Mechanism of Riboswitch Regulatory Domains and Application for Developing Novel Biosensors

Pablo Ceres Borau

University of Colorado at Boulder, ceresp@hotmail.com

Follow this and additional works at: https://scholar.colorado.edu/chem_gradetds

Part of the Biochemistry Commons, and the Cell Biology Commons

Recommended Citation

https://scholar.colorado.edu/chem_gradetds/96

This Dissertation is brought to you for free and open access by Chemistry & Biochemistry at CU Scholar. It has been accepted for inclusion in Chemistry & Biochemistry Graduate Theses & Dissertations by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
Mechanism of Riboswitch Regulatory Domains and Application for Developing Novel Biosensors

by

Pablo Ceres Borau

B.S., Bethel College 1999

A thesis submitted to the Faculty of the Graduate School of the University of Colorado in partial fulfillment of the requirement for the degree of Doctor of Philosophy Department of Chemistry and Biochemistry 2013
This thesis entitled:

Mechanism of Riboswitch Regulatory Domains and Application for Developing Novel Biosensors
Written by Pablo Ceres Borau
has been approved for the Department of Chemistry and Biochemistry

______________________
Robert T. Batey

______________________
Dylan J. Taatjes

Date______________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline
Abstract

Ceres, Pablo (Ph.D., Biochemistry, University of Colorado, Boulder)
Mechanism of Riboswitch Regulatory Domains and Application for Developing Novel Biosensors

Thesis directed by Professor Robert T. Batey

Riboswitches are structural elements of gene control located in the 5’-untranslated region of some mRNAs capable of regulating the expression of important genes. The mode of regulation occurs by binding a cellular metabolite if its concentration excess a threshold and controlling genes involved in the biogenesis or transport of this ligand. A common riboswitch is composed of two functionally distinct units: i) the aptamer domain, which recognizes and binds the cellular metabolite and ii) an expression platform responsible for interaction with the cellular machinery and control of gene production. The aptamer domain of riboswitches confers high degree of affinity and specificity towards the correct ligand. This binding event rearranges the secondary structure of the expression platform, which in turn allows regulation of gene expression. Riboswitch genetic control is observed in the three domains of life. However, it is most important in bacteria.

Previously, the characterization of riboswitches has focused on the biophysical study of different aptamer domains and the changes that occur when the ligand binds the RNA. These studies have revealed the intricate recognition of essentially every functional group of a diverse set of cellular metabolites by RNAs. However, the signal transduction from the aptamer domain to the expression platform remains uncharacterized. Most importantly, the secondary structural switch that an expression platform undergoes has been reduced to a simple variation of secondary interactions. In this dissertation, I have functionally characterized, both in vitro and in vivo, two different riboswitches in their entirety, demonstrating novel findings about the regulatory response of these ribo-regulators. The results show that small secondary structures formed within the expression platform are important for gene control and that the ability of the switching mechanism is innate within this domain (independent of ligand presence).

The application of these learned principles of riboswitch regulation for the development of novel RNA-based biosensors has also been accomplished. To establish engineering strategies capable of developing general, reliable, and modular RNA devices, I mixed and matched natural and SELEX-derived (systematic evolution of ligands by exponential enrichment) aptamers with a variation of natural expression platforms. These engineered riboswitches are functional in vitro and in E. coli, and present a new avenue to control cellular regulation for different applications including synthetic biology.
Dedication

To my daughters Ellen and Grace because nothing is as good as your hugs and smiles when I arrive home after a long day. I hope to give you as much love as you have given me.

A mis hijas Ellen y Gracia porque nada es tan bueno como vuestros abrazos y sonrisas cuando llego a casa después de un largo día. Espero poder daros tanto amor como vosotras me habéis dado.
Acknowledgements

I would like to thank my mentor Professor Robert Batey Ph.D. for his support throughout the time I have spent in his laboratory. His knowledge, resourcefulness and ideas have been a tremendous help in the technical and intellectual advancement of the projects that I completed in his laboratory. I know that without his help I would not have accomplished all that I have. I would also like to acknowledge my committee members for their advice Tom Blumenthal Ph.D., Shelley Copley Ph.D., Dylan Taatjes Ph.D., and Marcelo Sousa Ph.D.

During the last several years I have taken advantage of the structural and biochemical characterization previously performed by past members of the Batey laboratory. I would like to thank them for their concise work and for showing me what can be achieved with consistency, tenacity and hard work. These people include Sunny Gilbert Ph.D., Rebecca Montange Ph.D., Colby Stoddar Ph.D., Andrea Edwards Ph.D., Francis Reyes Ph.D. and James Johnson Ph.D. I specially would like to mention Andrew Garst Ph.D. for all our discussions about RNA structure and function, riboswitches and everything else that came to mind. Three current members of the Batey laboratory have contributed to the work presented in this dissertation: Joan Marcano, who assisted during in vivo characterization of two chimeric riboswitches, Jeremy Trausch, who contributed during the development of an additional modular expression platform from a riboswitch, and John Zinder, who helped with the in vivo characterization of the metE riboswitch.
I would also like to acknowledge my family because graduate school can be a very demanding time. The sacrifices that my wife, Anne, and my children, Ellen and Grace, have made were certainly appreciated. My parents, brothers and sisters were very supportive throughout this time and finally my in-laws who were so caring during difficult times.
Disclaimer

The work contained in this Dissertation has been in part previously published in the following works:


Table of contents

Acknowledgements.....................................................................................................................................v

Table of Contents......................................................................................................................................viii

List of Figures...........................................................................................................................................xi

List of Tables............................................................................................................................................xiii

Chapter 1. Introduction to riboswitches..............................................................................................................1
  1.1. Non-coding RNA regulation of cellular processes.................................................................2
  1.2. Riboswitches as genetic control units.........................................................................................4
  1.3. Applications of RNA based cellular regulation...........................................................................8
  1.4. Study of riboswitches may become beneficial for different applications...............................13
      1.4.1. Structural analysis of a SAM recognizing aptamer..........................................................13
      1.4.2. The metE sequence conforms to the general principles of action of the SAM-I riboswitch family.................................................................16
  1.5. Summary of thesis............................................................................................................................17

Chapter 2. Revealing the relationship between structure and regulatory activity of a SAM-responsive riboswitch..................................................................................................................20
  2.1. Introduction.....................................................................................................................................20
  2.2. Results............................................................................................................................................24
      2.2.1. Choice of riboswitch containing common features of an expression platform: metE..................................................................................26
      2.2.2. Single turnover transcription assay was employed to assess functionality........................28
      2.2.3. The metE riboswitch is functional during single round in vitro transcription assays........32
      2.2.4. Formation of a tertiary interaction is responsible for organization of the aptamer domain.........................................................................39
          2.2.4.1. Length of P2a is important for regulation...............................................................40
          2.2.4.2. Sequence composition of the kink-turn of the metE riboswitch shows deficient activity.................................................................43
          2.2.4.3. Adjacent residues to the kink-turn are also important for functionality......................44
      2.2.5. Functionality of riboswitch is controlled by the co-transcriptional folding of the P1 helix.................................................................46
      2.2.6. Validation of computationally predicted P1a and P1b.....................................................48
          2.2.6.1. Role of P1a and P1b..............................................................................................49
          2.2.6.2. Stabilization of P5 can be achieved anti-termination..............................................55
      2.2.7. In vivo characterization of the metE riboswitch.................................................................56
2.2.7.1. P1b helix controls gene expression of metE in vivo..57
2.2.8. Other riboswitch regulates gene expression similarly........60
2.3. Discussion.................................................................63
2.4. Materials and Methods................................................66
   2.4.1. Preparation of metE sequence variations......................66
   2.4.2. In vitro transcription assay.....................................67
   2.4.3. Cell based assay..................................................67

Chapter 3. Modularity of select riboswitch expression platform enables facile engineering of novel genetic regulatory devices...................................................69
   3.1. Introduction to RNA biosensors....................................69
   3.2. Results and Discussion..............................................73
      3.2.1. Considerations in design of modular expression platforms..73
      3.2.2. Chimeric riboswitches regulate transcription in vitro...80
      3.2.3. Synthetic aptamers can control the regulatory switch.....82
      3.2.4. Chimeric riboswitches function in vivo.....................84
      3.2.5. Optimizing riboswitch performance by altering P1 helix stability...............................................85
   3.3. Conclusion.....................................................................92
   3.4. Materials and Methods...............................................93
      3.4.1. Synthesis of DNA templates for transcription..............93
      3.4.2. In vitro transcription assays...................................93
      3.4.3. In vivo assay......................................................94

Chapter 4. Engineering of inducible modular regulatory RNA switches using biological components........................................97
   4.1. Introduction.....................................................................97
   4.2. Results..........................................................................101
      4.2.1. Design of chimeric riboswitches containing the pbuE expression platform..............................................101
      4.2.2. pbuE can be biologically engineered and yields functioning chimeric switches in vitro..................................................102
      4.2.3. pbuE engineered RNAs response can be tuned...........104
      4.2.4. Another approach guided by the SAH riboswitch........108
      4.2.5. In vivo functional chimeras.......................................111
      4.2.6. pbuE expression platform can be directed by a synthetic aptamer.................................................................120
      4.2.7. Engineering of chimeric ribo-regulators using the SAH riboswitch......................................................120
         4.2.7.1. Functionality of chimeric riboswitches created with the metH expression platform.................................126
      4.2.8. In vivo functional chimeras created with the expression platform from pbuE and metH.................................128
   4.3. Discussion.....................................................................133
   4.4. Materials and methods................................................136
      4.4.1. DNA template construction.......................................136
4.4.2. *In vitro* transcription assays..................................................137
4.4.3. *In vivo* assay............................................................................137

