of the downstream coding regions by preventing formation of the terminator helix. This model is supported by structural analysis of in vitro transcribed S box leader RNAs (McDaniel et al., 2003; Epshtein et al., 2003; Winkler et al., 2003; Chapter 3). SAM-induced structural changes in the antiantiterminator domain consistent with the model were detected by antisense oligonucleotide-dependent RNase H cleavage assays, performed by Irina Artsimovitch (McDaniel et al., 2003), as well as structural mapping with RNase A (Chapter 3). It also appears that the SAM-leader RNA interaction is stable in vitro, as ¹⁴C-SAM bound to S box leader RNA was retained in the SAM binding assays after repeated washes. It is possible, however, that although the leader RNA interacts with SAM under the SAM binding assay conditions, additional factors could be involved in this interaction or could facilitate folding of the leader RNA in vivo.
CHAPTER 3

ANALYSIS OF CIS-ACTING SEQUENCE AND STRUCTURAL ELEMENTS REQUIRED FOR S-ADENOSYLMETHIONINE-DIRECTED TRANSCRIPTION TERMINATION

3.1 Introduction

The *B. subtilis* yitJ leader RNA can fold into the stem loop of an intrinsic terminator or a competing antiterminator structure. Formation of the antiterminator structure is inhibited by an additional helical structure that competes with the antiterminator and functions as an anti-antiterminator (Grundy and Henkin, 1998; Fig. 1.12). All S box leaders identified to date are capable of forming similar secondary structures, and exhibit high conservation of primary sequence which is concentrated in the helix 1-4 region of the leader RNA. Initial mutational analysis of the *yitJ* leader indicated that the helix 1-4 region could act as the binding site for a negative regulatory factor that stabilizes the anti-antiterminator structure (helix 1) during growth in the presence of methionine (Grundy and Henkin, 1998, 2003; Fig. 1.13). As described in Chapter 2, we identified SAM as the molecular effector for S box transcription
termination. Having demonstrated that SAM directly binds to and induces a conformational change in the leader RNA, and that termination in a purified in vitro transcription system is specifically promoted by addition of SAM in the absence of any trans-acting factors (McDaniel et al., 2003), our next objective was to investigate the importance of conserved cis-acting sequence and structural features in the helix 1-4 region of S box leader RNAs in the SAM/leader RNA interaction, and SAM-directed transcription termination in vivo and in vitro.

Although SAM binding assays demonstrated that SAM binds specifically and directly to the helix 1-4 region of S box leader RNAs (McDaniel et al., 2003; Chapter 2), the specific cis-acting sequence and structural features in this region that are important for the SAM/leader RNA interaction were unknown. Comparison of S box leader RNAs revealed some variability in the conserved sequence elements in the helix 1-4 region, including the length and complexity of helix 3 (Grundy and Henkin, 1998). Further analysis of the variability in the helix 1-4 region could be used as a tool to identify the minimal SAM binding site.

It was previously demonstrated that two yitJ leader variants, yitJ-Pst-1 and yitJ-Pst-2, result in loss of repression of S box gene expression during growth in the presence of methionine (Grundy and Henkin, 1998). The observed derepression of S box gene expression suggested that the cis-acting sequence and structural features disrupted in these variants are required for terminator activity. However, the effect of these mutations on SAM binding and SAM-directed transcription termination in vitro was unknown.

One of the conserved cis-acting structural features in S box leaders RNAs is the GA motif found in helix 2, which consists of two short helices separated by an
asymmetrical internal loop with conserved GA dinucleotides on either side of the internal loop (Winkler et al., 2001). This motif fits the pattern of a kink-turn element, which causes a kink in the phosphodiester backbone and can mediate RNA tertiary structure interactions (Klein et al., 2001). The GA motif is highly conserved in S box leader RNAs, although its importance in S box gene regulation was not known. This motif is also highly conserved in T box leader sequences, where it was shown to be important for tRNA-directed antitermination (Winkler et al., 2001), and is present in the lysine-binding L box sequences in a position in helix 2 similar to that found for the S box RNAs (Grundy et al., 2003).

The kink-turn motif in helix 2 could facilitate potential tertiary interactions in S box leader RNAs that could play a role in S box gene regulation. Covariation analysis is commonly used to identify both secondary and tertiary structural elements (James et al., 1988; Michel and Westhof, 1990; Woese and Pace, 1993). Covariation of nucleotides in the terminal loop of helix 2 and the unpaired region between helices 3 and 4 was observed in all S box leaders, and it was postulated that this covariation could indicate a tertiary interaction between these elements (Grundy and Henkin, 1998). Although sequence covariation and the presence of the kink-turn motif indicated a possible interaction between these sequence elements in the tertiary structure of S box leader RNAs, this hypothesis had yet to be experimentally tested. In experiments described in this chapter, we used in vivo and in vitro assays to determine the effect of mutations in conserved cis-acting sequence and structural features in S box leader RNAs on the SAM/leader RNA interaction, SAM-induced structural changes in the leader RNA, and SAM-directed transcription termination in vivo and in vitro.
3.2 Materials and Methods

3.2.1 Bacterial strains and growth conditions

*B. subtilis* strain BR151MA (*lys*-3 *trpC2*) was the source of chromosomal DNA for amplification by PCR. Oligonucleotide primers were purchased from Integrated DNA Technologies (Coralville, IA). Mutations in the *yitJ* leader were generated by PCR using oligonucleotides containing the desired sequence changes. Three pairs of complementary oligonucleotides were used to generate mutations in the *ykrW* leader sequence as previously described (Yousef *et al.*, 2003). Each pair of oligonucleotides was designed to have a 4-nt 3’ overhang complimentary to the 5’ overhang of the next pair. The 5’ ends of all internal oligonucleotides were phosphorylated with T4 polynucleotide kinase (United States Biochemicals) using the conditions recommended by the manufacturer. Complementary pairs of oligonucleotides (60 pmole in 70 mM Tris-HCl, pH 7.6, 10 mM MgCl₂, 5 mM DTT, 50 µM rATP) were mixed, heated to 95°C for 5 min, and slow-cooled to room temperature for 70 min. Sets of oligonucleotides that formed the template for each *ykrW* variant (10 pmole each) were then mixed, incubated at 45°C for 5 min, and allowed to cool to room temperature for 30 min. T4 DNA ligase (800 U, New England Biolabs) was added and the mixture was incubated overnight at 16°C. The resulting DNA was used as a template for PCR amplification using the outside primers. PCR products were purified by a Qiagen PCR cleanup kit (Qiagen, Chatsworth, CA) and all mutations were verified by DNA sequencing (Genewiz Inc., North Brunswick, NJ).

Wild-type and mutant *yitJ-lacZ* transcriptional fusions including the *yitJ* promoter were generated in plasmid pFG328 (Grundy *et al.*, 1993) and integrated in single copy
into the chromosome of strain BR151 (*lys-3 metB10 trpC2*) using specialized transducing phage SPβ, as previously described (Grundy and Henkin, 1993; Grundy and Henkin, 1998). Strains were propagated on tryptone blood agar base medium (Difco) or in Spizizen minimal medium (Anagnostopoulos and Spizizen, 1961). Strains containing *lacZ* fusions were grown in the presence of chloramphenicol (5 µg/ml).

### 3.2.2 *In vitro* transcription assays

*In vitro* transcription assays were carried out as described in Chapter 2. DNA templates for *in vitro* transcription were generated by PCR using oligonucleotide primers that contained the *B. subtilis glyQS* promoter sequence (Grundy *et al*., 2002b) and hybridized within the leader region of *yitJ* or *ykrW* to generate a transcription start site 17 or 11 nt upstream of the start of helix 1 for *yitJ* or *ykrW*, respectively (Fig. 3.7). The fusion junctions were designed to allow initiation with a dinucleotide (ApC) corresponding to the +1/+2 positions of the transcript and a halt at position +16 for *yitJ* by omission of GTP and at position +12 for *ykrW* by omission of CTP (McDaniel *et al*., 2003). The *yitJ* template (479 bp) included 248 bp downstream of the termination site; the *ykrW* template (234 bp) included 20 bp downstream of the termination site. Templates for *in vitro* transcription of the *yitJ* and *ykrW* leader region variants were generated by PCR with DNA templates containing the appropriate alleles. The initiation reactions contained 1X transcription buffer (Grundy *et al*., 2002a), ApC (150 µM; Sigma), UTP (0.75 µM), [α<sup>32</sup>P]UTP (800 Ci/mmol; 1 Ci = 37 GBq) at 0.25 µM, and the remaining nucleotides required for elongation to the halt position (CTP and ATP for *yitJ*, GTP and ATP for *ykrW*) at 2.5 µM. The initiation reactions were incubated at 37°C for
15 min, then placed on ice. Heparin (20 µg/ml; Sigma) was added to block reinitiation, and MgCl₂ was added to 30 mM at the same time heparin was added. Elongation was triggered by the addition of NTPs to 10 µM final concentration in the presence of other reagents as indicated. Elongation reactions were incubated at 37°C for an additional 15 min and were terminated by extraction with phenol. Transcription products were resolved by denaturing polyacrylamide gel electrophoresis (PAGE) and visualized by PhosphorImager (Molecular Dynamics) analysis. Percent termination (%T) was calculated from the percentage of product in the terminated band relative to the total of the terminated and read-through products. Reproducibility was ±5%.

3.2.3 SAM binding assays

SAM binding assays were performed as described in Chapter 2. DNA templates for T7 RNAP transcription were generated by PCR utilizing a primer containing a T7 RNAP promoter initiating at tandem G residues fused to position +14 for the yitJ leader RNAs, position +2 for the ykrW leader RNAs, and position +25 for the yoaD leader RNAs. The endpoint of the yitJ product was +166 (just 5’ to the start of the terminator helix). The endpoint of the ykrW product was +125 (four nt 3’ to the base of helix 1; the endpoint of the ykrW-ΔI variant was +115, four nt 3’ to the base of helix 4). The endpoint of the yoaD product was +129 (three nt 5’ to the base of the extra stem loop); for the yoaD variant including the extra stem loop the endpoint was +175 (just 3’ to the base of the extra stem loop).
3.2.4 β-galactosidase measurements

Cells containing lacZ fusions were grown in Spizizen minimal medium containing methionine (50 µg/ml) until early exponential growth and were then harvested by centrifugation and resuspended in the same medium in the presence or absence of methionine. Samples were collected at 1 h intervals and were assayed for β-galactosidase activity as described by Miller (1972) using toluene permeabilization of the cells. β-galactosidase assays were carried out in triplicate, and variation was <10%.

3.2.5 Leader RNA structural analysis

RNAs synthesized by T7 RNAP transcription were gel-purified (Yousef et al., 2003) and 5’ end-labeled with [γ-32P]ATP (6000 Ci/mmol) using a KinaseMax kit (Ambion). RNA was passed through a MicroSpin G-50 column (Amersham Biosciences) to remove unincorporated nucleotide, ethanol precipitated, and resuspended in water. Labeled RNAs (0.2 µM) were heated and slow-cooled in 1X transcription buffer as described for the SAM binding assays. SAM was added to 1.5 µM final concentration where indicated, and the samples were incubated at room temperature for 5 min. RNase A (0.1 pg/µl, Ambion) was added and samples were then incubated at 37°C for 1 to 15 min, ethanol precipitated, and resuspended in gel loading buffer. Products were resolved by denaturing polyacrylamide gel electrophoresis (6% and 10%) and visualized by PhosphorImager (Molecular Dynamics) analysis. Size standards were generated by digestion of denatured RNAs with RNase T1 (0.1 U/µl Ambion).
3.3 Results

3.3.1 Determination of the minimal SAM binding site in \textit{ykrW} and \textit{yoaD} leader RNAs

We wanted to determine the minimal binding site in the helix 1-4 region of S box leader RNAs to identify which \textit{cis}-acting structural features in this region are important for the SAM/leader RNA interaction, as comparison of S box leader RNAs revealed variability within this region (Grundy and Henkin, 1998). Deletion analyses were carried out for \textit{ykrW}, the product of which participates in the 5’-methylthioadenosine recycling pathway (Murphy \textit{et al.}, 2002; Grundy and Henkin, 2002; Ashida \textit{et al.}, 2003; Chapter 5; Fig. 3.1A), and \textit{yoaD}, an S box gene of unknown function. We chose the \textit{yoaD} leader because it differs from all other S box leaders identified to date in that it possesses a tetraloop in place of the larger helix at the top of helix 3 in all other S box leaders, and an extra stem loop between the base of helix 1 and the terminator (Grundy and Henkin, 1998; Fig. 3.1B). We chose the \textit{ykrW} leader because it has the smallest helix 3 of all the remaining S box leaders. Other S box leaders have been identified that have a cruciform structure at the top of helix 3 (Grundy and Henkin, 1998), which suggests that significant variability is tolerated in helix 3. Three deletion mutants of the \textit{ykrW} leader were generated: \textit{ykrW-\Delta I}, in which the 3’ half of helix 1 is deleted; \textit{ykrW-\Delta III}, in which the top half of helix 3 is deleted; and \textit{ykrW-\Delta IV}, in which helix 4 is deleted (Fig. 3.1A). In the \textit{ykrW-\Delta III} variant, the region at the top of helix 3 was replaced with the \textit{yoaD} helix 3 tetraloop sequence (Fig. 3.1A). The region deleted at the top of helix 3 is variable in the S box leaders, unlike the sequence at the bottom of helix 3, which is conserved. All of
the $ykrW$ deletion mutants exhibited loss of SAM binding \textit{in vitro} (Fig. 3.2A), indicating that helices 1, 3, and 4 in the $ykrW$ leader are required for SAM binding.
Figure 3.1. Secondary structure models of ykrW and yoaD leader RNAs. Numbering of residues is relative to the predicted transcription start point (+1). The sequence is shown in the terminator conformation; red and blue residues illustrate the alternate pairing required for formation of the antiterminator, shown above the terminator. Asterisks indicate regions of covariance. Helices 1-4 are identified by boxed numbers; AAT, anti-antiterminator; AT, antiterminator; T, terminator. A) B. subtilis ykrW. Boxed residues were deleted in the ykrW-ΔIII and ykrW-ΔIV leader variants; sequence introduced to replace the deleted residues in the ykrW-ΔIII variant are shown with lowercase letters; the normal T7 RNAP transcription product endpoint for ykrW and the endpoint for the ykrW-ΔI variant are denoted by arrows. B) B. subtilis yoaD. The T7 RNAP transcription product endpoint for yoaD and the endpoint for the variant including the extra stem loop are denoted by arrows.
Figure 3.1
Figure 3.2. Binding of $^{14}$C-SAM by ykrW and yoaD leader RNAs. $^{14}$C-SAM (8 µM) and T7 RNAP-transcribed leader RNA (8 µM) were incubated together in 1X transcription buffer containing MgCl$_2$ (10 mM) in the presence or absence of an excess of nonradioactive SAM (filled bars). SAM binding is expressed as the percentage of $^{14}$C-SAM retained after filtration through a Nanosep 10K filter and 5 washes with 150 µl 1X transcription buffer relative to the amount of $^{14}$C-SAM added to the binding reaction. SAM retention in the absence of leader RNA was <0.1%. A) ykrW leader RNA. B) yoaD leader RNA.
No SAM binding was observed (Fig. 3.2B) with a \textit{yoaD} construct missing the additional stem loop before the terminator (Fig. 3.1B). This is in agreement with the loss of SAM binding observed for the \textit{ykrW-\DeltaIII} variant, which has a truncated helix 3 analogous to that of \textit{yoaD}. However, an intact \textit{yoaD} leader (\textit{yoaD} + stem loop; Fig. 3.1B) exhibited approximately eight-fold higher SAM binding than the wild-type \textit{ykrW} construct (Fig. 3.2B). These results demonstrate that the extra stem loop in \textit{yoaD} is required for SAM binding by the \textit{yoaD} leader, and suggest that the interaction between SAM and the \textit{yoaD} leader is more stable than the interaction between SAM and the \textit{ykrW} leader.

\subsection*{3.3.2 \textit{yitJ} leader mutations affect SAM-directed transcription termination}

Two \textit{yitJ} leader variants, \textit{yitJ-Pst-1} and \textit{yitJ-Pst-2}, were previously generated that result in read-through of the leader region terminator during growth in the presence of methionine (Grundy and Henkin, 1998). In the \textit{yitJ-Pst-1} variant, sequence changes were introduced in the 5’ side of helix 1 that were predicted to prevent formation of the anti-antiterminator without affecting the antiterminator (Grundy and Henkin, 1998; Fig. 3.3). In the \textit{yitJ-Pst-2} variant, sequence changes were made in conserved residues in helix 2, leaving both the anti-antiterminator (helix 1) and the antiterminator intact (Grundy and Henkin, 1998; Fig. 3.3). Both the \textit{yitJ-Pst-1} and \textit{yitJ-Pst-2} variants resulted in high read-through \textit{in vitro} in the presence or the absence of SAM, demonstrating loss of SAM-directed transcription termination (McDaniel \textit{et al.}, 2003; Fig. 3.4). These results are consistent with the loss of repression during growth in the presence of methionine observed \textit{in vivo}. 