Chapter 5. Future directions. .................................................................139
  5.1. Summary of thesis. ........................................................................139
  5.2. Future directions. ..........................................................................145
    5.2.1. Biological implications of riboswitch based genetic regulation........................................145
    5.2.2. Possible applications for chimeric riboswitch control systems........................................146
    5.2.3. Possible application of the novel bio-sensory devices............147

References........................................................................................................151

Appendix........................................................................................................164

  Appendix 1.
  Sequences of metE and mutations used for functional characterization of the riboswitch.........................164

  Appendix 2.
  Sequences of wild type riboswitches unresponsive during *in vitro* transcription assay..........................169

  Appendix 3.
  Sequences of chimeric riboswitches utilized to demonstrate the modularity of aptamer domains and expression platforms....................................................170

  Appendix 4
  Sequences of chimeric riboswitches that allow gene expression when bound to their cognate ligand........177
List of Figures

Chapter 1.

Figure 1.1. Cis- and trans-acting RNAs control the fate of mRNAs ................. 4
Figure 1.2. Mechanism of CRISPR adaptive immunity .................................. 6
Figure 1.3. Mechanism of a transcription controlling riboswitch ...................... 8
Figure 1.4. Secondary and X-ray crystal structures of the SAM-I riboswitch ....... 11

Chapter 2.

Figure 2.1. Sequence and secondary structure of the metE riboswitch .......... 26
Figure 2.2. Representation of data obtained during a typical in vitro transcription experiment .......................................................... 30
Figure 2.3. The metE riboswitch responds to SAM during a single round in vitro transcription .......................................................... 33
Figure 2.4. Regulation of metE riboswitch mutant U157A .............................. 35
Figure 2.5. Sequence adjacent to the binding pocket assists during binding .... 36
Figure 2.6. Functionality of mutation A85G in metE ................................. 37
Figure 2.7. The peripheral extension of P3 is important for function ............... 39
Figure 2.8. Secondary structure of the metE riboswitch highlighting the regions subjected to mutational analysis ................................................. 40
Figure 2.9. Secondary structure of metE identifying the mutations performed in P2 and P4 .......................................................... 41
Figure 2.10. The kink-turn composition is relevant for regulation ................. 44
Figure 2.11. Helical region P1 of metE controls the transcriptional switch ....... 48
Figure 2.12. Sequence and secondary structure of the expression platform of metE illustrating the mutational strategy ................................................. 48
Figure 2.13. Single round in vitro transcription data of the metE riboswitch and mutants ....................................................... 51
Figure 2.14. Mutation strategy of metE riboswitch to study the formation of the helix P5 .......................................................... 54
Figure 2.15. Map of the parental riboswitch reporter plasmid (pRR1) ............ 58
Figure 2.16. In vivo genetic control of the metE riboswitch .......................... 59
Figure 2.17. Sequence and secondary structure of the pbuE riboswitch .......... 61
Figure 2.18. Graphical representation of the response of the pbuE riboswitch and the alterations of the transcription start site ................................................. 64

Chapter 3.

Figure 3.1. Riboswitches regulate transcription through a ligand-dependent secondary structural switch .......................................................... 72
Figure 3.2. A chimeric xpt/metE riboswitch regulates transcriptional termination .......................................................... 77
Chapter 4.

Figure 4.1. In vitro characterization of the pbuE riboswitch..........................104
Figure 4.2. Sequence and secondary structure of the pbuE riboswitch and the engineering design of pbuE/pbuE.................................................................107
Figure 4.3. In vitro characterization of the chimeric riboswitch strategy to couple aptamers to the pbuE expression platform ..............................................110
Figure 4.4. In vitro characterization of P1 mutants of pbuE expression platform..........................................................................................113
Figure 4.5. The response of yitJ/pbuE chimeric riboswitch can be tuned by mutating the poly-uridine region...............................................................116
Figure 4.6. Different aptamers can direct the expression platform of pbuE.....118
Figure 4.7. SELEX-derived aptamers are capable of directing gene expression when fused to pbuE.................................................................121
Figure 4.8. Sequence and secondary structure of the metH riboswitch from D. aromatica and the new engineering design for allowing gene expression control ..........................................................................................122
Figure 4.9. Engineering of the sahH/metH chimeric riboswitch..................124
Figure 4.10. Design of the chimeric riboswitch xpt/metH and engineered mutations enhancing the switching mechanism.............................................126
Figure 4.11. In vitro characterization of chimeric riboswitches constructed with the expression platform of the metH riboswitch..........................127
Figure 4.12. In vivo characterization of the pbuE riboswitch and the functional chimeras that allow gene expression......................................................129
List of Tables

Table 2.1. Data quantification of A85 mutations in the metE riboswitch..................37
Table 2.2. Data quantification of the metE riboswitch and mutants located in P2 ...............................................................42
Table 2.3. Data quantification of the metE riboswitch and mutations within P2 and P4..........................................................................................................................................................................................45
Table 2.4. Truncations of metE varying the transcription start site change the regulation of the riboswitch..........................................................................................................................................................................................50
Table 2.5. Quantitation of metE riboswitch and mutations during single round in vitro transcription........................................................................................................................................................................................................53
Table 2.6. Data obtained in titrations of metE and mutations located in the P5 region........................................................................................................................................................................................................56

Table 3.1. Quantitation of in vitro transcription data and isothermal titration calorimetry........................................................................................................................................................................................................96

Table 4.1. in vitro data characterization of initial trials for creating chimeric riboswitches using the pbuE expression platform.................................................................111
Table 4.2. Data quantification of yitJ/pbuE and mutants located in the poly-uridine tract........................................................................................................................................................................................................115
Table 4.3. xpt (C74U)/pbuE and pbuE/pbuE chimeric riboswitches response can be tuned by mutating the poly-uridine tract in the expression platform.............117
Table 4.4. Data quantification of in vitro titrations of chimeric riboswitches employing the expression platform of pbuE.......................................................................................119
Table 4.5. Data quantification of chimeric riboswitches created by fusing aptamers to the metH expression platform.........................................................................................128
Table 4.6. In vivo titration data of pbuE mutants and chimeric riboswitches.....131
Chapter 1

Introduction to RNA-based biology

The adaptation of an organism to the constantly changing environment is crucial for its survival. To adjust to these variations a cell must undergo physiological changes, accomplished through alterations in gene expression. This task was originally thought to pertain to proteins only; however, it can also be performed by nucleic acid. In recent decades, numerous RNAs have been discovered and characterized that control or regulate important cellular processes (Gripenland et al., 2010; Storz et al., 2011). A diverse set of functions has been attributed to non-coding RNAs (ncRNAs), and non-protein coding elements of RNA, ranging from sensing temperature, pH and metabolite fluctuation to transcriptional interference, translation inhibition, alternative splicing, or even mRNA stability (Grundy and Henkin, 2006).

Non-coding RNAs (ncRNAs) have been discovered in bacteria and biochemically characterized yielding a myriad of functions and mechanisms of action (Gripenland et al., 2010). Some of the better-characterized RNAs are cis-acting antisense RNAs, which act within the same RNA they are a part of. Another example of well-studied regulation is that of trans-acting small ncRNAs, which may influence a transcript distant from their origin. Both cis- and trans-acting RNAs can modulate gene expression. Additionally, Clustered Regularly Interspace Short Palindromic Repeats (CRISPR) represent an RNA-associated defense mechanism for bacterial cells against bacteriophages and horizontal gene transfer (Barrangou et al., 2007; Deveau et al., 2010). Although there are still many questions to be answered about the regulatory
mechanisms of these RNAs, the trend in the expansion of the functions performed by RNA should be noted.

In this chapter, a brief introduction of some of the most important ncRNAs will be given, which will be followed by a more in-depth description of riboswitches as they are the focus of this dissertation. After that, I will illustrate recent advancements of RNA-based engineered devices to control different genetic cellular outcomes. To this end, I have employed components derived from natural riboswitches to engineer synthetic biosensors capable of gene regulation in bacteria.

1.1. Non-coding RNA regulation of cellular processes

RNA-based gene regulatory mechanisms are conserved across the bacterial kingdom and possibly outnumber proteins in quantity and diversity (Barrick and Breaker, 2007; Storz et al., 2011). One common feature of these RNAs is that they may be unable to function alone and require the assistance of chaperones and/or other protein factors such as Hfq or RNases, which may stabilize these ncRNAs (Beisel and Storz, 2010). Generally, these RNAs interact with their targets via canonical Watson-Crick base pairing. Thus the mode of initial recognition is difficult to achieve since there are many potential targets in the cell (Gripenland et al., 2010). To overcome this problem, a fast and high-affinity interaction is initially achieved by the recognition of exposed nucleotides usually presented as a stable local fold of the target mRNA, the ncRNA or in some cases both (Storz et al., 2011). For example, a C-rich region contained within loops mediates the interaction of small RNAs (sRNAs) in *Staphylococcus aureus*. After the initial contact, a rearrangement of the RNA can occur and results in more base pair
formation, suggesting that a multi-step interaction may be required (Brantl, 2007). Fully describing the function and mechanism of all these ncRNAs is beyond the scope of this chapter. However, some of the better-characterized ncRNAs will be introduced.