67
Figure 3.3. *B. subtilis* *yitJ* Pst leader variants. Numbering of residues is relative to the predicted transcription start point (+1). The sequence is shown in the terminator conformation; red and blue residues illustrate the alternate pairing required for formation of the antiterminator, shown above the terminator. Helices 1-4 are identified by boxed numbers; AAT, anti-antiterminator; AT, antiterminator; T, terminator. Sequence changes introduced in the Pst-1 and Pst-2 alleles are shown with lowercase letters.
**Figure 3.4. In vitro transcription of B. subtilis yitJ Pst leader variants.** Read-through (RT) and terminated (T) transcripts are labeled. Percent termination (%T) is shown at the bottom of each lane. Templates were transcribed with *E. coli* RNAP. SAM was added at 150 µM (+). Lanes 1 and 2, wild-type *yitJ* template; lanes 3 and 4, *yitJ*-Pst-1; lanes 5 and 6, *yitJ*-Pst-2.

The *yitJ*-Pst-2 variant was tested for ability to bind SAM *in vitro* to determine the effect of mutations in conserved residues in the helix 1-4 region outside of the anti-antiterminator itself on SAM binding. Binding of SAM by *yitJ*-Pst-2 leader RNA was reduced approximately twenty-fold compared to wild-type *yitJ* leader RNA (McDaniel *et al.*, 2003; Fig. 3.5). This reduction in SAM binding is consistent with the observations that this variant demonstrated loss of repression during growth in the presence of methionine and loss of SAM-dependent termination *in vitro*. The residual binding of
SAM by the *yitJ*-Pst-2 leader RNA was specific, as addition of nonradioactive SAM prevented $^{14}$C-SAM binding, while addition of SAH had no effect.

**Figure 3.5. Binding of $^{14}$C-SAM by *yitJ*-Pst-2 leader RNA.** $^{14}$C-SAM (8 µM) and T7 RNAP-transcribed leader RNA (8 µM) were incubated together in 1X transcription buffer containing MgCl$_2$ (10 mM) in the presence or absence of nonradioactive SAM (filled bars) or SAH (hatched bars) at 400 µM. SAM binding is expressed as the percentage of $^{14}$C-SAM retained after filtration through a Nanosep 10K filter and 5 washes with 150 µl 1X transcription buffer relative to the amount of $^{14}$C-SAM added to the binding reaction. SAM retention in the absence of leader RNA was <0.1%.
3.3.3 Role of the kink-turn motif in S box gene regulation in vivo

The kink-turn motif causes a kink in the phosphodiester backbone and can mediate RNA tertiary structure interactions (Klein et al., 2001), but the role of this motif in S box gene regulation was not known. Site-directed mutations were therefore introduced into the kink-turn element of a yitJ-lacZ transcriptional fusion to determine the importance of the conserved sequence and structural features. The G50C and C58G mutations (Winkler et al., 2001; Fig. 3.6A), which should prevent pairing of these two bases and disrupt formation of helix 2, resulted in loss of repression during growth in the presence of methionine, and reduced expression in the absence of methionine (Table 3.1). The G50C/C58G double mutant (Fig. 3.6A), which restores pairing but not primary sequence in helix 2, partially restored repression during growth in the presence of methionine, indicating that both the formation of helix 2 and its primary sequence are important for regulation. These results support the observation that elements within helix 2 that fit the pattern for a kink-turn motif are important for S box gene regulation in vivo.
Figure 3.6. Kink-turn motif sequence and consensus pattern. A) *B. subtilis yitJ* leader kink-turn motif sequence. The *yitJ* kink-turn motif is shown in the hatched box. Arrows indicate mutations introduced. B) S box kink-turn motif consensus pattern. Consensus pattern was derived from 49 S box leader sequences. Positions are numbered according to the intact *yitJ* leader. Boxes indicate >90% conservation, circles indicate >80% conservation, underlined positions indicate >70% conservation. R=G, A; Y=C, U; D=G, A,U; N = any nucleotide.
Figure 3.6

A) vitJ leader

B) S box GA motif consensus
Several similarities were noted in the consensus patterns of the kink-turn motifs found in S box and T box leaders (Winkler et al., 2001). Phylogenetic analysis determined that position C62 in the S box leader kink-turn motif exhibited a pattern of conservation (74% U, 14% C, 12% A, 0% G) similar to that of the analogous base, A21, in the T box leader kink-turn motif (80% U, 14% A, 4% C, 2% G; Winkler et al., 2001). A mutation made at this position in the *B. subtilis tyrS* T box leader (A21G) results in a

**Table 3.1. Effect of kink-turn motif mutations on expression of yitJ-lacZ transcriptional fusions.** Fusions were integrated in single copy in strain BR151 (*lys-3 metB10 trpC2*). Cells were grown in Spizizen minimal medium (Anagnostopoulos and Spizizen, 1961) containing methionine, harvested by centrifugation, and resuspended in Spizizen minimal medium in the presence (+ Met) or absence (- Met) of methionine. Samples for the β-galactosidase assay were taken 4 h after the cultures were split. β-galactosidase activity is expressed in Miller units (Miller, 1972).

<table>
<thead>
<tr>
<th>Fusion</th>
<th>β-galactosidase activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>- Met</td>
</tr>
<tr>
<td>wild-type</td>
<td>330</td>
</tr>
<tr>
<td>yitJ-G50C</td>
<td>140</td>
</tr>
<tr>
<td>yitJ-C58G</td>
<td>110</td>
</tr>
<tr>
<td>yitJ-G50C/C58G</td>
<td>240</td>
</tr>
<tr>
<td>yitJ-C62G</td>
<td>470</td>
</tr>
</tbody>
</table>
severe defect in tRNA<sub>Tyr</sub>-directed antitermination (Winkler et al., 2001). An equivalent mutation (C62G; Fig. 3.6A) in the yitJ leader had no significant effect on expression in vivo (Table 3.1), demonstrating that the requirements at this position differ for T box and S box leaders.

### 3.3.4 yitJ and ykrW leader covariation mutants

Covariation of at least two nucleotides in the terminal loop of helix 2 (designated “region 1”) with nucleotides in the unpaired region between helices 3 and 4 (designated “region 2”; Fig. 3.7) was observed in all S box leaders identified to date; 96% of S box leaders have at least three covarying residues, and 66% have at least four covarying residues (Grundy and Henkin, 1998; unpublished results). This pattern of covariation suggests a possible tertiary interaction between these two regions. The most common potential pairing, 5’-CNGG-3’ in region 1 and 5’-CCNG-3’ in region 2, is found in 48% of S box leader sequences identified to date, with N being a U in region 1 and an A in region 2 in 54% of leaders in this class, including yitJ (Fig. 3.7A). The covarying residues in the ykrW leader (5’-CGAG-3’ in region 1 and 5’-CUUG-3’ in region 2; Fig. 3.7B) represent an alternate pattern. To test the hypothesis that these residues interact with each other in the three-dimensional structure, we replaced residues in region 1 of the yitJ leader with the corresponding ykrW residues to generate the yitJ-W1 (U54G/G55A) variant, and replaced residues in region 2 of the yitJ leader with the corresponding ykrW residues to generate the yitJ-W2 (C116U/A117U) variant. Both sets of residues were replaced for the yitJ-W1W2 variant (U54G/G55A/C116U/A117U), so that base-pairing between regions 1 and region 2 was predicted to be restored (Fig. 3.7A). We also
introduced an A113U mutation into the yitJ-W2 and yitJ-W1W2 constructs, so that the sequence in the region 2 loop exactly matches that of ykrW (Fig. 3.7A). Similarly, residues in regions 1 or 2 of the ykrW leader were replaced with the corresponding yitJ sequences to generate the ykrW-J1 (G48U/A49G) and ykrW-J2 (U96C/U97A) variants, and both sets of mutations were introduced in the ykrW-J1J2 variant (G48U/A49G/U96C/U97A; Fig. 3.7B). In each case, the W1, W2, J1 and J2 nomenclature indicates the source of the sequence in regions 1 or 2. The yitJ wild-type leader therefore has the same sequences in regions 1 and 2 as the ykrW-J1J2 variant, while yitJ-W1 matches ykrW-J2, yitJ-W2 matches ykrW-J1, and yitJ-W1W2 matches the wild-type ykrW leader in regions 1 and 2. The sequences present in regions 1 and 2 and potential pairing between these sequence elements in each of these constructs are shown in Figure 3.8. We predicted that variants with mutations that could affect base-pairing (yitJ-W1, yitJ-W2, ykrW-J1 and ykrW-J2) would exhibit reduced stability of pairing between regions 1 and 2. We also predicted that pairing should be restored in the variants with compensatory mutations (yit-W1W2 and ykrW-J1J2).
Figure 3.7. Secondary structure models of \( yitJ \) and \( ykrW \) S box leader RNAs. Numbering is relative to the predicted transcription start-site (+1). The sequences are shown in the terminator conformation; red and blue residues illustrate the alternate pairing required for formation of the antiterminator, shown above the terminator. Helices 1-5 are identified by boxed numbers; AAT, anti-antiterminator; AT, antiterminator; T, terminator; K-turn, kink-turn motif. Green and purple residues (bold) denote positions of covariation in region 1 (terminal loop of helix 2) and region 2 (unpaired region between helices 3 and 4), respectively. Arrows indicate sequence alterations generated in this study. A. \( B. \textit{subtilis} \) \( yitJ \). W1 indicates sequence changes introduced into region 1 to generate the \( yitJ\)-W1 variant, which matches the sequence found in region 1 of the \( ykrW \) leader; W2 indicates sequence changes introduced into region 2 to generate the \( yitJ\)-W2 variant, which matches the sequence found in region 2 of the \( ykrW \) leader. The \( yitJ\)-W1W2 variant contains the sequence alterations in both regions 1 and 2, to restore the pairing pattern found in the wild-type \( ykrW \) leader. The A113U mutation was introduced into the \( yitJ\)-W1 and \( yitJ\)-W1W2 variants to replace the sequence in the entire region 2 internal loop with that of \( ykrW \). B. \( B. \textit{subtilis} \) \( ykrW \). J1 indicates sequence changes introduced into region 1 to generate the \( ykrW\)-J1 variant, which matches the sequence found in region 1 of the \( yitJ \) leader; J2 indicates sequence changes introduced into region 2 to generate the \( ykrW\)-J2 variant, which matches the sequence found in region 2 of the \( yitJ \) leader. The \( ykrW\)-J1J2 variant contains the sequence alterations in both regions 1 and 2, to restore the pairing pattern found in the wild-type \( yitJ \) leader.
Figure 3.7
3.3.4.1 In vitro transcription of yitJ leader covariation mutants

In vitro transcription assays were performed to determine the effect of these leader mutations on SAM-directed transcription termination. Transcription of wild-type yitJ and ykrW leader constructs under the control of the B. subtilis glyQS promoter with B. subtilis RNAP resulted in inefficient termination at the leader region terminator, while

---

**Figure 3.8. Covarying regions in S box leader RNAs.** A) Positions of covarying residues in the S box leader RNA anti-antiterminator structure. Residues in the terminal loop of helix 2 (region 1) are shown in green, residues in the internal loop between helices 3 and 4 (region 2) are shown in purple. B) Potential pairing between regions 1 and 2. The leader RNA sequence present in regions 1 (green) and 2 (purple) are shown for the yitJ and ykrW leader variants. J1 and J2 indicate sequence elements from yitJ regions 1 or 2, respectively, replacing the corresponding positions of ykrW; W1 and W2 indicate sequence elements from ykrW regions 1 or 2 replacing the corresponding positions of yitJ. Potential Watson-Crick pairing is denoted with straight lines, G•U pairing is denoted by filled circles.
addition of SAM promoted transcription termination (McDaniel et al., 2003; Fig. 3.9A, lanes 1, 2; Fig. 3.9B, lanes 1, 2). Use of the glyQS promoter allows more efficient transcription in vitro without affecting the response to SAM (McDaniel et al., 2003; Chapter 2). Addition of SAM to the in vitro transcription assay for the yitJ-W1 variant failed to stimulate termination (Fig. 3.9A, lanes 3, 4). The sequence changes in region 1 of yitJ therefore prevent SAM-dependent transcription termination, despite the fact that the ykrW leader RNA has this sequence in the helix 2 terminal loop (McDaniel et al., 2003). In contrast, in vitro transcription results for the yitJ-W2 template (Fig. 3.9A, lanes 5, 6) were similar to those for the wild-type template, indicating that the sequence changes in region 2 had no effect on SAM-dependent transcription termination in vitro under these conditions. The yitJ-W2 template has an additional G•U base pair as compared to the yitJ-W1 variant (Fig. 3.8B), which could be responsible for the response of the yitJ-W2 construct to the addition of SAM. Introduction of the sequence changes in region 2 in the yitJ-W1W2 template, to restore potential base-pairing, also restored SAM-dependent termination (Fig. 3.9A, lanes 7, 8). Similar results were obtained using variants lacking the A113U mutation (data not shown), indicating that the sequence change at this position has no effect.
Figure 3.9. *In vitro* transcription of *B. subtilis* yitJ and ykrW leader variants. Read-through (RT) and terminated (T) transcripts are labeled. Percent termination (%T) is shown at the bottom of each lane. Templates were transcribed with *B. subtilis* RNAP. SAM was added at 1.5 µM (+). A) Lanes 1 and 2, wild-type yitJ template; lanes 3 and 4, yitJ-W1 template; lanes 5 and 6, yitJ-W2 template; lanes 7 and 8, yitJ-W1W2 template. B) Lanes 1 and 2, wild-type ykrW template; lanes 3 and 4, ykrW-J1 template; lanes 5 and 6, ykrW-J2 template; lanes 7 and 8, ykrW-J1J2 template.
In related experiments performed by Frank Grundy, the effects of the sequence changes in regions 1 and 2 on the response to SAM were further investigated using titration of the SAM concentration required for SAM-dependent transcription termination of the \textit{yitJ} variants. The wild-type template exhibited half-maximal termination efficiency at 0.15 µM SAM, while the \textit{yitJ}-W2 and \textit{yitJ}-W1W2 templates exhibited half-maximal response at 2.2 and 1.5 µM SAM, respectively. The higher SAM concentrations required for half-maximal response of the \textit{yitJ}-W2 and \textit{yitJ}-W1W2 templates suggest that the sequence changes in these variants result in a reduced ability to respond to SAM \textit{in vitro} as compared to the wild-type construct. The \textit{yitJ}-W1 template required 1000-fold more SAM for the half-maximal response than the wild-type template (150 µM), indicating that the ability of the \textit{yitJ}-W1 variant to respond to SAM is reduced 1000-fold as compared to wild-type \textit{yitJ}. The requirement for higher SAM for efficient transcription termination of the \textit{yitJ} leader variants \textit{in vitro} suggests that sequences in regions 1 and 2 could affect affinity for SAM and the stability of the SAM-RNA interaction. The suppression of the severe defect in transcription termination conferred by the region 1 mutations in \textit{yitJ}-W1 by introduction of compensatory changes in region 2 in the \textit{yitJ}-W1W2 construct supports the prediction that pairing between regions 1 and 2 is important for function.

Analysis of the \textit{ykrW} constructs resulted in a pattern analogous to that observed for the \textit{yitJ} variants. The \textit{ykrW}-J2 variant (which has the same base-pairing disruption as \textit{yitJ}-W1; Fig. 3.8B) exhibited a major defect in SAM-dependent termination while the \textit{ykrW}-J1 variant (which has the base-pairing pattern of \textit{yitJ}-W2; Fig. 3.8B) was similar to wild-type \textit{ykrW} (Fig. 3.9B, lanes 3-6). This is consistent with the prediction that the
ykrW-J2 allele is more disruptive of base-pairing between regions 1 and 2 (Fig. 3.8B). The ykrW-J1J2 double mutant (which has sequences in regions 1 and 2 equivalent to those found in the wild-type yitJ leader) exhibited efficient SAM-dependent termination (Fig. 3.9B, lanes 7, 8), demonstrating that alterations in region 1 can suppress the effect of alterations in region 2, consistent with the results for the yitJ variants.

3.3.4.2 Binding of 14C-SAM to yitJ leader covariation mutant RNAs

The effect of mutations in regions 1 and 2 on SAM binding was determined by in vitro binding assays. The yitJ-W1 variant exhibited complete loss of binding of 14C-SAM (Fig. 3.10A), in agreement with the defect in SAM-dependent termination in vitro. The yitJ-W2 and yitJ-W1W2 RNAs bound SAM with an efficiency similar to that of wild-type yitJ RNA, consistent with the observation that the yitJ-W2 mutation did not significantly affect SAM-dependent transcription termination in vitro (although higher SAM concentrations were required) and the yitJ-W2 mutation suppresses the yitJ-W1 defect (Fig. 3.10A). Variants lacking the A113U mutation exhibited similar results, indicating that this mutation does not affect SAM binding (data not shown). Wild-type ykrW, ykrW-J1 and ykrW-J1J2 RNAs were all capable of binding 14C-SAM, although binding in ykrW-J1 and ykrW-J1J2 was reduced two-fold and seven-fold, respectively, compared to wild-type, indicating that these variants do have some effect on SAM binding in vitro. In contrast, ykrW-J2 RNA was completely defective in SAM binding (Fig. 3.10B), again consistent with the in vitro transcription results and with the effects of the analogous sequence rearrangement for regions 1 and 2 in the context of the yitJ leader. Binding activity was specific to SAM since in each case addition of a fifty-fold
molar excess of nonradioactive SAM blocked binding of $^{14}$C-SAM, while nonradioactive SAH failed to compete (Fig. 3.10B). These results demonstrate that both SAM binding and SAM-dependent transcription termination \textit{in vitro} respond in a similar way to sequence changes in regions 1 and 2.
Figure 3.10. Binding of $^{14}$C-SAM by yitJ and ykrW leader RNAs. $^{14}$C-SAM (8 μM) and T7 RNAP-transcribed leader RNA (8 μM) were incubated together in 1X transcription buffer containing MgCl₂ (10 mM) in the presence or absence of nonradioactive SAM (filled bars) or SAH (hatched bars) at 400 μM. SAM binding is expressed as the percentage of $^{14}$C-SAM retained after filtration through a Nanosep 10K filter and 5 washes with 150 μl 1X transcription buffer, relative to the amount of $^{14}$C-SAM added to the binding reaction. SAM retention in the absence of leader RNA was <0.1%. A) yitJ leader RNA variants. B) ykrW leader RNA variants.
3.3.4.3 RNase A mapping of \textit{yitJ} and \textit{ykrW} leader covariation mutant RNAs

Nuclease digestion was performed for structural analysis of \textit{yitJ} and \textit{ykrW} leader variant RNAs. RNase T1 digestion of RNA transcripts generated by transcription with \textit{B. subtilis} RNAP (performed by Frank Grundy; McDaniel \textit{et al.}, in press) revealed changes in cleavage at certain positions in response to SAM, indicating structural changes induced by SAM binding. However, this analysis was complicated by the presence of RNAs of different lengths due to pausing and termination. We therefore carried out structural mapping of 5’ end-labeled RNAs generated by T7 RNAP, using the conditions for RNA refolding employed in the SAM binding assays. RNase A, which cleaves after U and C residues, was used to provide information about cleavage susceptibility of residues different from those probed by RNase T1, which is G-specific. There were few strong cleavage sites in the leader RNAs, consistent with the complex structural arrangement predicted by the S box leader models. RNase A digestion products for \textit{yitJ}-wild-type RNA revealed efficient cleavage at U33 and C36 (in helix 1) in the absence of SAM (Fig. 3.11A, lanes 3-7) but decreased cleavage at these residues in the presence of SAM (Fig. 3.11A, lanes 8-12). Efficient cleavage was observed at U19 and C21 in the presence and the absence of SAM (Fig.3.11A, lanes 3-12) as expected, since these are predicted to be unpaired residues located 5’ of helix 1 in the \textit{yitJ} leader RNA (Grundy and Henkin, 1998). Efficient cleavage of wild-type \textit{yitJ} RNA was also observed in the presence and absence of SAM at residues C92 (in helix 3) and at positions C125 and C129 (in the helix 4 terminal loop; Fig. 3.11A, lanes 3-12; Fig. 3.11B, lanes 2-11). The \textit{yitJ} leader variants also exhibited efficient cleavage at these positions, demonstrating that the overall structure of the RNA was not grossly affected. An increase in cleavage at U82 (in helix
3) in the presence of SAM was observed for \textit{yitJ} wild-type RNA as well as the \textit{yitJ}-W2 and \textit{yitJ}-W1W2 variants, indicative of a SAM-induced structural change. In contrast, no change in cleavage at U82 was observed for the \textit{yitJ}-W1 RNA (Fig. 3.11B, lanes 12-21). Residues C53 and U54 in region 1 were not cleaved by RNase A in the wild-type \textit{yitJ} leader RNA in the presence or absence of SAM (Fig. 3.11A, lanes 3-12; Fig. 3.11B, lanes 2-11). However, the \textit{yitJ}-W2 RNA, which has the same sequence in region 1 as wild-type \textit{yitJ} RNA, exhibited efficient cleavage at C53 and U54 in the absence of SAM, and reduced cleavage in the presence of SAM (Fig. 3.11B, lanes 22-31). This pattern is consistent with the results of the RNase T1 digestion of RNA transcripts generated by transcription with \textit{B. subtilis} RNAP (performed by Frank Grundy; McDaniel \textit{et al.}, in press), where introduction of sequence changes in region 2 affected the accessibility of residues in region 1. The predicted modest destabilizing effect of the \textit{yitJ}-W2 allele appears to allow increased cleavage in region 1 in the absence of SAM without disturbing the response to SAM.
Figure 3.11. RNase A mapping of \textit{yitJ} leader RNAs. 5’ end-labeled leader RNA (0.2 µM) was incubated in 1X transcription buffer containing MgCl$_2$ (10 mM) in the presence (+) or absence (-) of 1.5 µM SAM, followed by digestion with RNase A for the indicated times. T1, RNase T1 digestion of denatured wild-type leader RNA; certain G residues are labeled as size standards. A. \textit{yitJ} wild-type leader RNA. Products were resolved by denaturing 10% polyacrylamide gel electrophoresis. B. \textit{yitJ} leader RNA variants. Products were resolved by denaturing 6% polyacrylamide gel electrophoresis.
Figure 3.11
Figure 3.12. RNase A mapping of \textit{ykrW} leader RNAs. 5’ end-labeled leader RNA (0.2 µM) was incubated in 1X transcription buffer containing MgCl$_2$ (10 mM) in the presence (+) or absence (-) of 1.5 µM SAM, followed by digestion with RNase A for the indicated times. T1, RNase T1 digestion of denatured wild-type leader RNA; certain G residues are labeled as size standards. A. \textit{ykrW} wild-type leader RNA. Products were resolved by denaturing 10% polyacrylamide gel electrophoresis. B. \textit{ykrW} leader RNA variants. Products were resolved by denaturing 6% polyacrylamide gel electrophoresis.
Figure 3.12
The \emph{yitJ}-W1 RNA exhibited cleavage at C53 regardless of the presence of SAM (Fig. 3.11B, lanes 12-21), consistent with the observed loss of SAM binding and SAM-directed transcription termination \textit{in vitro} and the RNase T1 cleavage results for RNA transcripts generated by transcription with \textit{B. subtilis} RNAP (performed by Frank Grundy; McDaniel \textit{et al.}, in press). The \emph{yitJ}-W1W2 variant, which has the same sequence in region 1 as \emph{yitJ}-W1 but also has restored base-pairing to region 2, exhibited protection from digestion in the presence of SAM (Fig. 3.11B, lanes 37-41). These results again demonstrate that sequence changes in region 2 can suppress the negative phenotype of sequence changes in region 1, further supporting the hypothesis that interaction between regions 1 and 2 is required for SAM-dependent modulation of the leader RNA structure. Region 2 residues C116 and C119 were not cleaved in the \emph{yitJ} RNA in the presence or absence of SAM, and showed weak protection in the presence of SAM for \emph{yitJ}-W2 and \emph{yitJ}-W1W2. In contrast, cleavage of these residues was observed in \emph{yitJ}-W1 RNA in the presence and absence of SAM, supporting the proposal that changes in region 1 affect the structure of region 2. Cleavage at C62 (in the kink-turn element in helix 2) in each \emph{yitJ} leader variant exhibited the pattern observed for region 1 residues C53 and U54 (i.e., SAM-dependent protection in \emph{yitJ}-W2 and \emph{yitJ}-W1W2; Fig. 3.11B). This result is consistent with the model that the kink-turn element participates in a SAM-induced tertiary interaction between regions 1 and 2.

RNase A digestion of \emph{ykrW} RNAs was also carried out to determine if the inverse sequence alterations resulted in a pattern similar to that observed in the context of the \emph{yitJ} sequence. The general cleavage pattern was similar for all \emph{ykrW} variants, demonstrating
that the mutations do not cause a major disruption in RNA structure (Fig. 3.12). Efficient SAM-independent cleavage was observed at positions U75 and U77 (at the top of helix 3), and at U103 (in helix 4) regardless of the presence of SAM. The equivalent regions in \textit{yitJ} also exhibited cleavage in the presence or absence of SAM, providing support for the model that the different S box leader RNAs fold into similar structural arrangements.

Residues U26 and C29 in helix 1 in \textit{ykrW} wild-type RNA showed decreased cleavage in the presence of SAM vs. the absence of SAM (Fig. 3.12A, lanes 3-12), as did the residues in helix 1 in \textit{yitJ} wild-type RNA. Positions C99 in \textit{ykrW} and C119 in \textit{yitJ}-W1W2, which occupy analogous positions at the base of helix 4, exhibited weak protection from cleavage in the presence of SAM, providing additional evidence for structural conservation.

Cleavage at C47 in region 1 was observed in the \textit{ykrW} wild-type RNA in the absence of SAM (Fig. 3.12A, lanes 3-7; Fig. 3.12B, lanes 2-6), while addition of SAM resulted in protection from digestion (Fig. 3.12A, lanes 8-12; Fig. 3.12B, lanes 7-11). This is similar to the cleavage pattern for residue C53 in \textit{yitJ}-W1W2, which has the same sequence in regions 1 and 2 as wild-type \textit{ykrW} RNA. The \textit{ykrW}-J1 variant exhibited cleavage at C47 and U48 in the absence of SAM (Fig. 3.12B, lanes 12-16), and protection from digestion in the presence of SAM (Fig. 3.12B, lanes 17-21). This pattern is analogous to that observed for C53 in \textit{yitJ}-W2, which has the same sequence in regions 1 and 2 as \textit{ykrW}-J1. Similarly, the \textit{ykrW}-J2 variant, which has the same sequence in regions 1 and 2 as \textit{yitJ}-W1, exhibited cleavage at C47 in the presence and the absence of SAM (Fig. 3.12B, lanes 22-31), consistent with the more severe disruption of pairing predicted for this allele (Fig. 3.8B). Finally, the \textit{ykrW}-J1J2 variant, which has the same
sequence in regions 1 and 2 as \textit{yitJ} wild-type RNA, did not exhibit cleavage in region 1 by RNase A in the presence or absence of SAM (Fig. 3.12B, lanes 32-41), a pattern similar to that observed for wild-type \textit{yitJ}. The loss of cleavage at C47 in the \textit{ykrW}-J1J2 RNA demonstrates that sequence alterations in region 1 can compensate for alterations in region 2, analogous to the suppression of the phenotype of the \textit{yitJ}-W1 allele in \textit{yitJ}- W1W2. Cleavage at C39 (in the kink-turn motif in helix 2) in the \textit{ykrW} RNAs exhibited a pattern similar to that observed for region 1, supporting the prediction that SAM-dependent protection of region 1 correlates with an alteration in the reactivity of the kink-turn motif.

Cleavage within region 2 was inefficient in the \textit{ykrW} variant RNAs, precluding a detailed examination. Position C95 in region 2 of \textit{ykrW}-J2 RNA was cleaved in the presence and absence of SAM (Fig. 3.12B, lanes 22-31), consistent with the loss of SAM-dependent stabilization observed at region 1. This result is similar to the SAM-independent cleavage at residue C116 in \textit{yitJ}-W1 RNA. The RNase A digestion patterns observed in regions 1 and 2 for \textit{ykrW} leader RNAs therefore matched the patterns for the \textit{yitJ} leader RNAs with the equivalent sequence in these regions. This demonstrates that the region 1 and 2 sequence elements can be transplanted into different sequence contexts if both elements are introduced together.

3.3.4.4 \textit{In vivo} expression of \textit{yitJ-lacZ} fusions for \textit{yitJ} covariation mutants

The effect of mutations in regions 1 and 2 on \textit{yitJ-lacZ} expression \textit{in vivo} was tested to determine if the \textit{in vitro} effects on SAM binding and SAM-dependent transcription termination correlate with \textit{in vivo} expression. Expression of a wild-type
yitJ-lacZ transcriptional fusion is induced in response to starvation for methionine and is repressed greater than 1000-fold during growth in the presence of methionine (Grundy and Henkin, 1998; Chapter 1; Table 3.1; Table 3.2). Intracellular SAM pools decrease in cells starved for methionine (Wabiko et al., 1988) and an increase in SAM synthetase activity results in a delayed response to starvation for methionine (McDaniel et al., 2003; Chapter 4), consistent with the model that SAM is the molecular effector in vivo. The yitJ-W1 mutant exhibited a small decrease in expression during starvation for methionine relative to the wild-type fusion, while repression during growth in the presence of methionine was abolished (Table 3.2). This is consistent with the loss of SAM-directed transcription termination and SAM binding in vitro, and absence of alterations in the RNase cleavage pattern in the presence of SAM. The yitJ-W2 fusion was expressed at the same level as the wild-type during growth in the absence of methionine and expression was repressed only three-fold during growth in the presence of methionine. The yitJ-W2 allele had little effect on SAM-dependent transcription termination in vitro when SAM concentrations were high, but resulted in a 10-fold decrease in sensitivity to SAM in experiments performed by Frank Grundy. The lack of efficient repression in vivo under conditions when intracellular pools of SAM are expected to be relatively high may reflect the decrease in sensitivity to SAM observed in the in vitro transcription assays. The yitJ-W1W2 fusion was expressed at the same level as the wild-type fusion during growth in the absence of methionine, and exhibited five-fold repression during growth in the presence of methionine. The lack of full repression in vivo may again reflect the observed requirement for high SAM concentrations for efficient transcription termination in vitro. Although repression was not restored to the wild-type level, it
nevertheless appears that the \textit{yitJ-W2} mutation conferred significant suppression of the \textit{yitJ-W1} allele. Expression of a \textit{ykrW-lacZ} transcriptional fusion, which has the same sequence in regions 1 and 2 as \textit{yitJ-W1W2}, is not fully repressed during growth in the presence of methionine (Murphy \textit{et al.}, 2002; Chapter 5), consistent with lack of full repression of the \textit{yitJ-W1W2} fusion.

<table>
<thead>
<tr>
<th>Fusion</th>
<th>\textit{(\beta)-galactosidase activity}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>- Met</td>
</tr>
<tr>
<td>wild-type</td>
<td>240</td>
</tr>
<tr>
<td>\textit{yitJ-W1}</td>
<td>180</td>
</tr>
<tr>
<td>\textit{yitJ-W2}</td>
<td>280</td>
</tr>
<tr>
<td>\textit{yitJ-W1W2}</td>
<td>240</td>
</tr>
</tbody>
</table>

\textbf{Table 3.2. Expression of \textit{yitJ-lacZ} transcriptional fusions in response to methionine availability.} Fusions were integrated in single copy in strain BR151 (\textit{lys-3 metB10 trpC2}). Cells were grown in Spizizen minimal medium (Anagnostopoulos and Spizizen, 1961) containing methionine, harvested by centrifugation, and resuspended in Spizizen minimal medium in the presence (+ Met) or absence (- Met) of methionine. Samples for the \textit{\(\beta\)-galactosidase assay} were taken 4 h after the cultures were split. \textit{\(\beta\)-galactosidase activity} is expressed in Miller units (Miller, 1972).
3.4 Discussion

It was demonstrated through phylogenetic analysis that S box leaders exhibit high conservation of primary sequence and secondary structural elements (Grundy and Henkin, 1998; Chapter 1). Early experiments confirmed that conserved elements in the anti-antiterminator were likely to play a role in effector recognition and binding (Grundy and Henkin, 1998, 2003). Having identified SAM as the molecular effector for S box gene regulation (McDaniel et al., 2003; Chapter 2), we decided to carry out experiments to further analyze the importance of S box leader cis-acting sequence and structural elements in SAM-directed transcription termination.

We made deletion mutants of ykrW to examine which elements in the helix 1-4 region are required for SAM binding. The results indicate that the entire helix 1-4 region is essential for SAM binding by the leader RNA. We also examined the ability of the yoaD leader to bind SAM. Of all the S box leaders identified to date, yoaD is the only leader that has a tetraloop at the top of helix 3 in place of the larger helix at the top of helix 3 in all other S box leaders, and an additional stem loop preceding the terminator. SAM binding assays performed for yoaD leader RNAs with or without the extra stem loop structure indicate that the extra stem loop is vital for SAM binding by yoaD leader RNA, despite the fact that this structure is absent in all other S box leaders. It is possible that in the case of yoaD, the extra stem loop somehow compensates for the shortened helix 3. The fact that the ykrW leader variant with a shortened helix 3, designed to mimic helix 3 in yoaD, was incapable of SAM binding provides further evidence that the shortened helix 3 in yoaD is not sufficient for SAM binding in the absence of the additional stem loop. The observation that the yoaD leader RNA including the stem loop
exhibited approximately five-fold greater SAM binding than wild-type \( ykrW \) suggests that the stability of the interaction between SAM and individual S box leader RNAs varies. It is possible, however, that the \( ykrW \) and \( yoaD \) leader variants that failed to bind SAM are not capable of folding correctly under the \textit{in vitro} conditions generated.