Trans-acting RNAs, which can interact with targets distant from their origin, control important cellular processes in bacteria. These ncRNAs may bind either an mRNA or a protein, which drastically change functionality. When bound to an mRNA, the complex may be targeted for degradation, leading to down regulation of the associated gene (Gripenland et al., 2010) (Figure 1.1). This interaction can take place near the ribosome binding site (RBS) or Shine-Dalgarno sequence and block the normal loading of ribosomes. However, recent studies have shown that these interfering interactions can occur 70 nucleotides upstream, or 15 nucleotides downstream, of the start codon in both *E. coli* and *Salmonella* (Bouvier et al., 2008). Also variable are the sites of interactions of other ncRNAs that activate translation of mRNA by preventing the formation of an inhibitory secondary structure (Frohlich and Vogel, 2009). Trans-acting ncRNAs may also interact within the coding region of some mRNAs, thereby regulating stability (Pfeiffer et al., 2009). These ncRNAs may influence the activity of mRNAs by directly establishing a complex.

Trans-acting ncRNAs may also achieve a variation in genetic regulation by affecting the function of a target protein after a complex has formed (Waters and Storz, 2009). Some of these interactions sequester proteins from their regular target such as RNA binding proteins. This sequestering process may be performed by the ncRNA because it may contain multiple copies of the protein recognition site. The specific sequence may be repeated 18 times as in the case of *E. coli* CsrB RNA for CsrA, which
is a protein known to control mRNA stability and translation (Babitzke and Romeo, 2007). More importantly CsrB-like RNAs and CsrA-like proteins have been identified in other species such as *Pseudomonas*, signifying that this mode of regulation may be widely spread (Sonnleitner et al., 2009). Together, these examples demonstrate the importance of *trans*-acting ncRNAs in bacterial gene regulation.

![Diagram showing RNA-based regulation](Image)

**Figure 1.1.** *Cis-* and *trans*-acting RNAs control the fate of mRNAs. *Cis*-acting ncRNAs are synthesized from the complementary strand of the DNA and can bind the 3' UTR of the mRNA, while *trans*-acting ncRNAs can vary in location origin and may bind in the ribosome binding site or Shine-Dalgarno (SD) sequence. RNases can then degrade the mRNA. This figure was taken from (Gripenland et al., 2010).

RNA-based regulation of cellular processes can also be performed by *cis*-acting ncRNAs. These RNAs can act by inhibiting the function of their target mRNAs. The origin of transcription of these ncRNAs may be a promoter located in the complementary strand of the DNA. The mode of function of these RNAs, as the previously described *trans*-acting, is established by annealing to their target on or near the ribosome binding site (RBS), and inhibiting the ribosomes from normally translating...
the mRNA (Figure 1.1). The size and location of interaction can greatly vary. Some of these RNAs may function by overlapping with more than one gene as their length may reach a few kilobases in size, which has been demonstrated in *Listeria monocytogenes* and *Bacillus subtilis* (*B. subtilis*) (Rasmussen et al., 2009; Toledo-Arana et al., 2009). Besides the importance of their length, these RNAs are also important due to their persistent transcription as occurs in *Escherichia coli* (*E. coli*), where there have been 1000 transcription start sites identified for antisense RNAs (Dornenburg et al., 2010). This is also the case in eukaryotic organisms where the presence of these RNAs has now been identified (Jacquier, 2009). The characterization of these *cis*-acting RNAs deserves attention, as they may be more important in number and function than originally hypothesized.

One more example of a crucial cellular process performed by a ncRNA occurs in the regulation of CRISPR in bacteria. This complex represents an adaptive immunity system to resist bacteriophage infection and horizontal gene transfer by bacteria. This system is present in 50% of bacteria and 85% of archaea whose genomes have been sequenced (Grissa et al., 2007). In general, this system contains a short sequence of the non-native DNA that becomes incorporated into the bacterial CRISPR sequence (Figure 1.2 A and B). The immune response occurs when these sequences are transcribed and processed by the CRISPR associated (Cas) proteins, forming a ribonucleoprotein complex. The RNA component of this complex targets a complementary sequence in the foreign DNA, which is cleaved and therefore its expression is stopped (Figure 1.2 C). This system was reviewed in (Richter et al., 2012).
The functional RNAs, briefly described above, perform a diverse set of biological processes and can be found in the three domains of life. This prevalence signifies that RNA may present an evolutionary advantage for regulating these processes over other functional macromolecules, such as proteins. Some of the factors that may favor RNA to develop these characteristics may be a low metabolic cost (the translation of a complete mRNA is not required), a greater number of regulation levels (transcription, translation, splicing, etc.), faster control, and the precise interaction and location that an

**Figure 1.2. Mechanism of CRISPR adaptive immunity.** A) During the invasion of foreign DNA (phages or plasmids), small sequences of these elements are incorporated into the leader end of the CRISPR sequence and the repeat duplicated. B) CRISPR RNAs are transcribed into pre-crRNA and then processed into mature crRNA. C) Complex formation of the crRNA and the Cas proteins that binds the target RNA and targets it for degradation. Figure was taken from (Richter et al., 2012).
RNA can form through base-pairing (Beisel and Storz, 2010). All these illustrations of the varied RNA-based functions signify that RNA is more than a carrier of genetic information. However, the focus of this dissertation is the genetic control exhibited by riboswitches, which will be introduced in the next section.

1.2. Riboswitches as genetic control units

Riboswitches are regulatory elements located in the leader sequence of some mRNAs. They are composed of two domains of distinct function. The aptamer domain acts as a receptor and is capable of binding the effector molecule, and the expression platform, which communicates the binding event in the aptamer domain to the cell machinery controlling gene expression (Figure 1.3) (Roth and Breaker, 2009). The aptamer domain is a highly conserved region within the same riboswitch family, which confers high specificity and affinity for the cellular metabolite. These ribo-regulators are capable of sensing a wide assortment of cellular metabolites such as ions (Baker et al., 2012; Cromie et al., 2006; Dann et al., 2007), metal cofactors (Regulski et al., 2008), purines (Batey et al., 2004; Mandal and Breaker, 2004), co-enzymes (Nahvi et al., 2002; Wang et al., 2008; Winkler et al., 2003), sugars (Roth and Breaker, 2009), cellular secondary messengers (Lee et al., 2010), and amino acids (Roth and Breaker, 2009).

The expression platform of a riboswitch is highly variable in sequence composition as well as in the mode in which gene regulation is performed (Barrick and Breaker, 2007). Transcriptional riboswitches (the focus of this dissertation) control the formation of an intrinsic rho-independent terminator, translational controllers regulate access to the ribosome binding site (RBS), alternative splice site modulators direct
splicing variants in eukaryotes, and finally these RNAs may alter their own stability by inducing self-cleaving events in a riboswitch-ribozyme combination (Serganov and Nudler, 2013; Serganov and Patel, 2012a, b).

**Figure 1.3: Mechanism of a transcription controlling riboswitch.** Cartoon representation of the control mechanism performed by a transcription controlling riboswitch. On the left, the aptamer domain (AD) is not folded into its highly structured arrangement in the absence of ligand, which corresponds to the formation of an anti-terminator in the expression platform (EP). This secondary structure of the RNA allows normal transcription of the gene. However, if the concentration of a ligand reaches a threshold the riboswitch binds the metabolite. The binding event correlates with a significant conformational change in the structure of the aptamer forming a helical region called P1, which incorporates the switching sequence (represented in red) in its 3’-side. This bound aptamer domain rearranges the secondary structure of the expression platform forming a rho-independent terminator, which is recognized by the RNA polymerase. The enzyme will stop transcription at this site.

Riboswitches were discovered a decade ago by the Breaker group (Winkler et al., 2002a; Winkler et al., 2002b), but were originally investigated before by other laboratories (Franklund and Kadner, 1997; Gelfand et al., 1999; Grundy and Henkin, 1998; Miranda-Rios et al., 2001; Mironov et al., 2002; Nou and Kadner, 1998). The Kadner group convincingly reported a change in genetic outcome due to the presence or absence of a cellular metabolite (Nou and Kadner, 2000). The authors identified that the *btuB* mRNA, which encodes a cobalamin transporter in *E. coli*, changed its gene
expression depending on the presence or absence of adenosylcobalamin, and not cyanocobalamin (Nou and Kadner, 2000). Soon after that, and with the help of bioinformatics, the identification of highly conserved sequences previously shown to bind thiamine pyrophosphate in bacteria was illustrated in eukaryotic mRNAs (Sudarsan et al., 2003a; Winkler et al., 2002a). The discovery of the “S-box” is of special interest for this dissertation, since in chapter 2 an S-adenosylmethionine (SAM)-binding riboswitch will be functionally characterized (Grundy and Henkin, 1998).