Genetic analysis showed that mutations that disrupt conserved primary sequence and structural elements in the helix 1-4 region of the S box leader result in loss of repression during growth in the presence of methionine (Grundy and Henkin, 1998). \textit{In vitro} transcription assays performed for two of these mutants, \( \textit{yitJ-Pst-1} \) and \( \textit{yitJ-Pst-2} \), indicated that these mutations also result in loss of SAM-directed transcription termination. The observation that the \( \textit{yitJ-Pst-1} \) and \( \textit{yitJ-Pst-2} \) mutants exhibit loss of repression during growth in the presence of methionine and loss of SAM-dependent termination \textit{in vitro} indicates that at least some of the conserved leader features that are required for regulation \textit{in vivo} are also important for SAM-directed transcription termination \textit{in vitro}. The loss of SAM binding exhibited by the \( \textit{yitJ-Pst-2} \) mutant suggests that at least some of the determinants in helix 2 required for regulation \textit{in vivo} are also important for binding of SAM by the leader RNA \textit{in vitro}. Winkler \textit{et al}. (2003) also demonstrated that mutations that disrupt the helix 1-4 region of \( B.\ subtilis\ yitJ \) result in loss of SAM binding \textit{in vitro} and loss of repression of during growth in the presence of methionine, while compensatory mutations restore at least partial binding activity and repression during growth in the presence of methionine.

We also used site-directed mutagenesis to determine the importance of the kink-turn motif in helix 2 in S box gene regulation. Single nucleotide substitutions were made for residues in the \( \textit{yitJ} \) leader that are greater than 90 percent conserved in the kink-turn
motifs of all S box leaders. These mutations resulted in loss of repression during growth in the presence of methionine, and the level of expression for these mutants was lower than that of wild-type yitJ in the absence of methionine. It is not clear why these mutations also affect the level of expression in the absence of methionine. The fact that the double mutant, which restores secondary structure but not primary sequence, did not fully restore repression during growth in the presence of methionine indicates that both the conserved primary sequence and the formation of secondary structure are important for regulation. These results, in combination with phylogenetic analysis, confirm the importance of the kink-turn motif in S box gene regulation. The kink-turn motif was also shown to be important in the T box regulatory system, as disruption of this motif in the B. subtilis tyrS leader, a member of the T box family, resulted in loss of antitermination (Winkler et al., 2001). The results obtained with the C62G mutation showed that although the kink-turn motif is also critical in T box gene regulation, S box and T box kink-turn motifs have different sequence requirements (Winkler et al., 2001).

Nucleotides that covary in RNA secondary structures have the potential to participate in tertiary interactions, including pseudoknots (Michel and Westhof, 1990; ten Dam et al., 1992; Woese and Pace, 1993). However, sequence covariation does not always correlate with tertiary interactions (Chernyaeva and Murgola, 2000). Covariation between residues in the terminal loop of helix 2 (region 1) and residues in the unpaired region between helices 3 and 4 (region 2) was found in all S box leaders, indicating that an interaction between these regions could be important for formation of tertiary structure and SAM binding (Grundy and Henkin, 1998; McDaniel et al., 2003). As noted, mutational analysis supported our observation that conserved elements in an internal loop
within helix 2 fit the pattern for a kink-turn motif that could facilitate interaction of regions 1 and 2 by introducing a bend in the RNA. Lysine-binding RNAs in the L box family also have a kink-turn motif in helix 2 and an element in the helix 2 terminal loop that covaries with another unpaired leader element (Grundy et al., 2003), which suggests that helix 2 could play a similar role in both classes of RNAs.

To test the hypothesis that residues in region 1 interact with residues in region 2 in the tertiary structure of S box leader RNAs, we introduced mutations in these regions that could affect base-pairing, as well as compensatory mutations that would restore base-pairing but not the primary sequence. The mutagenesis strategy utilized sequences found in another S box leader to preserve naturally occurring primary sequence elements.

Consistent with the prediction, the \textit{yitJ}-W1 and \textit{ykrW}-J2 alleles, in which two of the four conserved base-pairs are eliminated (Fig. 3.8B), resulted in loss of SAM-directed transcription termination, SAM binding, and SAM-induced alterations in susceptibility to nuclease digestion \textit{in vitro}; the \textit{yitJ}-W1 mutant also exhibited complete loss of repression during growth in the presence of methionine. The \textit{yitJ}-W2 and \textit{ykrW}-J1 mutants, which retain two Watson-Crick base-pairs and a G\textbullet{}U base-pair, exhibited a pattern of SAM-directed transcription termination, SAM binding, and SAM-induced protection from nuclease digestion \textit{in vitro} similar to the wild-type constructs, although in related experiments performed by Frank Grundy, a 10-fold-higher SAM concentration was required for efficient SAM-dependent transcription termination. These results demonstrate that the sequence in regions 1 and 2 plays an important role in the S box regulatory system, and suggest that at least three base-pairs between regions 1 and 2 are required. The \textit{yitJ}-W2 and \textit{ykrW}-J1 mutations suppressed the phenotype of the \textit{yitJ}-W1
and \textit{ykrW}-J2 alleles, respectively, in all of the \textit{in vitro} assays. Similarly, the \textit{yitJ}-W2 mutation partially suppressed the phenotype of the \textit{yitJ}-W1 allele \textit{in vivo}. The suppression of the phenotype conferred by mutations in one region by compensatory changes in a second region is a key result that strongly supports the hypothesis that residues in region 1 interact with residues in region 2.

Analysis of \textit{yitJ} leader RNA by spontaneous cleavage in the presence of Mg$^{+2}$ (Winkler \textit{et al.}, 2003) resulted in a cleavage pattern consistent with the secondary structure model (Grundy and Henkin, 1998), and with our results using RNase T1 and RNase A. Cleavage at residues in the terminal loops of helices 3 and 4 in the presence or absence of SAM, and SAM-induced protection of specific residues in helices 1-4, including residues within the kink-turn element, were observed in both studies. Slight protection of residues C53 and U54 in region 1 was also reported, but the cleavage at these positions was weak even in the absence of SAM. Similarly, residues C53, U54 and G55 in region 1 exhibited low cleavage by RNase A or RNase T1 regardless of the presence of SAM, indicating that accessibility of these residues is low. In contrast, the \textit{yitJ}-W2 variant, which retains wild-type sequence in region 1, exhibited strong RNase cleavage at these residues in the absence of SAM, and SAM-dependent protection. This result clearly demonstrates that sequence changes in region 2 can significantly affect the structural arrangement of residues in region 1. A similar pattern was observed for residues in region 2, supporting this conclusion.

It is of interest to note that the accessibility of residues in regions 1 and 2 of wild-type \textit{yitJ} leader RNA to RNase A was lower than that of wild-type \textit{ykrW} leader RNA in the absence of SAM. This suggests that the \textit{yitJ} and \textit{ykrW} leader RNAs differ in their
predisposition to fold into the proposed tertiary structural arrangement prior to SAM binding. We also observed that leader RNAs with the same sequence in regions 1 and 2 responded similarly to RNase A digestion, which indicates that transplantation of sequence elements in regions 1 and 2 results in acquisition of the corresponding pattern of accessibility, so that the \( yitJ-W1W2 \) allele exhibited SAM-dependent protection similar to that of wild-type \( ykrW \) RNA, and \( ykrW-J1J2 \) RNA exhibited low cleavage regardless of the presence of SAM, similar to that exhibited by wild-type \( yitJ \) RNA. The propensity of certain S box RNAs to prefold into a tertiary arrangement similar to that induced by SAM binding in other RNAs could reflect different affinities of these RNAs for SAM and/or differential sensitivity to SAM. As noted, \( ykrW-lacZ \) fusion constructs are not as tightly repressed as \( yitJ-lacZ \) constructs in vivo under conditions where SAM pools are elevated, consistent with the model that prearrangement of the tertiary interaction (as proposed for \( yitJ \)) may facilitate the response to SAM. Since sequence alterations in regions 1 and 2 affect this tertiary arrangement, this interaction could have a major effect on the overall structure of S box leader RNAs, potentially affecting their ability to respond to SAM. The residues in regions 1 and 2 of the S box leaders exhibit covariation and appear to interact via canonical Watson-Crick pairing. The fact that these residues are protected from RNase cleavage in the absence of SAM in certain S box RNAs, like \( yitJ \), but not in others, like \( ykrW \), suggests that while the ability of these residues to interact is required for SAM binding, a preexisting interaction is not necessary; this conclusion is supported by the observation that binding of SAM was retained in variants of \( yitJ \) in which the tertiary interaction appears to be SAM-dependent.
The *yitJ*-W1 variant exhibited no repression during growth in the presence of methionine. The *yitJ*-W1W2 mutant, which restores base-pairing between regions 1 and 2 but switches the primary sequence to that of *ykrW*, exhibited partial repression during growth in the presence of methionine, but suppression of the mutations in region 1 was incomplete compared to that observed in the *in vitro* assays. Similarly, the *yitJ*-W2 variant exhibited a significant defect in repression *in vivo* under conditions where SAM pools are predicted to be high, but exhibited little defect in other assays. Since both the *yitJ*-W2 and *yitJ*-W1W2 variants required a 10-fold higher SAM concentration in the *in vitro* termination assay as compared to the wild-type *yitJ* template, it is likely that the sequence changes in regions 1 and 2 result in reduced affinity for SAM. The SAM concentration used in the *in vitro* assays is within the physiological range reported by Wabiko *et al.* (1988), but the cellular pool of free SAM that is available for binding to the S box leader RNA transcripts is likely to be significantly lower. It therefore appears that expression *in vivo* may be more sensitive to alterations in regions 1 and 2 than is apparent from the *in vitro* assays. The possibility also remains that although the *in vitro* conditions generated were sufficient for SAM binding to the leader RNAs and SAM-directed transcription termination, additional factors could be involved in S box gene regulation *in vivo*.

The crystal structures of the guanine- and adenine-sensing RNAs revealed that the RNA forms a purine-binding pocket that is anchored by interaction of phylogenetically conserved nucleotides in the terminal loops (Batey *et al.*, 2003; Serganov *et al.*, 2004). This interaction occurs in the presence or absence of the effector, and is required for and stabilized by effector binding (Batey *et al.*, 2004). Covarying regions have also been
identified in other effector-binding regulatory RNAs, suggesting that they too form tertiary interactions. However, these proposed interactions have not yet been verified experimentally. The current study provides the first information concerning the three-dimensional arrangement of structural elements within the S box RNAs, and provides clear evidence that the interaction between regions 1 and 2 in the S box leader RNAs is required for SAM binding and SAM-directed transcription termination.
CHAPTER 4

CHARACTERIZATION OF THE *Bacillus subtilis* metK GENE AND ITS EFFECT ON S BOX GENE EXPRESSION

4.1 Introduction

Genetic analyses of S box leader RNAs suggested that the helix 1-4 region was likely to be the target for binding of a negative regulatory factor or factors (Grundy and Henkin, 1998; Winkler *et al*., 2001; Chapter 3), which we identified as SAM (McDaniel *et al*., 2003; Chapter 2). The *B. subtilis* yitJ Pst-1 and Pst-2 mutations, which introduce changes in conserved elements in helix 1 and helix 2, respectively, both result in loss of repression of *yitJ* expression during growth in the presence of methionine (Grundy and Henkin, 1998) and loss of SAM-directed transcription termination *in vitro* (McDaniel *et al*., 2003; Chapter 3). Mutations in the conserved GA motif in helix 2 of *yitJ* also result in loss of repression during growth in the presence of methionine (Winkler *et al*., 2001; Chapter 3). The Pst-2 mutation in *yitJ* as well as mutations that delete helices 1, 3, or 4 in the *B. subtilis* ykrW leader result in loss of SAM binding *in vitro* (McDaniel *et al*., 2003;
Chapter 3). These findings provided additional evidence that the helix 1-4 region is essential for binding of SAM. While SAM was shown to be sufficient for RNA binding and transcription termination \textit{in vitro}, it was not clear whether other factors might be involved in S box gene regulation \textit{in vivo}.

In experiments described in this chapter, we attempted to identify \textit{trans}-acting mutations that would lead to loss of repression of S box gene expression during growth in the presence of methionine. We predicted that we could isolate \textit{trans}-acting mutations in genes that affect the molecular effector (SAM) or other genes involved in the folding of the leader RNA, the SAM/leader RNA interaction, or transcription termination in response to SAM. These experiments resulted in the isolation and characterization of one mutant that specifically affected S box gene expression. The genetic alteration in this isolate was identified as a single nucleotide substitution in the coding region of \textit{metK} (formerly designated \textit{metE}; Yocum \textit{et al}., 1996), the gene encoding SAM synthetase, which results in decreased SAM synthetase activity.

Since SAM was shown to be the molecular effector for S box termination \textit{in vitro}, we also analyzed the effect of increased SAM synthetase activity on S box gene expression \textit{in vivo}. Posnick and Samson (1999) established that increased SAM synthetase activity in \textit{E. coli} results in increased intracellular SAM pools. In experiments described in this chapter, we attempted to alter SAM pools in \textit{B. subtilis} by overexpression of SAM synthetase. We find that increasing SAM synthetase activity resulted in a delay in S box expression during methionine limitation, a phenotype opposite to that observed when SAM synthetase activity was decreased. These studies
provided additional support for the hypothesis that SAM directly affects S box gene transcription termination in vivo (McDaniel et al., 2003).

4.2 Materials and Methods

4.2.1 Bacterial strains and growth conditions

The B. subtilis strains used in this study were 168 (trpC2), BR151 (lys-3 metB10 trpC2), SA29 (trpC2 metK1 sacB::φPvegmetK+ [Yocum et al., 1996], obtained from J. Pero [OminiGene Bioproducts, Cambridge, MA]; metK was formerly designated metE; Yocum et al., 1996), and ZB307A (SPβc2del2::Tn917::pSK10Δ6), and IS56B (lys-3 metB relA1 sup-3 strA), a derivative of strain IS56 (lys-3 relA1 trpC2 [Smith et al., 1980] containing a mutation that partially suppresses the partial growth defect conferred by the relA1 allele (unpublished results). B. subtilis mapping strains 1A627 through 1A645 (Vandeyar and Zahler, 1986) were obtained from the Bacillus Genetic Stock Center (The Ohio State University) and generalized transducing phage PBS-1 was obtained from P. Setlow (University of Connecticut Health Center, Farmington, CT). B. subtilis strains were grown on tryptose blood agar base medium (TBAB; Difco), Spizizen minimal medium (Anagnostopoulos and Spizizen, 1961), 2XYT broth (Miller, 1972), or A3 medium (Antibiotic medium #3; Difco). Antibiotics were added as indicated at the following concentrations: chloramphenicol, 5 µg/ml; erythromycin, 1 µg/ml; lincomycin, 25 µg/ml; neomycin, 5 µg/ml; spectinomycin, 25 µg/ml. All growth was at 37°C and was measured with a Klett-Summerson colorimeter using the No. 66 red filter.
4.2.2 Genetic techniques

Transformation of *B. subtilis* was carried out as described by Henkin and Chambliss (1984). Chromosomal DNA was prepared using the DNeasy Tissue Kit (Qiagen). Wizard columns (Promega) were used for plasmid preparations. Oligonucleotide primers (Table 4.1) were purchased from Integrated DNA Technologies (Coralville, IA). Restriction endonucleases and DNA modifying enzymes were purchased from New England Biolabs and used as described by the manufacturer. Plasmid pGEM7Zf+ (Promega) was used for cloning. Mutations were identified by DNA sequencing (Genewiz Inc., North Brunswick, NJ).
<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>METKDS2</td>
<td>CCTCCTGCACAAGGCCTCCCCGAAAGACC</td>
</tr>
<tr>
<td>METKDS3</td>
<td>GCTCATCAATTTGATGTTAAAAACCGC</td>
</tr>
<tr>
<td>METKDS4</td>
<td>TAATTTCAGGGTGATGCTGAGTTGA</td>
</tr>
<tr>
<td>METKDS5</td>
<td>TGTTGATTGAGATTGACACAGGCTG</td>
</tr>
<tr>
<td>METKDS6</td>
<td>TGAACCCTATTTGTCGACGCTGTG</td>
</tr>
<tr>
<td>METKDS7</td>
<td>CGTTTCGTTGCACGCAGATCCGAACATTAATCC</td>
</tr>
<tr>
<td>METKUS2</td>
<td>TTGTTTTCCCTCTAGATAGGTGCT</td>
</tr>
<tr>
<td>METKUS3</td>
<td>GAGGGAGCGAGTTCCGTATCATATA</td>
</tr>
<tr>
<td>METKUS4</td>
<td>CGTGCAAATACGGATTTGATCGG</td>
</tr>
<tr>
<td>METKUS5</td>
<td>CCAGTCCGCATTGACGCGATTGTA</td>
</tr>
<tr>
<td>METKUS6</td>
<td>CAGCTTGCTTACGCGATCGGTGTG</td>
</tr>
<tr>
<td>METKUS7</td>
<td>GACAGCATTCTAGATGAAATTTAAAGAACCTC</td>
</tr>
</tbody>
</table>

Table 4.1. Oligonucleotide primers. Sequences are shown from 5’ to 3’. Underlined positions show variation from the wild-type sequence. Primers METKDS7 and METUS7 were used for cloning a 350 bp fragment of the *metK* gene. The remaining primers were used for sequencing of the *metK* gene.

Transcriptional fusions were generated in plasmid pFG328 (Grundy *et al.*, 1993), which contains the *cat* gene and an erythromycin resistance cassette, and integrated in single copy into the chromosome using specialized transducing phage SPβ, as previously
described (Grundy and Henkin, 1993; Grundy and Henkin, 1998). Strains containing 
lacZ fusions were grown in the presence of chloramphenicol, or spectinomycin for fusions in which the erythromycin resistance cassette is replaced with a spectinomycin resistance cassette. Plasmid pFGneo, a derivative of plasmid pFG328 that contains a neomycin resistance gene cassette in place of the cat gene, was introduced into strains by transformation to replace transcriptional fusions with a neomycin resistance cassette by homologous recombination.

Methanesulfonic acid (EMS; Sigma) was used for mutagenesis. Cells were grown in 2XYT to 30 Klett units before addition of EMS (1% final concentration). Cells were incubated with shaking for an additional 100 min at 37°C, pelleted by centrifugation, and washed twice with Davis salts (0.7% K₂HPO₄, 0.3% KH₂PO₄, 0.1% (NH₄)₂SO₄ and 0.005% MgSO₄·7H₂O). Cells were then resuspended in 2XYT and incubated with shaking for 3 h at 37°C before storage at -70°C in 10% glycerol. EMS mutagenized cells were later diluted in 2XYT, grown for 30 min and plated on TBAB containing X-gal as an indicator of lacZ expression.

Preparation of PBS-1 transducing lysates and transductions were carried out as described by Jamet and Anagnostopoulos (1969) with minor modifications (Henkin and Chambliss, 1984). PBS-1 transducing lysates were made from nineteen B. subtilis mapping strains containing chromosomal insertions of the transposon Tn917 (Vandeyar and Zahler, 1986). The PBS-1 phage were passed through the donor strains at least three times before transduction of the recipient strain. Erythromycin and lincomycin were used to select for resistance to macrolide, lincomycin, and streptogramin B antibiotics (MLS resistance) encoded by the transposon.
The tblastn program of the BLAST suite of sequence analysis programs (Altschul et al., 1990) was used to identify protein homologs in unannotated sequence data. Signature amino acid patterns were retrieved from the PROSITE database of protein families and domains (Gasteiger et al., 2003).

4.2.3 β-galactosidase measurements

Cells containing lacZ fusions were grown in Spizizen minimal medium containing all required amino acids at 50 µg/ml until early exponential growth. The cells were harvested by centrifugation, then resuspended in fresh Spizizen minimal medium in the presence or absence of methionine. Samples were collected at 1 h intervals and assayed for β-galactosidase activity (Miller, 1972; Chapter 3), and variation was <10%.

4.2.4 SAM synthetase assay

Cells (400 ml) were grown in Spizizen minimal medium containing the required amino acids until the OD595 reached ~0.8. Crude cell extracts were prepared by passage through a French pressure cell (8,000 lb/in²) followed by removal of cell debris by centrifugation (22,000 x g for 20 min at 4°C). Protein concentration of cell extracts was measured by the Bio-Rad (Bradford) protein assay using a bovine serum albumin standard (New England Biolabs). Cell extracts were assayed for SAM synthetase activity using a protocol adapted from Ochi and Freese (1982) with minor modifications, by measuring incorporation of [35S]methionine into [35S]SAM. Samples were filtered through P81 phosphocellulose paper (Upstate Biotechnology) and washed with distilled water. Bound [35S]SAM was quantitated in Packard Bioscience Ultima Gold scintillation
fluid using a Packard Tri-Carb 2100TR liquid scintillation counter. The recovery of SAM after the wash with distilled water was 67% as determined with [methyl-\(^{14}\text{C}\)]-SAM (ICN); therefore the measured amounts of \(^{35}\text{S}\)SAM were multiplied by 1.5. The amount of \(^{35}\text{S}\)SAM generated in the reaction mixture was measured at 10 min intervals and used to determine the specific activity of SAM synthetase in the crude cell extract by calculating the amount of \(^{35}\text{S}\)SAM (in pmol) generated per mg of protein in the cell extract per minute.

4.3 Results

4.3.1 Isolation of S box-depressed mutants of *B. subtilis* strain IS56B

EMS mutagenesis was employed to generate random mutations in *B. subtilis* strain IS56B containing a *yitJ-lacZ* fusion, in an attempt to isolate *trans*-acting mutations that would affect S box gene expression. The strategy used to isolate and characterize these mutants is illustrated in Fig. 4.1. Four independent pools of cells were mutagenized with EMS and screened for isolates that formed blue colonies on TBAB (a rich medium that contains methionine) containing X-gal, indicating loss of repression of the *yitJ-lacZ* fusion during growth in the presence of methionine. At least one mutant exhibiting the derepressed phenotype was obtained from each EMS-mutagenized pool, and a total of 10 isolates that formed blue colonies on TBAB containing X-gal were isolated from approximately 7000 colonies.
Figure 4.1. Strategy for isolation of S box-specific mutants. Random mutagenesis of independent pools of strain IS56B containing a yitJ-lacZ fusion was carried out with EMS. After plating on TBAB containing X-gal, isolates that produced blue colonies were tested for cis-acting mutations in the yitJ-lacZ fusion by transferring the fusion into strain ZB307A. The isolates that did not have a mutation in the yitJ-lacZ fusion were tested to insure that the β-galactosidase activity did not result from activation of the lacA gene, a cryptic β-galactosidase gene encoded on the B. subtilis chromosome (Ogura et al., 1999). Plasmid pFGneo, a derivative of plasmid pFG328 (Grundy et al., 1993) with the cat gene replaced by a neomycin resistance gene cassette, was introduced into these isolates by transformation, selecting for replacement of cat and the yitJ leader in the fusion with a neomycin resistance cassette.

The next step was determining if any of the mutations were cis-acting mutations in the yitJ-lacZ fusion (Fig. 4.1). It was likely that we would isolate mutations in the yitJ leader in the transcriptional fusion (in the helix 1-4 region or in the terminator) that lead
to a loss of repression during growth in the presence of methionine, similar to the site-directed mutations previously introduced in this region (Grundy and Henkin, 1998; Winkler et al., 2001; Chapter 3). The fusions from the IS56B mutant strains were transferred into a clean background to screen for and eliminate isolates containing fusions that expressed β-galactosidase in the unmutagenized strain. Chromosomal DNA was prepared for each IS56B variant and introduced into *B. subtilis* strain ZB307A by transformation, selecting for recombination of the resident SPβ prophage in ZB307A with the SPβ sequences flanking the yitJ-lacZ fusion in the chromosome of the mutant strains. Eight of the ten isolates were eliminated at this stage since introduction of the yitJ-lacZ fusion from these strains into ZB307A produced variants that formed blue colonies on TBAB containing X-gal, an indication that the mutation leading to the S box-derepressed phenotype in these isolates was in the yitJ-lacZ fusion.

The two remaining isolates were tested to insure that the β-galactosidase activity did not result from activation of the lacA gene, a cryptic β-galactosidase gene encoded on the *B. subtilis* chromosome (Ogura et al., 1999). Plasmid pFGneo a derivative of plasmid pFG328 (Grundy et al., 1993) with the cat gene replaced by a neomycin resistance gene cassette, was introduced into these isolates by transformation, selecting for replacement of cat and the yitJ leader in the fusion with a neomycin resistance cassette. One isolate formed blue, neomycin-resistant, chloramphenicol-sensitive colonies on TBAB containing X-gal after removal of the yitJ-lacZ fusion, indicating that the mutation in this variant resulted in expression of lacA. The mutation in the remaining isolate was designated “S box-derepressed mutation 1” (SBD1), as it resulted in derepression of S box gene expression during growth in the presence of methionine.
4.3.2 Determination of the specificity of *B. subtilis* strain IS56B-SBD1

The specificity of the SBD1 allele for loss of repression of S box gene expression was tested by introducing a series of transcriptional fusions into the chromosome of IS56B-SBD1 from which the *yitJ-lacZ* leader had been removed. First the *yitJ-lacZ* fusion was reintroduced, resulting in blue colonies on TBAB containing X-gal, as previously observed. A *ykrT-lacZ* fusion was tested to determine if loss of repression during growth in the presence of methionine would be observed with another S box gene fusion; the derepressed phenotype was seen for *ykrT-lacZ* as well. A fusion containing a heterologous terminator from *B. subtilis rpsD* (which encodes ribosomal protein S4) was introduced into IS56B-SBD1 to test whether the SBD1 allele generally increases terminator readthrough, but no increase in β-galactosidase expression was observed. A *tyrS-lacZ* fusion was tested to determine if the SBD1 allele had any effect on the T box regulatory mechanism. Light blue colonies on TBAB containing X-gal were observed, indicating a low level of expression, consistent with the level of expression of this fusion in a wild-type strain under these growth conditions (Grundy and Henkin, 1993). The effect of the SBD allele on expression of a fusion containing the *yitJ*-Pst-3 mutation was also tested. The Pst-3 mutation prevents formation of the antiterminator, leaving the terminator intact and resulting in loss of *yitJ-lacZ* expression under any growth condition (Grundy and Henkin, 1998). IS56B-SBD1 containing the *yitJ*-Pst-3 fusion produced white colonies on TBAB containing X-gal, demonstrating that the SBD1 allele does not bypass the requirement for antiterminator formation for S box gene expression. These results confirmed that the SBD1 allele has a specific effect on S box gene expression.
4.3.3 *In vivo expression of yitJ-lacZ in strains containing the SBD1 allele*

Expression assays were performed for *yitJ-lacZ* in IS56B and IS56B-SBD1, to determine the relative levels of S box gene expression in these two strains during growth in the presence or absence of methionine. The *yitJ-lacZ* fusion in strain IS56B-SBD1 exhibited a small increase in expression during growth in the absence of methionine and a 260-fold increase in expression during growth in the presence of methionine as compared to the parent strain (Table 4.2).

<table>
<thead>
<tr>
<th>Strain</th>
<th>β-galactosidase activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>- Met</td>
</tr>
<tr>
<td>IS56B::yitJ-lacZ</td>
<td>82</td>
</tr>
<tr>
<td>IS56B-SBD1::yitJ-lacZ</td>
<td>123</td>
</tr>
<tr>
<td>BR151::yitJ-lacZ</td>
<td>520</td>
</tr>
<tr>
<td>BR151-SBD1::yitJ-lacZ</td>
<td>208</td>
</tr>
</tbody>
</table>

Table 4.2. **Effect of the SBD1 allele on expression of a yitJ-lacZ transcriptional fusion.** Fusions were integrated in single copy in strains IS56B (*lys-3 metB relA1 sup-3 strA*) and BR151 (*lys-3 metB10 trpC2*). Cells were grown in Spizizen minimal medium (Anagnostopoulos and Spizizen, 1961) containing methionine, harvested by centrifugation, and resuspended in Spizizen minimal medium in the presence (+ Met) or absence (- Met) of methionine. Samples for the β-galactosidase assay were taken 4 h after the cultures were split. β-galactosidase activity is expressed in Miller units (Miller, 1972).
The SBD1 allele was moved from strain IS56B into strain BR151 to look at the effect of the mutation on expression levels of \textit{yitJ-lacZ} in an otherwise unmutagenized background, in the absence of the \textit{relA1} and \textit{sup3} alleles. Chromosomal DNA from IS56B-SBD1 was introduced into strain BR151::\textit{yitJ-lacZ} by transformation, and transformants were screened for blue colonies on TBAB containing X-gal. Expression of the \textit{yitJ-lacZ} fusion was higher in the BR151 strain background, but was reduced two-fold in BR151-SBD1 as compared to BR151 during growth in the absence of methionine, in contrast to the small increase in expression conferred by the SBD1 allele in the IS56B strain background. These differences may be due to effects of the \textit{relA1} and \textit{sup3} alleles, or other unknown differences between the strains. The key result again is that the SBD1 allele results in a partial loss of repression during growth in the presence of methionine (Table 4.2). Expression of the \textit{yitJ-lacZ} fusion during growth in the presence of methionine in BR151-SBD1 was 560-fold greater than in wild-type BR151. The observation that the SBD1 allele elicits the same phenotype in the BR151 background as in strain IS56B indicates that the SBD1 allele is sufficient for this effect.

\subsection{Identification of the SBD1 allele}

The SBD1 allele was genetically mapped using a set of \textit{B. subtilis} mapping strains obtained from the \textit{Bacillus} Genetic Stock Center. Each of these mapping strains has an MLS resistance marker inserted at a unique location in the chromosome (Vandeyar and Zahler, 1986). Generalized transduction using phage PBS-1 was employed to determine the linkage between the SBD1 allele and the MLS resistance marker from each mapping strain. PBS-1 transducing lysates were generated in each of the mapping strains and used
to transduce BR151-SBD1::yitJ-lacZ. MLS-resistant transductants were screened for formation of white colonies on TBAB containing X-gal, indicating replacement of the SBD1 allele with the wild-type allele by cotransduction with the MLS resistance marker from the donor strain. Approximately 60% of the MLS-resistant transductants obtained using lysates from strain 1A642 formed white colonies on TBAB containing X-gal, indicating linkage of the mutation to the MLS resistance marker in this strain. The MLS resistance marker in strain 1A642 is 85% linked to ald (Vandeyar and Zahler, 1986), which is located at 3277.3 kb on the B. subtilis chromosome (Zhang et al., 1999). The metK gene, encoding SAM synthetase is located at 3128.1 kb on the B. subtilis chromosome; this suggested the possibility that the SBD1 allele represents an alteration in the metK gene.

Multiple PCR products covering the entire metK gene and leader region were generated from strains BR151 and BR151-SBD1, and the DNA sequence was determined. A single nucleotide substitution was identified in the metK coding region that would result in replacement of alanine with threonine at position 83 (A83T). A 350 bp fragment of metK containing the SBD1 allele was cloned into plasmid pGEM7Zf+ and the resulting construct was introduced into wild-type BR151::yitJ-lacZ by transformation. Transformants were screened for formation of blue colonies on TBAB containing X-gal, indicating introduction of the SBD1 allele into metK in the BR151 chromosome by homologous recombination with the metK fragment containing the SBD1 allele. Chromosomal DNA was isolated from transformants and the region of interest was isolated by PCR and screened for the presence of the SBD1 allele by DNA sequencing. All the transformants that formed white colonies on TBAB containing X-gal possessed
wild-type *metK* sequence, while all of the transformants that formed blue colonies (indicating loss of repression during growth in the presence of methionine) contained the A83T mutation. These results demonstrated that the A83T substitution in MetK is sufficient to elicit the phenotype observed for the SBD1 allele.

The *B. subtilis* and *E. coli* SAM synthetase enzymes are 57% identical at the level of amino acid sequence (Yocum *et al.*, 1996). SAM synthetase homologs were identified in 69 additional bacterial species representing 46 different genera. Two signature amino acid patterns found in all known SAM synthetases were used to determine which of the sequences obtained from the tblastn program were actually SAM synthetase homologs. The identity of the amino acid analogous to A83 in *B. subtilis* MetK was alanine in MetKs from 28 species representing 16 genera, serine in MetKs from 32 species representing 25 genera, threonine in MetKs from 8 species representing 4 genera, and glycine in *Lactobacillus plantarum* MetK. Of the 28 species containing alanine at this position in MetK, 20 are Gram-positive, including 6 additional *Bacillus* species. All 32 of the species containing serine at this position in MetK are Gram-negative, including *E. coli*. Of the 8 species containing threonine at this position in MetK, 6 species are Gram-negative; the two Gram-positive species are *Streptococcus mutans* and *Streptococcus pneumoniae*. The A83T mutation in the SBD1 allele appears to be a conservative change, consistent with SAM synthetase being an essential enzyme in *B. subtilis*, as the analogous residue in 8 other species was found to be threonine, which is similar to alanine (Yocum *et al.*, 1996). The sequences present in the region surrounding A83 from *B. subtilis* MetK and the MetKs from other representative species are shown in Figure 4.2.
Figure 4.2. Sequence comparison of bacterial MetK homologs. The region of *B. subtilis* MetK surrounding residue A83 is aligned with the analogous regions of MetKs from other bacterial species. Sequences were obtained from the tblastn program (Altschul et al., 1990). A83 in *B. subtilis* MetK and the analogous residues from the other bacterial MetKs are boxed. Residues that differ from the *B. subtilis* sequence are shown in red. Bsu, *B. subtilis*; Ban, *Bacillus anthracis*; Sau, *Staphylococcus aureus*; Eco, *E. coli*; Sty, *Salmonella typhimurium*; Pae, *Pseudomonas aeruginosa*; Spn, *Streptococcus pneumoniae*; Bpe, *Bordetella pertussis*; Bma, *Burkholderia mallei*; Lpl, *Lactobacillus plantarum*. (+), Gram-positive species; (-), Gram-negative species. Amino acid identity to *B. subtilis* MetK: Ban, 86%; Sau, 77%; Eco, 57%; Sty, 63%; Pae, 59%; Spn, 68%; Bpe, 58%; Bma, 59%; Lpl, 68%.