These ribo-regulators present a level of gene control that is mediated in the absence of protein cofactors and consists solely of an interaction occurring between an mRNA and a cellular metabolite. This is a factor of relevance because it allows for the quick and simple \textit{in vitro} characterization of these RNAs and the effect their cognate ligands have on gene regulation. Auxiliary proteins, such as NusA or Hfq are not required for function; however, they may assist in genetic regulation of a riboswitch (Nudler and Gottesman, 2002; Nudler and Gusarov, 2003).

The genes controlled by these RNA sequences are usually important for the metabolism, biosynthesis or transport of the cognate metabolites. Riboswitches are important not only because they can control up to 4% of the total genes in some Firmicutes such as \textit{B. subtilis}, but also because they control genes located virtually everywhere in the genome, and belonging to different operons (Mandal et al., 2003). For transcriptional regulators the binding of the metabolite arranges the aptamer in a highly structured form that stabilizes the formation of the helical region P1 (Figure 1.3). This re-arrangement is in turn communicated to the expression platform or functional domain and the gene control takes place by the formation of two mutually exclusive
structures: i) the rho-independent terminator, which induces the stalling of the polymerase and in turn sudden stop of transcription or ii) the formation of the anti-terminator, which allows the gene to be fully transcribed (Figure 1.3) (Serganov and Patel, 2012a).

In the Batey laboratory, previous efforts have been focused on the structural and biochemical characterization of riboswitches in general and particularly the aptamer domain. Some of the high resolution structures solved to date are the aptamer domains of the purine riboswitch (Batey et al., 2004), S-adenosylmethionine (SAM) illustrated in Figure 1.4 (Gilbert et al., 2008; Montange and Batey, 2006), lysine (Garst et al., 2008), 2' deoxiguanosine (Edwards and Batey, 2009), S-adenosylhomocysteine (SAH) (Edwards et al., 2010), tetrahydrofolate (THF) (Trausch et al., 2011) and most recently coenzyme B12 (Johnson et al., 2012). In these structures, the intricate patterns of recognition of the RNA for the ligand can be observed. Virtually all functional groups of the ligand are recognized by the aptamer forming electrostatic, hydrogen-bonds, van der Waals, and other interactions.

The biochemical studies of these riboswitches have resulted in a greater understanding of the folding (Heppell and Lafontaine, 2008; Lemay et al., 2006), thermodynamic (Gilbert et al., 2006b) and kinetic (Wickiser et al., 2005a) characteristics induced by the binding event. In general, natural aptamer domains from riboswitches rely heavily on the pre-organization of the binding pocket. This arrangement allows accommodating the cognate ligand and enabling the discrimination of similar cellular metabolites. This pre-formed binding pocket is probably assisted by the tertiary
interactions organizing the helical regions in a specific conformation, and by mono and
divalent cations present in the cell.

Figure 1.4. Secondary and X-ray crystal structures of the SAM-I riboswitch. Highly conserved residues (97%) are represented in capital letters on the left. The regions boxed in different colors represent conserved structural motifs. These color choices are also illustrated on the x-ray crystal structure on the right. Figure taken from (Batey, 2011)

1.3. Applications of RNA-based cellular regulation

The study and characterization of different biologically relevant RNAs has revealed the plasticity in structure and function that RNA can achieve. Demonstrating the mechanistic details of the function of different ncRNAs is important to advance our knowledge in biological processes, but also to apply these principles to yield specific desired functions in a cell. Many RNA-based devices have been and continue to be developed for different applications in synthetic biology (Babiskin and Smolke, 2011; Beisel et al., 2008; Chang et al., 2012; Weigand and Suess, 2007; Weigand et al., 2012; Wittmann and Suess, 2012). Some of these functional engineered RNAs mimic to
some degree a biologically important RNA, and take advantage of the ability of nucleic acid to adopt different conformations of similar stability, but with different function.

Smolke and co-workers illustrated a great screening strategy of an engineered RNA capable of sensing the product of an enzyme and coupling it to the expression of green fluorescence protein (GFP) in vivo (Michener and Smolke, 2012). The mutations engineered in the enzyme to select for increased functionality were screened in a high throughput manner by monitoring the expression of GFP obtained. This method allows quick assessment of the effect of mutations on the functionality. The evolved enzyme was capable of increasing activity by 33-fold, and specificity by 22-fold suggesting that RNA can be engineered to function as a selection tool for protein evolution. This is just one of the many examples of RNA-based regulators that this group has successfully developed.

However, RNA may perform even more unexpected functions, such as a small chromophore-binding RNA obtained through SELEX (systematic evolution of ligands by exponential enrichment) (Ellington and Szostak, 1990; Tuerk and Gold, 1990), which behaves similarly to green fluorescent protein. These chromophores have a weak fluorescence when free in solution due to their flexibility. Upon binding the aptamer, however, the small molecule is stabilized in such a way that it increases the quantum yield and the fluorescence emission by approximately 20-fold (Paige et al., 2011). This RNA functions in vitro and more importantly in vivo, and the fluorescence enhancement is not activated by other cellular metabolites (Paige et al., 2011). This aptamer, termed Spinach, can be coupled to other RNA aptamers, such as a natural SAM responsive aptamer. This new coupled sequence can fluoresce upon the binding of both
metabolites (SAM and the chromophore) (Paige et al., 2012). Under this arrangement, the Spinach aptamer can be employed to monitor the cellular concentration of SAM through the different states of cellular growth. This application of an engineered RNA illustrates concisely the possibilities achievable with reliable and composable modules of RNA.

Although no real medicinal or anti-bacterial application has resulted from the study of riboswitches to date, these ribo-regulators remain a great target for novel antibiotic development. Indeed, riboswitches are well spread in bacteria and control the expression of relevant genes. However, different factors have inhibited the development of new antibacterial therapies targeting riboswitches. For example, riboswitches control gene expression by different methods (transcription, translation, splicing or stability) preventing a general strategy applicable to all these RNAs. Importantly, not all genes controlled by these ribo-regulators are essential for cell survival. It should be noted that the Lafontaine and Breaker groups have demonstrated repression of several genes in different organisms by rationally designing compounds mimicking the natural ligand that targeted specific interactions with aptamer domains of riboswitches (Blount and Breaker, 2006; Blount et al., 2007; Kim et al., 2009; Mulhbacher et al., 2010). In the case of the flavin mononucleotide, a natural compound, roseoflavin, can act as an antibacterial as it can bind the aptamer domain with similar genetic effect as the natural ligand (Lee et al., 2009).

A plausible new therapy may emerge in the future from the engineering of CRISPR. Two different studies published this year demonstrated that engineering a human CRISPR, which can add a desired sequence into the genome, may be possible
(Cong et al., 2013; Mali et al., 2013). This new development may allow for genome editing and may be of great influence in the designing of a reliable and specific methodology for gene therapy, which has been referred to as the “Holy Grail” of human medicine (van der Oost, 2013). The possibility of identifying gene mutations responsible for a number of human diseases has greatly increased with new sequencing methods. On the contrary, the treatment therapy of these genetic variations to a non-phenotypic mutant is underutilized. The possibilities demonstrated by the two different groups previously mentioned could greatly improve gene therapy development and in turn the treatment of human disease.

As it can be observed, the field of synthetic biology and medicine can greatly benefit by incorporating RNA-based regulatory devices into their repertoire of tools. In synthetic biology the main goal is to produce a relatively difficult or expensive product by developing new cellular pathways (Keasling, 2010). In medicine, the production of new antibiotic targets that may not elicit a resistance response, and the possibility of developing new strategies for gene therapy are important issues worth investigating. Although RNA research is becoming more important, it is crucial that these new strategies are fully exploited and studied.

1.4. Study of riboswitches may become beneficial for different applications

This dissertation focuses on the study of riboswitches. The specific goal is to fully demonstrate the functionally important aspects of these RNA biosensors. Rather than biochemically characterizing an aptamer domain, the results obtained relate to the communication between the aptamer domain and expression platform during gene
regulation. One of the better-characterized aptamers to date is the SAM-I family. Since this family of riboswitches will be the focus of one of the chapters in this dissertation an introduction is needed to guide the reader.

1.4.1. Structural and binding analysis of a SAM recognizing aptamer

Much of what has been demonstrated about this riboswitch focuses on the structural and binding characteristics of the aptamer to the ligand (Grundy and Henkin, 1998; Heppell and Lafontaine, 2008; Lu et al., 2010; McDaniel et al., 2005). In the Batey laboratory the crystal structure of the aptamer was solved to 2.9 Å resolution and later improved to 2.4 Å (Figure 1.4) (Montange and Batey, 2006; Montange et al., 2010). The sequence used for this study is used by the organism Thermoanaerobacter tengcongensis to control sulfur metabolism, and contains all the major and most conserved structural motifs of this aptamer (Grundy and Henkin, 1998; Winkler et al., 2003). The riboswitch that will be functionally characterized in chapter 2 constitutes the metE riboswitch from B. subtilis and differs only slightly from the solved sequence, and therefore a comparison can be made. The RNA is able to bind the cognate ligand with high affinity (approximately 130 nM) and specificity, which allows discrimination between SAM and SAH by a factor of 100 (Lim et al., 2006; Winkler et al., 2003). The secondary structure reveals four helical regions, which are connected through a four-way junction (Figure 1.4). The structural arrangements of the aptamer in the bound and free confirmation reveal two sets of coaxially stacked helical regions (Lu et al., 2010; Montange and Batey, 2006). The first coaxial stack occurs between paired regions 1 and 4, and the second occurs by stacking P2 and P3. P2 presents a highly conserved kink-turn, of variable composition, which allows the formation of a pseudo
knot between the terminal loop of helix 2 and a bulged region located in P4 (Montange and Batey, 2006). This interaction may be responsible for the pre-organization of the binding pocket and the overall fold of the aptamer, as its formation is independent of ligand binding (Stoddard et al., 2010). The overall fold, tertiary contacts and ligand interactions observed in this structure were shown in another crystal structure of a different riboswitch, the *yitJ* aptamer from *B. subtilis*, which belongs to the same family (Lu et al., 2010).