4.3.5 Measurement of SAM synthetase activity in BR151-SBD1

SAM synthetase assays were performed to determine the level of SAM synthetase activity in BR151-SBD1 compared to the BR151 parent strain during growth in minimal medium containing methionine. SAM synthetase activity in the parent strain was 150±7 pmol min⁻¹ mg of protein⁻¹, consistent with the range of activity (128 to 200 pmol min⁻¹ mg of protein⁻¹) previously reported for strains of *B. subtilis* containing wild-type copies
of the *metK* gene (Ochi and Freese, 1982; Wabiko *et al.*, 1988; Yocum *et al.*, 1996). Activity in BR151-SBD1 was $10\pm2$ pmol min$^{-1}$ mg of protein$^{-1}$, which indicates that the SBD1 allele results in a 15-fold reduction in SAM synthetase activity. The growth rate of BR151-SBD1 was the same as the growth rate of BR151 in the presence of methionine, suggesting that the observed decrease in SAM synthetase activity does not have a significant effect on growth rate, as previously observed for other *metK* mutant strains (Wabiko *et al.*, 1988).

### 4.3.6 Effect of *metK* overexpression on *yitJ-lacZ* expression *in vivo*

Experiments were also performed to observe the effect of increased SAM synthetase activity on S box gene expression *in vivo*, since SAM was shown to be the molecular effector for S box termination *in vitro*, and to compare to the observed effect of decreased SAM synthetase activity on *yitJ-lacZ* expression *in vivo*. Expression of a *yitJ-lacZ* transcriptional fusion was assayed in *B. subtilis* strain SA29, in which *metK* is expressed under the control of the strong P$_{veg}$ promoter as compared to a wild-type strain (168). During growth in minimal media containing methionine, SAM synthetase activity in strain SA29 was determined to be increased six-fold as compared to the wild-type strain (Yocum *et al.*, 1996). Strains 168 and SA29 are methionine prototrophs, and S box genes are expressed at a lower level during growth in the absence of exogenous methionine in methionine prototrophs than in methionine auxotrophs, due to the methionine generated in the cell (Grundy and Henkin, 1998). Expression of the *yitJ-lacZ* fusion in strain 168 was induced rapidly after methionine was removed from the growth medium, while strain SA29 showed a delayed response to removal of methionine.
(McDaniel et al., 2003; Fig. 4.3). This is consistent with the prediction that strain SA29 should have increased SAM pools due to overproduction of SAM synthetase, and thus a longer growth period in the absence of methionine is required for depletion of SAM pools and subsequent induction of yitJ-lacZ expression.

Figure 4.3. Effect of overexpression of SAM synthetase on expression of a yitJ-lacZ transcriptional fusion. Fusions were integrated in single copy in strains 168 (trpC2) and SA29 (trpC2 metK1 sacB::φP_peg metK'). Cells were grown in Spizizen minimal medium (Anagnostopoulos and Spizizen, 1961) containing methionine, harvested by centrifugation, and resuspended in Spizizen minimal medium in the presence (filled symbols) or absence (open symbols) of methionine. Samples for the β-galactosidase assay were taken at 1 h intervals after the cultures were split. β-galactosidase activity is expressed in Miller units (Miller, 1972).
4.4 Discussion

In this study we attempted to isolate trans-acting mutations that would lead to loss of repression of S box gene expression during growth in the presence of methionine, either in genes that affect the molecular effector (SAM), or in other factors that are involved in S box gene regulation in vivo. A trans-acting mutation (SBD1) was isolated that led to loss of repression of S box gene repression during growth in the presence of methionine. The SBD1 allele was identified as a single nucleotide substitution in the metK gene, which encodes SAM synthetase. The metK gene product converts methionine and ATP into SAM, which is required for methyltransferase reactions in the cell as well as for polyamine biosynthesis (Matthews, 1996). The level of intracellular methionine in the cell can therefore be measured indirectly by sensing SAM levels, as methionine must be present for SAM to be synthesized. The key result is that the SBD1 allele results in a reduction in SAM synthetase activity, consistent with the prediction that SAM could act as a corepressor of methionine biosynthesis genes in B. subtilis, as in E. coli (Yocum et al., 1996; Grundy and Henkin, 1998, 2002, 2003), and our identification of SAM as the molecular effector for S box gene expression (McDaniel et al., 2003; Chapter 2). The metK gene is likely to be regulated by SAM (the metabolite synthesized by its gene product), as it is a member of the S box regulon, and SAM is the molecular effector of S box gene expression.

The only trans-acting mutation identified in this study was in metK, which suggests that perhaps no additional factors are involved in S box gene regulation in vivo. However, the possibility remains that other mutations were not identified because the putative additional factors are essential gene products that play a role in cellular
processes other than regulation of S box gene expression. *In vivo* expression of a yitJ-*lacZ* transcriptional fusion in strains containing the SBD1 allele indicated a partial loss of repression during growth in the presence of methionine. Inactivation of the *metK* gene is a lethal event, as some level of SAM synthetase activity is essential for viability (Yocum et al., 1996). Therefore it was only possible to isolate mutations that decrease SAM synthetase activity, not abolish it, and the residual SAM synthetase activity results in some repression during growth in the presence of methionine. The observation that an increase in SAM synthetase activity *in vivo* resulted in delayed induction of yitJ-*lacZ* expression in response to methionine starvation is consistent with the results observed for the SBD1 mutation, as increased SAM synthetase activity represses S box gene expression, while decreased SAM synthetase activity results in a partial loss of repression of S box gene expression during growth in the presence of methionine.

It is possible that the expression of *metK* is increased in strains containing the SBD1 allele, to compensate for the loss of SAM synthetase activity. Wabiko et al. (1988) observed a 25- to 200-fold decrease in SAM synthetase activity in *metK* mutant strains of *B. subtilis* compared to the parent strain, yet these strains exhibited only a three- to four-fold lower SAM pool during growth in the presence of methionine, which could be due to increased *metK* expression. An increase in methionine pools has been observed in strains with decreased SAM synthetase activity (due to either increased methionine production or decreased methionine utilization), and addition of extracellular methionine further increases the SAM pools in some of the mutant strains (Wabiko et al., 1988).
The amino acid sequence of the SAM synthetase enzyme is highly conserved among bacteria and eukarya. *E. coli* SAM synthetase consists of four identical subunits (Markham *et al*., 1980). The crystal structure of this enzyme revealed that two subunits form a tight dimer with two active sites located between them, and two dimers interact in an asymmetrical arrangement to form a peanut-shaped tetrameric enzyme (Takusagawa *et al*., 1996). The interactions between the dimers appear to be less extensive than the subunit-subunit interactions within each dimer (Takusagawa *et al*., 1996; Komoto *et al*., 2004). Residue S80, which is located at the tetramer interface, is the only residue that participates in hydrogen bonds between the dimers (Takusagawa *et al*., 1996; Fig. 4.4). The altered residue in the SBD1 variant, A83T, is analogous to S80 in *E. coli* MetK. It is therefore possible that substitution of the non-polar alanine residue with a polar threonine residue at position 83 in *B. subtilis* MetK could affect activity by affecting tetramerization. It is also possible that context plays a role in the potential effect of the A83T mutation, as other bacterial MetKs have threonine at this position.
Figure 4.4. Location of S80 in the crystal structure of the *E. coli* SAM synthetase ternary complex. The crystal structure of *E. coli* SAM synthetase with the ATP analogue AMPPNP and methionine in each active site is shown (Komoto *et al.*, 2004). The tetrameric enzyme is a dimer of dimers, with two active sites located between the two subunits in each dimer. The location of S80 (in tan) in each subunit is indicated by red arrows. S80 is the only residue that participates in hydrogen bond formation between the dimers of *E. coli* SAM synthetase. The image was created using the Cn3D v4.1 3D-structure viewer (NCBI).

Tetramerization of *E. coli* SAM synthetase is required for full activity (Markham and Satishchandran, 1988). Modification of both C90 (located at the tetramer interface) and C240 by *N*-ethylmaleimide results in dissociation of *E. coli* SAM synthetase active tetramers into inactive dimers. Site-directed mutagenesis of C90 to alanine or serine
results in a mixture of dimers and tetramers and a reduction in SAM synthetase activity; the purified dimers exhibit approximately 20-fold lower activity than the purified tetramers (Reczkowski and Markham, 1995), a reduction in activity similar to that observed for the SBD1 allele. C90 is a conserved amino acid residue in SAM synthetases, indicating that this residue could play an important role in activity of the enzyme (Reczkowski and Markham, 1995; Takusagawa et al., 1996). Mingorance et al. (1996) employed site-directed mutagenesis to determine the functional importance of cysteine residues in rat liver SAM synthetase, and found that replacement of C150, which is analogous to C90 in E. coli SAM synthetase, also results in a mixture of dimers and tetramers and decreased SAM synthetase activity.

Mutations that affect the oligomerization state of an enzyme can affect the activity of that enzyme in multiple ways. A mutation located at the dimer interface of tyrosyl-tRNA synthetase (TyrRS) results in reversible dissociation of the dimer into monomers at high pH. The monomers do not catalyze amino acid activation, while the dimer formed at low pH does catalyze amino acid activation (Jones et al., 1985). Residues located at the dimer interface of the tryptophanyl-tRNA synthetase (TrpRS) enzymes of B. subtilis (Chow et al., 1992), E. coli (Sever et al., 1996), and Bacillus stearothermophilus (Acchione et al., 2003) have been shown to be essential for proper structure and function of the enzyme, as mutation of these residues results in decreased or abolished enzyme activity. Mutation of the residues located at the dimer interface in E. coli TrpRS is proposed to affect dimerization and proper alignment of the active sites, which is necessary for efficient tryptophan binding and optimal catalysis (Sever et al., 1996). Mutation of the residues located at the dimer interface in B. stearothermophilus TrpRS is
proposed to affect cross-subunit communication involving the dimer interface that results in a structural change in the second subunit (Acchione et al., 2003). The A83T mutation could affect activity of \textit{B. subtilis} MetK in a similar manner, by affecting tetramerization of the enzyme and active site alignment, or by affecting communication between the two dimers.

In summary, we isolated a \textit{trans}-acting mutation in the gene that encodes SAM synthetase which resulted in a decrease in SAM synthetase activity and derepression of \textit{S} box gene expression during growth in the presence of methionine, consistent with our identification of SAM as the molecular effector for \textit{S} box gene expression. The mutation was identified as an amino acid substitution at the residue in SAM synthetase (A83) that is analogous to the only residue that participates in hydrogen bond formation between the dimers of \textit{E. coli} SAM synthetase, and thus could affect SAM synthetase activity by preventing tetramerization of the enzyme. The observation that increased SAM pools due to increased SAM synthetase activity resulted in delayed induction of \textit{S} box gene expression in response to starvation for methionine \textit{in vivo} is also consistent with our identification of SAM as the molecular effector for \textit{S} box gene expression. \textit{In vitro} transcription assays indicated that SAM promotes premature termination of \textit{S} box gene expression in the absence of any additional factors (McDaniel \textit{et al}, 2003; Chapter 2). We were unable to identify a \textit{trans}-acting mutation in any gene other than \textit{metK}, which suggests the possibility that no additional factors are involved in \textit{S} box gene regulation \textit{in vivo}, unless additional \textit{trans}-acting mutations were not identified because the putative additional factors are essential gene products that participate in cellular processes other than regulation of \textit{S} box gene expression.
CHAPTER 5

IDENTIFICATION OF THE GENE ENCODING
5-METHYLTHIORIBOSE KINASE IN Bacillus subtilis

5.1 Introduction

Of the 26 genes originally identified as members of the S box family in B. subtilis, only three (cysH, metE [formerly metC], and metK [formerly metE]) had been shown to be involved in cysteine and methionine metabolism (Grundy and Henkin, 1998). Roles in pathways for cysteine and methionine biosynthesis were predicted for 11 of the genes of unknown function, based on similarity to genes of known function (Grundy and Henkin, 1998, 2003). It has been demonstrated that expression of several of these genes is repressed during growth in the presence of methionine (Grundy and Henkin, 1998; Murphy et al., 2002; Auger et al., 2002; Hullo et al., 2004; Chapter 1). Conservation of the regulatory pattern of S box leader RNAs suggested that genes with these conserved features have unknown functions in methionine metabolism. Since B. subtilis is efficient at scavenging of sulfur-containing compounds, it seemed likely that some of the S box genes of unknown function could encode products involved in alternate routes to biosynthesis of methionine.
5’-methylthioadenosine (MTA) is a sulfur-containing compound that is generated as a byproduct of polyamine biosynthesis and can be recycled to methionine by a number of organisms, including *B. subtilis* and *Klebsiella pneumoniae* (Trackman and Abeles, 1983; Furfine and Abeles, 1988; Sekowska and Danchin, 1999; Fig. 5.1). Decarboxylated SAM is generated by the *speD* gene product, and the aminopropyl moiety of decarboxylated SAM is then transferred to putrescine by spermidine synthase, the product of the *speE* gene (Sekowska et al., 1998), generating spermidine and MTA. MTA is cleaved by the *mtn* gene product, MTA/S-adenosylhomocysteine nucleosidase (the *pfs* gene product in *E. coli*) to yield adenine and 5-methylthioribose (MTR; Duerre, 1962; Cornell et al., 1996; Cornell et al., 1998; Sekowska and Danchin, 1999). MTR is excreted in *E. coli* (Schroeder et al., 1973; Sekowska et al., 2000), whereas *B. subtilis* can use either MTA or MTR as the sole sulfur source (Sekowska and Danchin, 1999). The gene products of *metK* (formerly designated *metE*; Yocum et al., 1996), *speD* (Sekowska et al., 2000), *speE* (Sekowska et al., 1998), and *mtn* (Sekowska and Danchin, 1999) catalyze the enzymatic steps generating MTR from methionine (Fig. 5.1).
Figure 5.1. Predicted methionine biosynthesis pathways in *B. subtilis*. Genes that are preceded by S box regulatory elements are shown with a boxed “S.” SAM, S-adenosylmethionine; SAH, S-adenosylhomocysteine; SRH, S-ribosylhomocysteine; MT, methylthio; KMTB, 2-keto-4-methylthiobutyrate; THF, tetrahydrofolate; MTA, methylthioadenosine; MTR, methylthioribose; α-KB, α-ketobutyrate. Modified from Murphy et al. (2002).
The MTA recycling pathway is best characterized in \textit{K. pneumoniae} (Trackman and Abeles, 1983; Furfine and Abeles, 1988; Sekowka \textit{et al.}, 2000, 2004). Once MTA has been converted to MTR, MTR phosphate (MTRP) is generated by MTR kinase (Ferro \textit{et al.}, 1978). At the start of this work, the intermediates in the conversion of MTRP to methionine in the \textit{K. pneumoniae} MTA recycling pathway were known, and the isomerase that catalyzes the conversion of MTRP to 5-methylthioribulose-1-phosphate had been identified (Furfine and Abeles, 1988). However, few of the genes involved had been identified. The \textit{ykrTS} and \textit{ykrWXYZ} operons in \textit{B. subtilis} (Fig. 5.2) are members of the S box family, and are regulated in response to methionine availability (Grundy and Henkin, 1998; Murphy \textit{et al.}, 2002). Initial searches for proteins related to \textit{B. subtilis} YkrT (performed by Frank Grundy) revealed similarity to aminoglycoside transferases. Since MTR is structurally related to aminoglycosides, this observation suggested that YkrT could function in phosphorylation of MTR to generate MTRP. This prediction was later corroborated by identification of the \textit{K. pneumoniae} gene encoding MTR kinase, which is closely related to \textit{B. subtilis} \textit{ykrT}; levels of MTR kinase in \textit{K. pneumoniae} were shown to be reduced during growth in the presence of methionine (Tower \textit{et al.}, 1993), consistent with a role in methionine production. The similarity of \textit{B. subtilis} \textit{ykrT} to the \textit{K. pneumoniae} gene encoding MTR kinase supported the prediction that \textit{ykrT} could be involved in the MTA recycling pathway (Grundy and Henkin, 2002). Cotranscription of \textit{ykrS} with \textit{ykrT} (Fig. 5.2) suggested that it could be involved in the same pathway. Homologs of \textit{ykrS} have been identified in many prokaryotes (Kyprides \textit{et al.}, 1998; Sekowska \textit{et al.}, 2004), and the similarity to eukaryotic eIF2Bα (Grundy and Henkin, 1998) has raised interesting questions about its distribution and functional role, since
eIF2Bα plays an important role in regulation of translation initiation in eukaryotic cells (Clemens, 1996).

**Figure 5.2.** Region of the *B. subtilis* chromosome containing the *ykrTS* and *ykrWXYZ* operons. Expression of the *ykrTS* operon is under the control of the *ykrT* promoter and leader region, and expression of the *ykrWXYZ* operon is under the control of the *ykrW* promoter and leader region (Grundy and Henkin, 1998). The *ykrU* and *ykrV* genes are located between the *ykrTS* and *ykrWXYZ* operons.

The *ykrTS* operon is located close to the divergently transcribed *ykrWXYZ* operon (Fig. 5.2). Two additional genes of unknown function, *ykrU* and *ykrV*, are located between these two S box operons (Kunst *et al.*, 1997). Comparison of the possible enzymatic functions of these gene products, derived through homology searches, with the
intermediates in the MTA recycling pathway detected in *K. pneumoniae* (Furfine and Abeles, 1988) suggested the possibility that many (or all) of these genes could be involved in this process (Grundy and Henkin, 2002; Fig. 5.1). The observation that YkrW is related to ribulose-1,5-bisphosphate carboxylase (RuBisCO, the key enzyme in CO₂ fixation; Grundy and Henkin, 1998) and similarity of 5-methylthioribulose-1-phosphate (an MTA recycling pathway intermediate) to ribulose-1,5-bisphosphate, the substrate of RuBisCO, suggested a role for *ykrW* in this pathway (Grundy and Henkin, 2002). The combination of the regulatory clues and sequence similarity to proteins of known function led us to investigate roles for the *ykrTS* and *ykrWXYZ* gene products in MTA recycling.

5.2 Materials and Methods

5.2.1 Bacterial strains and growth conditions

The bacterial strains used in this study were BR151 (*lys*-3 *metB10* *trpC2*), BR151MA (*lys*-3 *trpC2*), BR151-ZKO (*lys*-3 *metB10* *trpC2* *tyrZ::neo*; Grundy and Henkin, 1998), BR151-YkrSKO (*lys*-3 *metB10* *trpC2* *ykrS::neo*; this study), BR151-YkrTKO (*lys*-3 *metB10* *trpC2* Δ*ykrT* *tyrZ::neo*; this study), and BR151MA-YitJKO (*lys*-3 *trpC2* *yitJ::neo*; this study). *B. subtilis* strains were grown on tryptose blood agar base medium (TBAB; Difco), 2XYT broth (Miller, 1972), or Spizizen minimal medium (Anagnostopoulus and Spizizen, 1961). Antibiotics were added at the following concentrations: chloramphenicol, 5 µg/ml; neomycin, 5 µg/ml; spectinomycin, 200 µg/ml. MTA was purchased from Sigma. MTR was prepared from MTA by acid hydrolysis with assistance from Cedric Bobst (Schlenk *et al.*, 1973). Methionine was
added to 0.34 mM, and MTA and MTR were added to 0.4 mM in liquid minimal media. All growth was at 37°C and was measured with a Klett-Summerson colorimeter using the No. 66 red filter.

5.2.2 Genetic techniques

Restriction endonucleases and T4 DNA ligase were purchased from New England Biolabs and used as described by the manufacturer. Oligonucleotide primers were purchased from Integrated DNA Technologies (Coralville, IA). Wizard columns (Promega) were used for plasmid preparations. Transformation of B. subtilis was carried out as described by Henkin and Chambliss (1984). The yitJ-lacZ, ykrT-lacZ and ykrW-lacZ transcriptional fusions were generated in plasmid pFG328 (Grundy and Henkin, 1993) and integrated in single copy into the chromosome using specialized transducing phage SPβ, as previously described (Grundy and Henkin, 1993; Grundy and Henkin, 1998). Strains containing lacZ fusions were grown in the presence of chloramphenicol.

5.2.3 Gene inactivation

A 620 bp fragment containing the amino-terminal coding region of ykrS was generated by PCR using primers YKRTKO1 (GTTACTTCTAGAGGAATGATCTATTCATGACCCAT) and YKRTKO2 (GTAAGGCGTCGACCTTGCAGGACAGGCC), digested with XbaI and SalI, and inserted immediately downstream of the ykrT leader region terminator in the pFG328 construct used to generate the ykrT-lacZ fusion (Grundy and Henkin, 1998; Fig. 5.3). The resulting plasmid was introduced into strain BR151 by congression to generate strain
BR151-YkrTKO, with selection for introduction of the \textit{tyrZ::neo} allele of BR151-ZKO. Homologous recombination at the \textit{ykrTS} chromosomal locus resulted in deletion of the \textit{ykrT} coding region, and expression of \textit{ykrS} was left under the control of the \textit{ykrT} promoter and regulatory region. The \textit{ykrS::neo} mutant (Grundy and Henkin, 1998) was introduced into strain BR151 by transformation to generate strain BR151-YkrSKO. Strain BR151MA-YitJKO was generated by insertional inactivation using a variant of plasmid pBEST501 (Itaya \textit{et al.}, 1989) containing an internal segment of the \textit{yitJ} coding sequence. This plasmid was introduced into \textit{B. subtilis} strain BR151MA by transformation with selection for neomycin resistance, and integration of the plasmid by homologous recombination resulted in disruption of the chromosomal \textit{yitJ} locus.
Figure 5.3. Plasmid used to generate strain BR151-YkrTKO. A 620 bp fragment containing the amino-terminal coding region of \textit{ykrS} was generated by PCR, digested with \textit{XbaI} and \textit{SalI}, and inserted immediately downstream of the \textit{ykrT} leader region terminator in the pFG328 construct used to generate the \textit{ykrT-lacZ} fusion (Grundy and Henkin, 1998). Restriction sites are labeled with dashed lines; the locations of ampicillin (\textit{Amp}^r) and chloramphenicol (\textit{Cm}^r) resistance cassettes are shown. P, promoter; T, terminator.
5.2.4 β-galactosidase measurements and growth experiments

Cells were grown in Spizizen minimal medium containing all required amino acids at 50 µg/ml until early exponential growth. The cells were harvested by centrifugation, then resuspended in fresh Spizizen minimal medium in the presence or absence of methionine, MTA, or MTR. β-galactosidase activity was assayed in cells containing lacZ transcriptional fusions as described in Chapter 3. Klett readings were taken at 0.5 h intervals for growth experiments.

5.3 Results

5.3.1 Role of ykrTS and ykrWXYZ genes in growth on MTA or MTR

The ability of methionine auxotrophic strains of B. subtilis to use the MTA recycling pathway to generate methionine was tested. Analysis of utilization of alternate sulfur sources in B. subtilis is complicated by its efficient scavenging of sulfur-containing compounds. To avoid residual growth on contaminating compounds in the growth media, we used a genetic background in which the metB gene, encoding homoserine O-acetyltransferase, is defective, so that the normal pathway of methionine biosynthesis is blocked. Strains in which ykrT (BR151-YkrTKO) or ykrS (BR151-YkrSKO) were inactivated were generated and compared to an isogenic control strain (BR151-ZKO). For the ykrT mutant, expression of the downstream ykrS gene was left under the control of the ykrTS promoter and leader region, so that normal expression levels were maintained. No growth was observed for any strain containing the metB allele in media containing sulfate alone, while all of the strains grew equally well in the presence of methionine (Fig. 5.4). MTA or MTR could replace methionine for growth of strain
BR151-ZKO, but inactivation of either \textit{ykrT} or \textit{ykrS} resulted in a total loss of growth in the presence of MTA or MTR, indicating that both genes are required for utilization of these compounds to generate methionine (Murphy \textit{et al.}, 2002; Fig. 5.4). The defect in utilization of MTA and MTR confirmed a role for \textit{ykrT} in the MTA recycling pathway, and represented the first demonstration of a physiological role for the \textit{ykrS} gene in any organism. Insertional inactivation of the \textit{ykrWXYZ} operon (BR151-YkrWKO, generated by Ekta Sirohi) resulted in a partial defect in growth on MTA, indicating that these genes could also be involved in the MTA recycling pathway (Murphy \textit{et al.}, 2002; Fig. 5.4D).
Figure 5.4. Growth of \textit{ykrTS} mutants and \textit{ykrWXYZ} mutants. Cultures were grown overnight in Spizizen minimal medium (Anagnostopoulos and Spizizen, 1961) containing methionine, cells were collected by centrifugation and resuspended in media containing sulfate alone (open circles), methionine (filled circles), MTA (squares), or MTR (triangles). A. BR151-ZKO (\textit{metB}10). B. BR151-YkrTKO (\textit{metB}10 \Delta ykrT). C. BR151-YkrSKO (\textit{metB}10 ykrS::neo). D. BR151MA-YkrWKO (\textit{metB}10 ykrW::neo).
The *B. subtilis* yitJ gene is similar to *E. coli* metF, encoding methylene tetrahydrafolate reductase, which is required for conversion of homocysteine to methionine in the primary pathway of methionine biosynthesis (Grundy and Henkin, 1998, 2002; Fig. 5.1). Inactivation of this gene (in the MetB+ strain BR151MA) resulted in loss of growth in the absence of methionine, consistent with the predicted function of the yitJ product (Murphy *et al.*, 2002). Unlike the ykrTS mutants, growth on MTA was unaffected, indicating that the pathway for methionine generation from MTA/MTR is independent of the normal biosynthetic pathway.

### 5.3.2 Regulation of ykrTS and ykrWXYZ expression

Previous studies indicated that expression of ykrTS is repressed during growth in the presence of methionine (Grundy and Henkin, 1998). Identification of a role in the MTA recycling pathway led us to test expression during growth in the presence of MTA. In strain BR151-ZKO, expression was repressed 1300-fold during growth in the presence of methionine compared to growth in the absence of methionine (Murphy *et al.*, 2002; Table 5.1), as previously reported (Grundy and Henkin, 1998). The ykrT-lacZ fusion was expressed when MTA was supplied in place of methionine, but at a much lower level than in the absence of methionine, consistent with the ability of strain BR151-ZKO to utilize MTA to generate methionine. Growth in the presence of both methionine and MTA resulted in nearly complete repression of ykrT-lacZ expression, indicating that utilization of MTA is repressed during growth in the presence of exogenous methionine (Murphy *et al.*, 2002; Table 5.1). Similar results were obtained with MTR in place of MTA (Murphy *et al.*, 2002).
Table 5.1. Regulation of \textit{lacZ} fusions in response to methionine and MTA. Cells were grown in Spizizen minimal medium (Anagnostopoulos and Spizizen, 1961) containing methionine (0.34 mM), harvested by centrifugation, and resuspended in Spizizen minimal medium in the presence or absence of methionine and/or MTA (0.4 mM). Samples for the $\beta$-galactosidase assay were taken 4 h after the cultures were split. $\beta$-galactosidase activity is expressed in Miller units (Miller 1972).

<table>
<thead>
<tr>
<th>Fusion</th>
<th>Strain</th>
<th>Relevant Genotype</th>
<th>- Met</th>
<th>+ Met</th>
<th>+ MTA</th>
<th>+ Met + MTA</th>
</tr>
</thead>
<tbody>
<tr>
<td>ykrT-lacZ</td>
<td>BR151-ZKO</td>
<td>metB tyrZ::neo</td>
<td>130</td>
<td>0.10</td>
<td>7.9</td>
<td>0.18</td>
</tr>
<tr>
<td>ykrT-lacZ</td>
<td>BR151-YkrTKO</td>
<td>metB $\Delta$ykrT tyrZ::neo</td>
<td>70</td>
<td>0.14</td>
<td>77</td>
<td>0.23</td>
</tr>
<tr>
<td>ykrT-lacZ</td>
<td>BR151-YkrSKO</td>
<td>metB ykrS::neo</td>
<td>50</td>
<td>0.11</td>
<td>61</td>
<td>0.11</td>
</tr>
<tr>
<td>ykrW-lacZ</td>
<td>BR151-ZKO</td>
<td>metB tyrZ::neo</td>
<td>38</td>
<td>0.63</td>
<td>8.0</td>
<td>0.92</td>
</tr>
<tr>
<td>yitJ-lacZ</td>
<td>BR151</td>
<td>metB</td>
<td>340</td>
<td>0.17</td>
<td>10</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Expression of the \textit{ykrT-lacZ} fusion responded normally to methionine availability in strains BR151-YkrTKO or BR151-YkrSKO, in which the \textit{ykrT} or \textit{ykrS} genes were inactivated (Murphy \textit{et al.}, 2002; Table 5.1), although expression under inducing conditions was reduced somewhat, as previously reported for a \textit{ykrS} mutant (Grundy and Henkin, 1998). Addition of MTA had no repressive effect on expression in either strain, in contrast to the results for BR151-ZKO. This is consistent with the inability of the
mutant strains to convert MTA to methionine, so that expression of \( ykrTS \) remained fully derepressed during growth in the presence of MTA (Murphy et al., 2002; Table 5.1). Similar results were obtained with MTR in place of MTA (Murphy et al., 2002).

Expression of a \( ykrW-lacZ \) fusion in strain BR151-ZKO was also induced by starvation for methionine (Murphy et al., 2002; Table 5.1), consistent with the presence of an S box leader and a role in methionine biosynthesis. Expression during growth in the presence of methionine was not as tightly repressed by methionine availability as was expression of \( ykrT \). The expression pattern was similar to that of \( ykrT \) during growth in the presence of MTA.

The effect of MTA on expression of the \( yitJ \) gene, which is in the primary pathway of methionine biosynthesis (Fig. 5.1), was also tested to determine if this compound generally affects expression of S box genes. Expression of a \( yitJ-lacZ \) fusion was induced by starvation for methionine, as previously reported (Grundy and Henkin, 1998), and was moderate during growth in the presence of MTA (Murphy et al., 2002; Table 5.1), as was observed for the \( ykrT-lacZ \) fusion. This indicates that the partial repression of \( ykrTS \) expression observed during growth on MTA and MTR is likely to be a consequence of the intracellular methionine generated by the recycling pathway, and reflects the general pattern of regulation of genes in the S box family rather than a specific regulatory effect for genes directly involved in this pathway.

### 5.4 Discussion

This study demonstrated a direct role for the \( B. subtilis ykrT, ykrS \) and \( ykrW \) genes in the MTA recycling pathway. Identification of these genes as members of the S box
regulon stimulated the initial investigation of the response to methionine limitation, and targeted these genes as likely to function in some way in methionine metabolism. Similarity of \textit{ykrT} to \textit{K. pneumoniae} MTR kinase supported our prediction that YkrT could also function as MTR kinase, which was confirmed by the phenotype of the \textit{ykrT} mutant. Our identification of \textit{ykrT} as the gene encoding MTR kinase in \textit{B. subtilis} was corroborated by Sekowska \textit{et al.} (2001), who also demonstrated that the \textit{ykrT} gene product exhibits MTR kinase activity \textit{in vitro}. Crystallization and preliminary X-ray analysis of MTR kinase from \textit{B. subtilis} have been completed (Ku \textit{et al.}, 2004). Our results on regulation of \textit{ykrTS} expression differ from those of Sekowska \textit{et al.} (2001), who found little effect of addition of exogenous methionine. However, those studies were performed using a methionine prototroph, so that intracellular methionine pools were high during growth in the presence of sulfate, masking the substantial effect of methionine deprivation in cells unable to generate endogenous methionine.

Inactivation of \textit{ykrS} was previously shown to result in reduced expression of several S box genes during methionine limitation, suggesting a possible role in the transcription termination control system (Grundy and Henkin, 1998). The demonstration that inactivation of \textit{ykrS} results in a complete block in MTA/MTR utilization indicated that this gene is likely to play a direct role in the MTA recycling pathway in \textit{B. subtilis}. Ashida \textit{et al.} (2003) confirmed the role of \textit{ykrS} in the MTA recycling pathway, and proposed that it is the gene encoding methylthioribose-1-phosphate isomerase (Fig. 5.1). The basis for its effect on S box gene expression remains to be determined. Since inactivation of \textit{ykrT} resulted in a similar decrease in \textit{ykrT-lacZ} expression, this effect is not specific to \textit{ykrS}, and could be due to disruption of the MTA recycling pathway.
Disruption of the MTA recycling pathway could affect polyamine biosynthesis, resulting in accumulation of SAM, a precursor of polyamine biosynthesis.