The ligand is virtually recognized in its entirety, as almost every functional group forms interactions with the aptamer. These contacts are important for the function of the riboswitch and will be studied in chapter 2. Although there are slight sequence variations between the consensus SAM-I aptamer and *metE*, the nucleotides involved in ligand binding remain unchanged. The adenosyl moiety of SAM forms a base triple with an internal loop located within P3 and which sequence is 5’-AA/U. These residues are 100% conserved among the SAM-I riboswitch family (Montange and Batey (2006)). Besides these contacts, the ligand also interacts with the minor groove of P1 through the positively charged sulfur atom, as well as the amino moiety. These interactions illustrate the intricate recognition of the ligand by the aptamer.

The crystal structure of the aptamer is an information rich coordinate map of the residues forming the aptamer domain, however the understanding of the regulation mechanism lags behind. Careful study of this structure hints towards a rational functionality of the entire riboswitch that needs to be further tested. Ligand binding by the aptamer is responsible for the stabilization of P1, and as a consequence the disruption of the anti-terminator structure in the expression platform occurs (Figure 1.3).
This conformation directs the configuration of the terminator hairpin, which is followed by a poly-uridine sequence. This element pauses transcription by the RNA polymerase, leading to its disengagement from the DNA template halting RNA synthesis. The hypothesis postulating P1 as one of the major control mechanisms for riboswitch function needs to be tested in a functional assay. Other factors that may influence the relation between structure and function of this riboswitch (particularly of the expression platform) also need to be studied.

1.5. Summary of thesis

In this thesis, I will show that I have further advanced our understanding of the control mechanism achieved by riboswitches. These new findings are of different nature than those acquired to date since they are focused on the functionality of the entire sequence, and not only on the aptamer domain in isolation. This is an important variation from previous studies, and in our opinion the difference is crucial for properly illustrating the mode of regulation by these ribo-regulators.

Chapter 2 will focus on the characterization of a SAM-responsive riboswitch, \textit{metE}, which belongs to the SAM-I family of riboswitches. This choice was originated from previous data illustrating the ability of this riboswitch to control gene expression as a function of the presence or absence of its natural ligand. The functional characterization is performed using the entire sequence and allowing for the natural folding mechanism to take place, as RNAs fold during their synthesis. This methodology allows verifying the secondary structural prediction obtained when simulating the co-transcriptional folding of this sequence \textit{in silico} with Kinefold (Xayaphoummine et al., 2005). The data obtained suggest that both the aptamer and
expression platform of *metE* have evolved to establish a narrow dynamic range of gene regulation.

The focus of chapter 3 is to employ the deep understanding of riboswitch function gained in the previous chapter to create novel ribo-regulators capable of sensing small molecules and functioning as gene control units. The data presented in this chapter show that expression platforms from different riboswitches function in an intrinsic co-transcriptional manner. This property allows the secondary structural re-arrangement to occur when the native aptamer domain is replaced by different natural or SELEX-derived sequences. In addition to this “mix-and-match” strategy, the ability to tune the response of these biosensors is developed by altering the stability of one of the major control regions of the expression platform. This is the first demonstration that components obtained from riboswitches, aptamers and expression platforms, are modular and capable of independent function.

The fourth chapter presents the work performed with two different expression platforms from riboswitches that, unlike those mentioned in the previous sections, induce gene expression when bound to their cognate ligand or analogs. In this case, the engineering process to convert two different expression platforms into independently functional modules required different strategies than those used in the previous chapter. This is due to the different structural arrangements present in these two expression platforms. These modular sequences are capable of hosting a variety of aptamers to control gene expression *in vitro* and in *E. coli*.

Chapter 5 considers the biological importance of these novel findings about the mechanistic controls of riboswitch function. The possible future directions that may be
beneficial for the further characterization of riboswitches are also discussed. In addition, I also examine the possible real-life applications of these biosensors created in the Batey laboratory (and ones that could be built by others) to change the cellular circuits.
Chapter 2

Revealing the relationship between structure and regulatory activity of a SAM-responsive riboswitch

The emphasis of riboswitch characterization has focused on the aptamer domain, and considering the expression platform as a simple secondary structural switch. However, a few recent studies demonstrated an effect on activity of different riboswitches when the expression platform was altered (Blouin et al., 2011; Lemay et al., 2011; Sherman et al., 2012). This is the reason that this chapter is not only focused on the aptamer domain, but also on the expression platform of a SAM-sensing riboswitch from \textit{B. subtilis}, \textit{metE}. The function of different nucleotides or RNA structural motifs important for riboswitch genetic control was elucidated via mutational analysis and the effect monitored using a single round \textit{in vitro} transcription assay. To date several riboswitches that bind different ligands have been functionally characterized with the help of this method (Blouin and Lafontaine, 2007; Dann et al., 2007; Lemay et al., 2011; Wickiser et al., 2005b). This chapter presents a comprehensive biochemical survey of sequence and structural elements that were hypothesized to be important for genetic regulation of the \textit{metE} riboswitch, a member of the SAM-I family.

2.1 Introduction

The binding event of the ligand to the aptamer domain necessary for the regulation performed by riboswitches has been characterized by several groups (Gilbert et al., 2006a; Heppell and Lafontaine, 2008; Montange and Batey, 2006; Winkler et al.,...
However, the regulatory structural switch of these ribo-regulators has been mostly ignored. This is the functional domain of the RNA, which controls the cellular machinery by forming or disrupting a rho-independent terminator. Notably, the sequence composition of these domains is poorly conserved within a family of riboswitches. Since the sequences of aptamer domains are highly conserved, it has been hypothesized that the expression platform is responsible for the variance of riboswitches regulatory ability (Mulhbacher and Lafontaine, 2007; Tomsic et al., 2008). The few reports that have investigated the function of these domains have shown the crucial involvement of expression platforms in gene regulation. However, the structure and function relations of these regulatory domains lack the details describing different aptamer domains (Sherman et al., 2012).

The few studies focused on the expression platform reveal that proper modulation of gene expression relies on a switching mechanism that is largely driven by the formation of helix 1, or P1 (see Figure 1.4). The \textit{in vitro} regulatory activity of the adenine and lysine riboswitches greatly varied with the addition or subtraction of base pairs to this helix (Blouin et al., 2011; Lemay et al., 2011). The over-stabilization of this region allowed the formation of the terminated product independently of ligand presence, indicating the formation of P1. However, introducing mutations in this helix that subtract base pairs, inhibited P1 formation. Additionally, in the lysine riboswitch a helical region outside the aptamer domain was hypothesized to act as a nucleator stem for the terminator hairpin. Increasing or decreasing the stability of this helix by adding or subtracting base pairs influenced the ability to induce the terminated or anti-terminated structures respectively (Blouin et al., 2011). Therefore, altering the stability
of helical regions in the secondary structural switch varied the activity of two different riboswitches \textit{in vitro}.

Finally, another illustration of the importance of the expression platform was performed in the glycine riboswitch. This RNA is composed of two aptamers that function in tandem and cooperatively to control gene expression (Butler et al., 2011; Mandal et al., 2004). However, during in-line probing experiments, which are indicative of secondary structure, different patterns were observed by the presence or absence of the leader sequence (Sherman et al., 2012). This region of the RNA is highly conserved and has been shown to form a helix with the linker region located between the two aptamers. This interaction increases the affinity of the aptamers by 4 to 80 fold for their ligand and additionally, the cooperativity between the two aptamers is no longer present when the full sequence is used (Baird and Ferre-D'Amare, 2013; Sherman et al., 2012). These reports suggested that sequence and structure in the expression platform had significant effects on riboswitch activity and genetic regulation.

Henkin and co-workers indirectly emphasized the importance of studying the expression platforms by illustrating the variation of gene regulation obtained when comparing the SAM-binding riboswitches present in the \textit{B. subtilis} genome (Tomsic et al., 2008). The variation of ribo-regulator activity was demonstrated by \textit{in vitro} transcription in the presence and absence of SAM, as well as \textit{in vivo} by reporting the activity of LacZ, which is located immediately downstream of the riboswitch. The authors demonstrate that the levels of gene modulation differ considerably, which had not been shown before. The hypothesis postulated that the observed differences in gene regulation might be due to the control mechanism imposed by evolutionary
constraints of each gene. The conclusion reached was that important genes that regulate biosynthetic pathways establish tighter controls than genes controlling transport systems (Tomsic et al., 2008). Although this is certainly possible, an in-depth demonstration of this hypothesis is needed, where the aptamer and expression platform of each riboswitch are studied to identify their characteristics and differences.