The partial defect in growth on MTA that resulted from insertional inactivation of the \textit{ykrWYZX} operon supported a role for these genes in MTA recycling, as did S box regulation of \textit{ykrWXYZ} in a pattern similar to that of \textit{ykrTS}. Subsequent studies confirmed our prediction that \textit{ykrU} and \textit{ykrV} and the genes in the \textit{ykrWXYZ} operon are involved in the MTA recycling pathway in \textit{B. subtilis} (Sekowska and Danchin, 2002; Berger et al., 2003; Ashida et al., 2003). Sekowska and Danchin (2002) demonstrated that YkrU, YkrW, YkrX, YkrY, and YkrZ are all required for MTA recycling in \textit{B. subtilis}. YkrV catalyzes the final step in the MTA recycling pathway, conversion of ketomethiobutyrate to methionine by an aminotransferase reaction (Berger et al., 2003; Fig. 5.1), as predicted by Grundy and Henkin (2002). Ashida and coworkers (2003) identified the gene product of \textit{ykrW} as the enzyme that catalyzes the 2,3-diketo-5-methylthiopentyl-1-phosphate enolase reaction in the methionine salvage pathway (Fig. 5.1). It was also demonstrated that this function of the RubisCO-like protein YkrW is conserved in the RubisCO from the photosynthetic bacterium \textit{Rhodospirillum rubrum}, as the photosynthetic RubisCO rescues the defect in the ability of a \textit{B. subtilis ykrW} mutant to grow with MTA in place of methionine (Ashida et al., 2003). A RubisCO-like protein that is involved in sulfur metabolism and the oxidative stress response has also been identified in \textit{Chlorobium tepidum} (Hanson and Tabita, 2001, 2003). Ashida et al. (2003) also confirmed roles for \textit{ykrX} (2-hydroxy-3-keto-5-methylthiopentenyl-1-phosphate-phosphatase), \textit{ykrY} (methylthioribulose-1-phosphate dehydratase), and \textit{ykrZ} (1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase) in MTA recycling (Fig. 5.1).
In mammalian cells MTA is converted in one step to adenine and MTRP by MTA phosphorylase (Pegg et al., 1969). MTR kinase was identified as a target for antimicrobial agents since it is required for methionine salvage in certain prokaryotes but is absent in mammalian cells (Gianotti et al., 1990). 5-Trifluoromethylthioribose, a structural analog of MTR, inhibits the growth of *K. pneumoniae* and inhibits MTR kinase activity, demonstrating the potential for selective killing of MTR kinase-containing microorganisms (Gianotti et al., 1990). Analogs of MTR have also been identified as antiprotozoal agents (Riscoe et al., 1988). Further analysis of the structure and function of MTR kinase, with the potential insight provided from the well-characterized eukaryotic protein kinases, may be helpful in the development of this system as an antimicrobial target. YkrW has also been identified as a target for new antimicrobial drugs, as MTR becomes extremely toxic to the cell in the absence of YkrW (Sekowska and Danchin, 2002). This is of great interest, as methionine salvage pathways have been identified in several other organisms in addition to *B. subtilis* and *K. pneumoniae*, including *Bacillus anthracis*, *Pseudomonas aeruginosa*, *Leptospira interrogans*, *Thermoanaerobacter tencongensis*, *Xylella fastidiosa*, and *Arabidopsis thaliana* (Sekowska et al., 2004).

This study demonstrated that identification of sequence elements involved in regulation of gene expression in response to a particular physiological signal can be a useful tool in determining the physiological roles of genes of unknown function. Expression patterns or effects of gene inactivation often provide insight into gene function; however, in the absence of this additional information it is difficult to predict functional roles for newly identified open reading frames in genomic sequence. The S
box regulon represents an example where easily recognized regulatory signals provide a clear indication of metabolic function (Grundy and Henkin, 2002; Murphy et al., 2002).
SUMMARY AND FUTURE DIRECTIONS

Several genetic systems have been recently identified in which the nascent transcript directly senses a regulatory signal in the absence of any accessory proteins, and expression of the downstream coding region(s) is controlled by modulation of the RNA structure in response to the regulatory signal. Because these systems exhibit a high degree of specificity and affinity for their particular effector molecules, they allow the cell to specifically and efficiently regulate gene expression in response to changes in the environment. The main goal of this project was to reveal details of the molecular mechanism of transcription termination control governing the S box regulon, which enabled us to show that the S box system is a member of this growing family of gene regulation systems.

*In vitro* transcription assays demonstrated that SAM is sensed by S box leader RNAs, in the absence of any additional regulatory factors, and that SAM promotes premature termination of transcription of S box genes (McDaniel *et al.*, 2003; Chapter 2). SAM binding assays demonstrated that SAM binds specifically and directly to the helix
1-4 region of S box leader RNAs. S box leader RNA is able to discriminate against closely related compounds, which is essential for an appropriate regulatory response in the cell. Compounds related to SAM were unable to act as competitive inhibitors of SAM-directed transcription termination of the \textit{ykrW} leader, although addition of sinefungin resulted in a slight increase in read-through in the presence of SAM. Testing the ability of other S box leaders to respond to sinefungin and other SAM-related compounds could give some insight into the leader elements required for discrimination. The determinants for specificity of the SAM/leader RNA interaction could also be identified by \textit{in vitro} selection for leader variants that are able to bind SAH or other potential SAM competitors. It is possible that the properties of SAM that are important for interaction with the leader RNA could be identified using SAM analogs, including \textit{S}-adenosylcysteine and azaSAM, that have not yet been tested and could potentially compete with SAM in the \textit{in vitro} assays. Studies of the kinetics of transcription of S box leaders could provide additional insight into the molecular mechanism of transcription termination, such as investigating the time window during which SAM can be added and still promote termination of transcription. It is probable that the rate of transcription is an important factor in S box gene regulation \textit{in vivo}, and pausing may be required to allow time for the nascent transcript to fold and for SAM to bind. In the \textit{in vitro} transcription assays, the rate of transcription was artificially slowed with low NTP concentrations.

Genetic and biochemical approaches were utilized to examine the importance of conserved \textit{cis}-acting elements in the helix 1-4 region of S box leader RNAs for the SAM/leader RNA interaction, SAM-induced structural changes in the leader RNA, and SAM-directed transcription termination \textit{in vivo} and \textit{in vitro} (Chapter 3). SAM binding
assays using deletion variants of the \( ykrW \) leader indicated that the entire helix 1-4 region is important for SAM binding. SAM binding assays carried out for \( yoaD \) leader RNAs indicated that the extra stem loop in \( yoaD \) is essential for SAM binding, even though this structure is absent in all other S box leaders. The extra stem loop could compensate for the shortened helix 3 in \( yoaD \). The SAM binding assays indicated that the stability of the interaction between SAM and individual S box leader RNAs is variable. Some variability was also observed in the efficiency of termination in the absence of SAM and in the degree of the response to SAM for different S box leaders in the \textit{in vitro} transcription assays (Chapter 2). These results, in combination with the variability observed in some S box leader structures by phylogenetic analysis, provide further evidence that some S box genes may respond to other regulatory signals in addition to methionine availability, as is the case for \( \textit{cysH} \), which primarily responds to \( O\)-acetyl-serine (Mansilla \textit{et al.}, 2000; Chapter 2). Further analysis of these differences observed for individual S box leaders could provide additional information about the S box leader elements required for SAM-directed transcription termination.

Mutational analysis of \textit{B. subtilis yitJ} indicated that changes in conserved elements in helices 1 and 2, including the kink-turn motif, lead to loss of repression during growth in the presence of methionine, loss of SAM binding, and loss of SAM-dependent termination \textit{in vitro}. It is not clear why the single nucleotide substitutions introduced in the kink-turn motif resulted in decreased expression in the absence of methionine, as these mutations are not predicted to affect antitermination. Our results confirmed the importance of the kink-turn motif in S box gene regulation, and we predicted that the kink-turn motif could facilitate formation of the S box leader RNA
tertiary structure by introducing a bend in the RNA. Analyzing the effect of mutations in
the kink-turn motif on SAM binding, SAM-dependent transcription termination, and
leader RNA structure could provide additional information concerning the role of the
kink-turn motif in SAM-directed transcription termination.

Covariation of nucleotides in the terminal loop of helix 2 and the unpaired region
between helices 3 and 4 was observed in all S box leaders, and it was postulated that this
covariation could indicate a tertiary interaction between these elements (Grundy and
Henkin, 1998), which could be assisted by the kink-turn motif. Analysis of *B. subtilis*
*yatJ* and *ykrW* leader variants demonstrated that interaction between these two covarying
regions is required for SAM binding, SAM-directed transcription termination, and SAM-
induced structural changes (Chapter 3). The observed SAM-induced structural changes
are consistent with the proposed secondary structure model. This study provided the first
information concerning the tertiary structure of S box leader RNAs, and our results
suggest that S box leader RNAs may form a SAM-binding pocket, similar to the purine-
binding pockets observed for the purine-sensing riboswitches (Batey *et al.*, 2003;
Serganov *et al.*, 2004; Chapter 3). More detailed structural analysis of the S box leader
RNA using NMR spectroscopy and X-ray crystallography would provide a wealth of
information about the tertiary structure of S box leader RNAs and the SAM/leader RNA
interaction. The *yatJ*-W2 and *ykrW*-J1 variants exhibited a significant defect in
repression *in vivo* under conditions where SAM pools are predicted to be high, but
exhibited little or no defect in the *in vitro* assays. It is likely that the sequence changes in
regions 1 and 2 result in reduced affinity for SAM, since both the *yatJ*-W2 and *yatJ-
W1W2 variants required a 10-fold higher SAM concentration in the *in vitro* termination
assay as compared to the wild-type yitJ template. Although the SAM concentration used in the \textit{in vitro} assays is within the physiological range reported by Wabiko \textit{et al.} (1988), the cellular pool of free SAM that is available for binding to the S box leader RNA transcripts is likely to be significantly lower. It therefore appears that expression \textit{in vivo} may be more sensitive to alterations in regions 1 and 2 than is apparent from the \textit{in vitro} assays. The possibility also remains that although the \textit{in vitro} conditions generated were adequate for SAM binding and SAM-directed transcription termination, additional factors could be involved in S box gene regulation \textit{in vivo}. These factors could control the rate of transcription, modulate the SAM/leader RNA interaction, or assist in folding of the leader RNA.

Previous mutational analysis indicated that the helix 1-4 region of S box leader RNAs was likely to be the target for binding of a negative regulatory factor during growth in the presence of methionine (Grundy and Henkin, 1998; Winkler \textit{et al.}, 2001). However, it was not known how methionine availability is sensed. The information obtained from these studies led to the development of a new model for S box gene regulation, in which SAM binds directly to the helix 1-4 region of the S box leader RNA, stabilizing the anti-antiterminator structure, resulting in premature termination of transcription at the leader region terminator (Fig. 6.1). The anti-antiterminator is destabilized when SAM is not available for binding, allowing formation of the antiterminator, which results in readthrough of the downstream coding region(s) by preventing formation of the terminator helix. This model is supported by structural analysis of \textit{in vitro} transcribed S box leader RNAs (McDaniel \textit{et al.}, 2003; Epshtein \textit{et al.}, 2003; Winkler \textit{et al.}, 2003; Chapter 3). It is possible that a single binding site exists
for SAM in the leader RNA, as a significant molar excess of SAM was not required for
SAM binding in the in vitro assay.

Figure 6.1. The proposed model for regulation of S box gene expression in response
to SAM. The antiterminator structure (AT, red-blue) forms in the absence of SAM,
allowing expression of the downstream coding region(s). Binding of SAM stabilizes the
anti-antiterminator structure (AAT), which sequesters sequences (red) required for
formation of the antiterminator and frees sequences (blue) required for formation of the
less stable terminator helix (T). The asterisk represents SAM, although it is not known
exactly where in the helix 1-4 region SAM binds. Modified from Grundy and Henkin
(2004b).
Much remains to be done to identify all of the *cis*-acting sequence and structural features of the leader RNA that are involved in the molecular mechanism of S box gene regulation. Further site-directed mutagenesis of conserved elements, guided by phylogenetic analysis, in combination with *in vivo* and *in vitro* assays, is necessary to identify which elements in the helix 1-4 region of the leader RNA are required for SAM binding, discrimination against SAM related compounds, SAM-induced structural changes, and SAM-directed transcription termination. Additional characterization could also be carried out for existing leader variants, such as *in vitro* transcription assays and *in vivo* expression assays for the *ykrW* and *yoaD* variants that were generated to determine the minimal SAM binding site. Further analysis of mutants obtained in the search for *trans*-acting mutations that affect S box gene regulation (see below; Chapter 4) should also be carried out. Some mutants exhibited a loss of repression during growth in the presence of methionine due to a mutation in the *yitJ-lacZ* fusion itself, and the leader regions from these fusions should be sequenced to provide additional information about the elements required for SAM-directed transcription termination.

We predicted that we could isolate *trans*-acting mutations that would lead to a loss of repression of S box gene expression under S box-repressing conditions, either in the negative regulatory factor itself, or in other factors that are involved in S box gene regulation *in vivo*. A *trans*-acting mutation (SBD1) was isolated in the gene that encodes SAM synthetase (*metK*). This mutation, which resulted in a 15-fold decrease in SAM synthetase activity, resulted in derepression of S box gene expression during growth in the presence of methionine, consistent with our identification of SAM as the molecular effector for S box gene expression (Chapter 4). The fact that the only *trans*-acting
mutation isolated was in \textit{metK} suggests that no factors in addition to SAM may be involved in S box gene regulation \textit{in vivo}. However, the possibility remains that other mutations were not isolated because the required factors are essential gene products. The SBD1 allele was identified as an amino acid substitution at the residue in SAM synthetase (A83) that is analogous to the only residue that participates in hydrogen bonds between the dimers of \textit{E. coli} SAM synthetase to form the active tetramer. The amino acid substitution could affect SAM synthetase activity by preventing tetramerization of the enzyme, as observed for some mutants of \textit{E. coli} SAM synthetase (Markham and Satishchandran, 1988; Reczkowski and Markham, 1995; Chapter 4). The effect of the SBD1 allele on the oligomerization state of SAM synthetase should be analyzed using native polyacrylamide gel electrophoresis. The effect of the decrease in SAM synthetase activity caused by the SBD1 allele on SAM pools should be quantitated, using HPLC analysis. The decrease in the SAM pools as compared to a wild-type strain could be relatively small, as a 25- to 200-fold decrease in SAM synthetase activity in \textit{metK} mutant strains of \textit{B. subtilis} results in only a three to four-fold lower SAM pool during growth in methionine (Wabiko \textit{et al.}, 1988).

We also analyzed the effect of increased SAM synthetase activity on S box gene expression \textit{in vivo}, to compare to the observed effect of decreased SAM synthetase activity. Increased SAM pools due to increased SAM synthetase activity resulted in delayed induction of S box gene expression in response to starvation for methionine \textit{in vivo}, which is also consistent with our identification of SAM as the molecular effector for S box gene expression (McDaniel \textit{et al.}, 2003; Chapter 4). The \textit{metK} gene in \textit{B. subtilis} is predicted to be regulated by SAM (the metabolite synthesized by its gene product) as it
is a member of the S box regulon; however, it is possible that metK expression is regulated at other levels as well.

A secondary goal of this project was investigation of the physiological roles of S box genes of unknown function in B. subtilis (Chapter 5). We investigated roles for the ykrTS and ykrWYZ gene products in MTA recycling based on the presence of the S box regulatory elements and sequence similarity to proteins of known function. Direct roles were demonstrated for the B. subtilis ykrT, ykrS and ykrW genes in the MTA recycling pathway, and ykrT was identified as the gene encoding MTR kinase. Ashida et al. (2003) confirmed roles for ykrS and ykrW in the MTA recycling pathway, identifying ykrS as the gene encoding methylthioribose-1-phosphate isomerase, and identifying ykrW as the gene encoding 2,3-diketo-5-methylthiopentyl-1-phosphate enolase. This study demonstrated that identification of genes of unknown function as members of the S box regulon can be used to predict physiological roles in sulfur metabolism (Murphy et al., 2002; Chapter 5).

MTR kinase is a target for antimicrobial agents since it is required for methionine salvage in certain prokaryotes but is absent in mammalian cells (Gianotti et al., 1990). Analogs of MTR kinase have been identified that inhibit the growth of K. pneumoniae (Gianotti et al., 1990) and act as antiprotozoal agents (Riscoe et al., 1988). Further analysis of the structure and function of MTR kinase may allow the development of this enzyme as an antimicrobial target. YkrW has also been identified as a novel target for antimicrobial drugs, as MTR becomes toxic to the cell in the absence of YkrW (Sekowska and Danchin, 2002). This is of great interest, as methionine salvage pathways have been identified in several pathogenic organisms, including Bacillus anthracis.
In summary, this work revealed details of the molecular mechanism of transcription termination control governing the S box system, and the tertiary structure of S box leader RNAs. We showed that methionine availability is sensed by direct binding of SAM to the S box leader RNA, and demonstrated the importance of conserved cis-acting elements in the helix 1-4 region. Roles in the MTA recycling pathway were demonstrated for S box-regulated genes of unknown function in *B. subtilis*, which confirmed that the identification of the S box regulatory motif can be used to predict physiological roles in sulfur metabolism.
LIST OF REFERENCES