One component of importance for riboswitch function is the aptamer domain and the affinity for binding its cognate ligand. The slight sequence variation within this domain may be responsible for differences in affinity, which would manifest small variations in gene control. Another feature crucial for regulation is whether a specific riboswitch functions under a thermodynamic or kinetic regime. This condition has been debated in the field and points towards the hypothesis that transcriptional riboswitches may be controlled by kinetic parameters, most of which depend on RNA transcriptional speed. A transcriptional ribo-regulator controls the formation of an intrinsic rho-independent terminator that will cause mRNA synthesis to stop. This helical region followed by a uridine repeat will occur at a specific location in the sequence of the RNA and the formation, or not, of the terminator is not reversible. For transcriptional riboswitches the time that an aptamer utilizes to sense the cellular conditions is crucial for function (Garst and Batey, 2009). Factors that may influence this sensing time may be RNA synthesis speed regulated by NTP concentrations, RNA polymerase, transcriptional protein factors such as NusA, and intrinsic pausing of the transcribed RNA (Wickiser et al., 2005a; Wickiser et al., 2005b). For the reasons listed and because this study is focused on transcriptional regulators the experiments presented in this
chapter maintained all these important factors constant to allow comparing different sequences.

The aptamer of this specific riboswitch, \textit{metE}, has not been characterized in depth, but as mentioned earlier other binding domains of the same family and with significant sequence similarity have been the subjects of different studies. These characterizations have probed the structural characteristics of the aptamer domain, however, they were performed exclusively in the context of these domains in isolation (Heppell and Lafontaine, 2008; Lu et al., 2010). This is perhaps the most problematic factor, because it assumes that the aptamer functions and folds similarly in isolation as it does in the presence of the expression platform.

The goal of this study is to identify the mechanisms of action of a riboswitch (binding or sensing the metabolite and also gene control). The data obtained will show that most of the sequence variations that significantly alter gene expression by the \textit{metE} riboswitch are located within the expression platform. In contrast, mutations in the aptamer domain resulted in either a complete loss of regulatory activity due to the disruption of ligand binding, or resulted in relatively moderate changes in functionality of the riboswitch.

2.2. Results

To examine the functional characteristics of the \textit{metE} riboswitch the entire sequence of this RNA was used. To guide our study, an RNA folding algorithm was employed to predict the secondary structure of the riboswitch. We chose to exploit KineFold as it mimics co-transcriptional folding of the RNA and the speed of the
synthesis can be controlled to resemble that of the RNA polymerase (Xayaphoummine et al., 2005). Additionally, this simulating software was proven useful in the modeling of competing structures of RNA in bacterial systems (Dawid et al., 2009). The metE riboswitch secondary structure obtained is illustrated in Figure 2.1 and it shows interactions not predicted before by other methods such as mfold, which focuses on the energetically favored structures, or by co-variation studies (Zuker, 2003). The new interactions predicted by the software are located in the expression platform of the riboswitch, which is not well conserved across phylogeny. The 5’- side of the expression platform, which is transcribed before the aptamer domain, forms several interactions with the 3’-side of the expression platform. These two sequence regions are not included in the main switching mechanism and do not participate in the anti-terminator helix (P-AT) or the terminator helix (P-T). In Figure 2.1 these two helices can be observed, located adjacent to P1 and therefore are termed P1a and P1b. This terminology will be employed in following sections.
2.2.1. Choice of riboswitch containing common features of an expression platform: *metE*

**Figure 2.1. Sequence and secondary structure of the metE riboswitch.** The riboswitch is modeled in the presence (left) and absence (right) of SAM, with the formation of P1 and P-Anti-Terminator respectively (P-AT). The paired regions are labeled following the order of appearance, 1 through 4. The sequence regions responsible for the secondary structural rearrangement are highlighted in red and green. The residues that formed direct contact with the ligand are highlighted in cyan. The pseudoknot formed between L2 and P4 is illustrated with a line. The regions of secondary structure not previously predicted are highlighted in pink.
To study a riboswitch in depth and determine how its function is achieved, a general target containing common elements of a transcriptional regulator needed to be determined. First, a well-distributed riboswitch family across phylogeny was desired to ensure that the principles learned could be related to different bacterial species. Second, the target riboswitch needed to be active in a single-turnover transcription assay (see next section), which is important for quickly assessing functionality of different regions. Third, the riboswitch of choice needed to harbor secondary structural motifs of a canonical transcriptional regulator (P-AT, PT). Finally, the crystal structure of an aptamer from this family of riboswitches bound to its ligand had been previously solved (Lu et al., 2010; Montange and Batey, 2006). This information is crucial in order to relate structure to function of the aptamer and the entire riboswitch. All these characteristics allow comparing this riboswitch to other ribo-regulators and demonstrating general principles of riboswitch mechanism of action. These characteristics are crucial for identifying general properties of riboswitch function, and ensure applicability of control principles to a wide range of riboswitches.

To ensure that these principles were met the SAM-I family riboswitch was chosen, which is well represented in different clades of bacteria (Barrick and Breaker, 2007). Almost all the members of this regulator family were shown to regulate gene expression in previous reports (McDaniel et al., 2003; Tomsic et al., 2008). The reported levels of transcription control were such that enhancements or deterioration of this regulation could be monitored during in vitro transcription assays (Epshtein et al., 2003; McDaniel et al., 2003; Tomsic et al., 2008). This group of riboswitches also satisfies the universal composition requirement as most of them contain well-defined
anti-terminator and terminator paired regions and their alternative formation occurs as is illustrated in Figure 2.1. Finally, high-resolution crystal structures were reported, demonstrating that several structural points of interest were present in the aptamer domain such as a kink-turn, pseudoknot and a P1 formed by the switching sequence in its 3′-side (Lu et al., 2010; Montange and Batey, 2006; Stoddard et al., 2010) (Figure 2.1). The chosen target, which meets all of the requirements imposed by our general design, was the metE riboswitch. This ribo-regulator has proven to behave similarly to other SAM-I riboswitches, although containing differences in sequence and function.

2.2.2. Single turnover transcription assay was employed to assess functionality

The approach used to interrogate the functionality of each region of interest of the riboswitch was mutagenesis. The sequence variations were performed to alter the suspected interaction formed, either by enhancing or inhibiting the interaction. The effect of each mutant was assessed by a single round in vitro transcription assay. In this method, RNA polymerase was bound to a DNA template containing a strong constitutive T7A1 promoter. After this event has occurred, a mixture of NTPs with the desired concentration of ligand was added and the reaction given 10 minutes to proceed. To avoid re-initiation and to ensure single-turnover transcription, heparin was added as part of the previous mixture. After the reaction was quenched, the 32P-labeled transcriptional products were separated and visualized on a denaturing polyacrylamide gel. The amount of RNA corresponding to the terminated (T) and read through products (RT) were quantified (see Materials and Methods) and the data fitted to a two-state binding equation (Figure 2.2). This analysis yielded two important parameters: T_{50} (the amount of ligand required to elicit the half-maximal regulatory response, and the
dynamic range or DR (the amplitude of the response as the difference between percent terminated at low and high ligand concentrations) (Figure 2.2). These two quantities describe the important functional parameters of a riboswitch, which can be compared with the mutant RNA sequences that may alter the behavior of the ribo-regulator.

While the methodology described above is a convenient and simple means of assaying activity, the transcription conditions in the reaction are necessarily different from those encountered by riboswitches in the cell. A potentially important difference in the assay is the use of E. coli RNA polymerase instead of the B. subtilis polymerase. Different and independent laboratories have explored the use of these two different polymerases. During the characterization of the flavin mononucleotide (FMN) riboswitch, Breaker and co-workers found the mid-point of the ligand titration to be either 0.48 µM with the E. coli polymerase and 1.8 µM with the B. subtilis (Wickiser et al., 2005b). However, the Henkin group demonstrated that for several SAM riboswitches the levels of read-through and terminated products during in vitro transcription were similar regardless of which polymerase was employed (McDaniel et al., 2003). Furthermore, an adenine riboswitch from B. subtilis presents a T_{50} of 2.1 µM with the RNA polymerase from the same organism and 0.5 µM with the E. coli polymerase, but with the same levels of read through and terminated product in the absence and presence of the ligand (Lemay et al., 2011). The lack of important differences in their behavior demonstrate that the characteristics needed for regulation are similar in both polymerases, and that the T_{50}s and dynamic ranges do not vary considerably (Artsimovitch and Henkin, 2009; Epshtein et al., 2003; Grundy and Henkin, 2003; Lemay et al., 2011; Wickiser et al., 2005b).
Another factor of difference is that the transcription conditions are chosen to match physiological ionic conditions as closely as possible including lowering the $[\text{Mg}^{2+}]$ to 2.5 mM. However, the NTP concentrations are low (200 µM total or 50 µM each) in comparison to their intracellular concentration in rapidly growing bacteria (>1 mM) (Buckstein et al., 2008). This slows the rate of RNA synthesis by RNA polymerase, increasing the time that the aptamer domain can employ to fully equilibrate with respect to effector binding prior to the regulatory decision being made. The polymerase will reach the poly-uridine tract of the rho-independent terminator (P-T) and stall briefly and

---

**Figure 2.2.** Representation of data obtained during a typical *in vitro* transcription experiment. **a,** Polyacrylamide gel electrophoresis (PAGE) is used to separate the read-through (RT) and terminated (T) transcripts obtained during a standard single-round *in vitro* transcription assay. The variation in transcription levels of the products as a function of ligand can be observed. The intensity of the transcripts is quantified and fitted to the two state model presented in **b.** These data can be plotted as a function of ligand concentration and fitted to the model in **c,** where two important parameters can be determined: $T_{50}$ and the dynamic range (or range) of the regulatory response.

\[
\% \text{Term} = T_{\text{min}} + (T_{\text{max}} - T_{\text{min}}) \left( \frac{[\text{ligand}]}{[\text{ligand}] + T_{50}} \right) \times 100
\]
disengage from RNA synthesis if the helical region folds properly (Garst and Batey, 2009; Wickiser et al., 2005a; Wickiser et al., 2005b). The consequence of not reaching equilibrium prior to the decision point is to push the riboswitch into a kinetic regulatory regime where $T_{50} > K_D$ (as opposed to a thermodynamic regime in which $T_{50} \sim K_D$).

Another difference between in vitro and cellular transcription is the lack of accessory transcription factors in the in vitro assay, which alters the reaction. Proteins such as NusA decrease the elongation rate and promote pauses (Grundy and Henkin, 2004; Gusarov and Nudler, 2001; Yakhnin and Babitzke, 2002). NusA increases the efficiency of terminators by forming contacts with the polymerase and in turn with the nascent RNA. The most important effect of NusA is its assistance during hairpin formation by suppressing steric hindrances of RNA folding (Gusarov and Nudler, 2001). Previous work has shown that the absence of NusA in the in vitro transcription assay can reduce the $T_{50}$ of a riboswitch; for the B. subtilis FMN riboswitch this was about 2-fold (Wickiser et al., 2005b). These caveats do not invalidate the use of the in vitro transcription assay to assess riboswitch activity, but signify that the results obtained should be taken as an initial description of the activity of riboswitches and how structure of these RNAs directly influences their regulatory ability.

To improve some of the weaknesses detailed above several changes were implemented in this method. The pause of the transcription reaction was eliminated. This was previously performed by mutating the initial sequence (20-30 nucleotides) and eliminating the presence of one nucleotide, this technique allowed the formation of a stable transcription-initiation complex. By discontinuing this practice, the mutations of the sequence within this initial region of the RNA are not needed and thus, I
accomplished studying the riboswitch in its native sequence composition. This approach allows monitoring the entire sequence and the interactions that may form during the native mode of regulation by a riboswitch, which is co-transcriptional folding. This process has been described to influence the RNA structure formation process (Xayaphoummine et al., 2007). These advantages are factors of great importance as they allow the characterization of the sequence in the native regulation mode of transcriptional riboswitches.

2.2.3. The metE riboswitch is functional during single round in vitro transcription assays

To ensure that the *B. subtilis* metE riboswitch is functional in a single round *in vitro* transcription, initial assays were performed. The results were expected to recapitulate, with minor differences, what was previously obtained by the Henkin and Nudler group (Epshtein et al., 2003; McDaniel et al., 2003). However, a significant reduction in the regulatory activity of the riboswitch was obtained (Figure 2.3a). The $T_{50}$ obtained was a high concentration of SAM (86 µM), and the dynamic range was low (73-95% termination) versus the 4 to 99% termination reported by McDaniel *et al.* To explain this difference we identified the following discrepancies in the methodology: usage of a different promoter region glyQS, pausing of the transcription reaction after initiation and stable complex formation (20-30 nucleotides), mutating this initial sequence to remove one nucleotide from this element, or perhaps the fact that the origin of transcription was randomly chosen. This is indicated by the imprecise location of the initiation site as “~20 nucleotides upstream of the start of helix 1” (McDaniel *et al.*, 2003). The transcription start site following their description is probably located at C17,
which removes a significant region of the RNA (Figure 2.1). This difference in the sequence could potentially be responsible for the remarkably different functionality.

![Image of PAGE gel showing metE wild type and metE Δ16 as a function of increasing concentration of SAM.](image)

**Figure 2.3.** The *metE* riboswitch responds to SAM during a single round *in vitro* transcription. **a** and **b**, Show PAGE of *metE* wild type and the Δ16 as a function of increasing concentration of SAM. The data in panels **a** and **b** are quantified, plotted as a function of SAM concentration and fitted to a two state model in the graph in **c**. The difference in T<sub>50</sub> and dynamic range can be appreciated.

To test the role of this sequence predicted to form a small helix (P1b), a construct lacking the initial 16 residues was used for comparison (Figure 2.1). Indeed, the *in vitro* transcription assay of the deletion mutant impeded the formation of the P1b helix, regulated RNA synthesis (Stoddard, 2009) as described earlier (Figure 2.3b). The Δ16 transcript recapitulates the switching ability reported before with a range in termination of 15-92% (Figure 2.3c). However, the T<sub>50</sub> cannot be compared, as it was not determined (McDaniel et al., 2003). These data demonstrate that, under these conditions, the 5′- sequence of the riboswitch can significantly influence its regulatory activity, in direct contrast with the assumption by Henkin and co-workers. Since *metE*
Δ16 regulates the transcription of this gene with greater dynamic range. It is easier to quantify the differences between the wild type and the mutants in this context. However, most variants will be tested in the context of the wild type to ensure that the true effect of each mutant is fully described.

To verify that the observed termination in the transcription assay is due to specific binding by SAM, a mutation that impedes normal SAM binding as guided by the crystal structure was designed (Montange and Batey, 2006). A uracil residue responsible for recognizing and binding the adenosyl moiety of SAM is located in helical region 3 (Figure 2.1 cyan region), and is universally conserved across phylogeny (Grundy and Henkin, 1998) (Figure 2.4). This is a mutation that should result in no alteration of the global structure of the aptamer as only one nucleotide is varied. Indeed, the mutant U157A is incapable of regulating a switch during the transcription assay, as can be observed in Figure 2.4. The fact that the level of transcription of the mutant, in the absence of SAM, does not vary from that of the wild type is indicative of the overall similarity of stability and fold of the U157A and wild type. This is in contrast to another report, where levels of transcription were altered to favor termination in the presence and absence of ligand (Lu et al., 2010). Therefore, the U157A mutation demonstrates that the riboswitch termination activity in vitro is directly due to the binding of SAM with the RNA.
The adenine binding pocket loop located in P3 is conserved universally in the SAM-I riboswitch family, but other motifs are as well. Flanking this loop is a C-G base pair situated immediately up- and downs-tream of the binding pocket (Figure 2.5). Both of these base pairs are universally conserved and the crystal structure explains their purpose. They both form base triples that help to recognize the ligand. The C47•C83-G158 triple interacts with the methionine moiety of the ligand and the A85•C86-G156 stabilizes the ligand by forming stacking interactions. To investigate the role of these base pairs in ligand binding the riboswitch was mutated in this region to transverse these nucleotides. The in vitro transcription of the transversion of either base pair, or the combination of both resulted in a total loss of regulatory activity and the transcript observed was always the read-through product (Figure 2.5). These results
indicate that the formation of the triples is crucial for binding and therefore regulation of the riboswitch.

Figure 2.5. Sequence adjacent to the binding pocket assists during binding. a, Sequence of the adenosyl moeity binding pocket. b, Illustration of the H-bond pattern formed during the interaction formed with the methionine moiety of the ligand, adapted from (Montange and Batey, 2006).

The secondary structure of this aptamer has evolved to bind the ligand with high affinity and specificity; however the conservation pattern is difficult to explain in some cases. There is a residue in P3 that is highly conserved and the structure reveals no specific function, although it is part of the triple previously mentioned A85•C86-G156 (Figure 2.5). To reveal the function of this residue it was mutated to C, G and U and also removed from the sequence. All these substitutions considerably increased the $T_{50}$ value, while maintaining a similar dynamic range of the regulatory response. Notably, the A85G substitution slightly increased the dynamic range but the $T_{50}$ was also considerably greater (Table 2.1 and Figure 2.6), demonstrating that this position is crucial for function. The most interesting change in regulation occurs when this residue is deleted, as A85Δ is incapable of responding to high concentrations of SAM (Table 2.1). This demonstrates that the binding is impaired, likely due to the formation of a stable A84–U157 pair as the helix is forming during the co-transcriptional folding of the
riboswitch. This base pair formation would interfere with the ligand binding event as it normally occurs. These data are consistent with the in-line probing results obtained by Colby Stoddard (Stoddard, 2009), who show similar reactivity by another aptamer. This trend was mirrored by the metE Δ16 construct when it was subjected to an identical mutational analysis.

Table 2.1. Data quantification of A85 mutations in the metE riboswitch.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Effect</th>
<th>T_{50}, µM</th>
<th>% Term - SAM</th>
<th>% Term + SAM</th>
<th>Dynamic Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>entire sequence</td>
<td>86</td>
<td>73</td>
<td>95</td>
<td>22</td>
</tr>
<tr>
<td>A85C</td>
<td>disturb binding pocket</td>
<td>1800</td>
<td>72</td>
<td>99</td>
<td>27</td>
</tr>
<tr>
<td>A85U</td>
<td>disturb binding pocket</td>
<td>280</td>
<td>77</td>
<td>99</td>
<td>22</td>
</tr>
<tr>
<td>A85G</td>
<td>disturb binding pocket</td>
<td>580</td>
<td>64</td>
<td>98</td>
<td>34</td>
</tr>
<tr>
<td>A85Δ</td>
<td>impede binding pocket formation</td>
<td>n.d.</td>
<td>75</td>
<td>77</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2.6. Functionality of mutation of metE A85G. The substitution of residue A85 to a G recapitulates the wild type behavior with a slight increase in dynamic range but a drastic variation of T_{50}.  

---

37
Another site of potential functional importance was a weakly conserved peripheral extension present in P3. However, the two crystal structures of this aptamer published present short extensions, indicating that those constructs may favor crystallization (Lu et al., 2010; Montange and Batey, 2006). To illustrate the function of this unique composition of P3 a truncation of this helix was designed and capped with a GAAA tetraloop (Figure 2.7a). This mutation was studied in the context of the metE Δ16 because of the small difference in the RNA composition and activity, which may not be measurable using the wild type sequence. The dynamic range of this riboswitch during in vitro transcription assays decreased from 77% obtained with the Δ16 sequence to 43% of the mutant lacking most of P3 and the T50 increased from 2 µM to 35 µM (Figure 2.7). This demonstrates that, although not predicted, the P3 extension has an effect on the functionality of the RNA. One could speculate that this helix probably presents a more stable fold with the extension, and that a longer sequence provides more time for the aptamer to interrogate the cellular environment for the presence of SAM.
2.2.4. Formation of a tertiary interaction is responsible for organization of the aptamer domain.

Even though there are regions of *metE* that present high degree of sequence divergence, other areas with similarities exist in the SAM-I riboswitch family. One RNA structural motif that is universally conserved in this riboswitch is a kink-turn in helix 2 that facilitates the formation of an interaction between the loop of this region, L2, and helix 4, forming a pseudoknot. These two motifs have been characterized in depth, demonstrating that they are responsible for crucial pre-organization of the binding
pocket formed by the RNA (Heppell and Lafontaine, 2008; McDaniel et al., 2005; Stoddard et al., 2010).

In the next two sections, different structural motifs important for this RNA function will be investigated. The first section will focus on the paired region 2 and the pseudoknot interaction formed with helix 4, which is guided by a kink-turn. This last motif divides P2 in two sub-helical regions, P2a and P2b in order of transcription (Figure 2.8 brown regions). These two RNA folds are important for the formation of the riboswitch binding pocket. Secondly, the role of P1 helix in the control mechanism of metE will be investigated (Figure 2.8, purple region).

2.2.4.1 Length of P2a is important for regulation

The paired region 2a of metE presents an increased number of base pairs in comparison with other SAM-binding natural aptamers (Grundy and Henkin, 1998; Winkler et al., 2003). The conservation pattern in this helix indicates that a length of 6
to 8 base pairs is well represented (Winkler et al., 2003). However, the *metE* riboswitch presents a 9 base pairs helix. This is a structural characteristic worth investigating to identify the evolutionary advantage of this variation. To reduce the length of this helix, one or two base pairs were removed from the sequence. These deletions of base pairs were performed in the middle of the helix to avoid interfering with the formation of the region, the binding pocket or with the kink-turn composition, see Figure 2.9. The deletion of one C-G base pair was highly beneficial for the activity of the riboswitch by lowering the concentration of SAM needed to obtain half-maximal response, which in this case was 8.4 µM, an approximate 10-fold reduction from that of wild type. The dynamic range remained poor (82-97% termination), indicating that perhaps the extra base pair may affect the formation of the binding pocket by allowing access to a greater conformational space during transcription in the original sequence. The addition of this base pair may have occurred during evolution to diminish the regulatory effectiveness of this riboswitch, more specifically the amount of SAM needed to trigger activity.

Figure 2.9. Secondary structure of *metE* identifying the mutations performed in P2 and P4. Nucleotides mutated are highlighted in brown, and the substitutions indicated. The sequence deletions are indicated in pink.
The deletion of the second base pair in this region yielded an RNA incapable of function. This result suggests that the pseudoknot interaction may not be achievable when the helix lacks certain length. Therefore, the position of L2 with respect to P4 is critical for function, which was previously shown by our laboratory (Stoddard et al., 2010). A similar effect was observed when the metE Δ16 sequence was employed yielding a functionally active riboswitch after the deletion of one base pair in P2a, but a non-responder RNA when deleting two base pairs from the helix (Table 2.2). The deletion of one base pair is consistent with the consensus sequence conservation. However, removing two base pairs from this paired region constitutes a helix length represented in phylogeny, which contrasts with the lack of regulatory activity of this mutant. This demonstrates that the evolution of these riboswitches is tuned to accomplish a desired degree of gene regulation.

Table 2.2. Data quantification of the metE riboswitch and mutants located in P2 in vitro data.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Effect</th>
<th>$T_{50}$, µM</th>
<th>% Term – SAM</th>
<th>% Term + SAM</th>
<th>Dynamic Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>entire sequence</td>
<td>86</td>
<td>73</td>
<td>95</td>
<td>22</td>
</tr>
<tr>
<td>P2 mutations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P2aΔ1bp</td>
<td>shorten P2a</td>
<td>7.2</td>
<td>82</td>
<td>97</td>
<td>15</td>
</tr>
<tr>
<td>P2aΔ2bp</td>
<td>shorten P2a</td>
<td>n.d.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>metE Δ16</td>
<td>eliminate P1b</td>
<td>1.7</td>
<td>15</td>
<td>92</td>
<td>77</td>
</tr>
<tr>
<td>P2aΔ1bp</td>
<td>shorten P2a</td>
<td>0.35</td>
<td>44</td>
<td>99</td>
<td>55</td>
</tr>
<tr>
<td>P2aΔ1bp</td>
<td>shorten P2a</td>
<td>n.d.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.2.4.2. The kink-turn of *metE* is deficient in regulatory activity

Another interesting difference found in the sequence of the *metE* aptamer was the composition of the kink-turn. This motif actively participates in the pre-organization of the binding pocket of the SAM-I riboswitch family, and is also important in the folding of other biologically active RNA molecules. The sequence conservation of these unpaired nucleotides is usually RNN (purine, followed by two non-specific nucleotides), while in the *metE* riboswitch the region presents a composition of UUU (Daldrop and Lilley, 2013; Daldrop et al., 2013; Grundy and Henkin, 1998). In order to study the contribution of this sequence difference I introduced the more represented sequence GUA in the *metE* sequence. The transcription assay of this mutant resulted in a remarkable difference in activity, yielding a $T_{50}$ of 5.0 µM and only a modest change in termination efficiency of 68-94%. This indicates that perhaps the sequence composition of *metE* does not favor the proper fold as much as the better conserved GUA. This sequence was also replaced by an AAU sequence, which resulted in a poor regulator in terms of dynamic range, but that improved the $T_{50}$ to 0.3 µM (*Figure 2.10* and *Table 2.3*). These results demonstrate that the composition of the kink-turn bulged nucleotides have a pronounced effect on the regulation of SAM-I riboswitches as reported by the Lilley group (Daldrop and Lilley, 2013; Daldrop et al., 2013; Lilley, 2012).
2.2.4.3. Residues adjacent to the kink-turn are also important for functionality

As shown in the previous section, the nucleotides that form the bulge of the kink-turn in \textit{metE} do not conform with consensus sequences, which is also the case for the residues adjacent to this motif. For example, the base pair immediately following the turn in the helix that is kinked is well conserved as G-C (Schroeder et al., 2012), and in \textit{metE} it is a non-canonical G•U base pair. Mutating this U residue to a C in position 68 would allow the comparison of these two sequences and illustrate the effect and

---

**Figure 2.10.** The kink-turn composition is relevant for regulation. PAGE of SAM titrations of three different \textit{metE} mutants are presented in a (P2a C55G•C74A), b (bulged nucleotides are GAU instead of the wild type UUU) and c (U68C). Data obtained are quantified and plotted as a function of SAM in d. The wild type is plotted for comparison purposes as well as another kink-turn mutant of the bulged residues by substitution of the wild type sequence UUU for AAU. The data shown are the average of at least 3 independent experiments and the error bars plotted represent the error in the measurement.
function of the U incorporation (Figure 2.9 and Table 2.3). The mutant riboswitch allowed in vitro transcription levels to vary from 16 to 98% termination in the absence and presence of SAM respectively, with an improved T_{50} of 4.9 µM. This result indicates that the riboswitch function is altered greatly by the composition of the kink-turn motif. It is remarkable that the substitution of only one residue greatly affects the transcription levels allowed and the concentration of ligand required to change the expression of the gene (Figure 2.9 and Table 2.3).

Table 2.3. Data quantification of the metE riboswitch and mutations within P2 and P4.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Effect</th>
<th>T_{50}, µM</th>
<th>% Term – SAM</th>
<th>% Term + SAM</th>
<th>Dynamic Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>entire sequence</td>
<td>86</td>
<td>73</td>
<td>92</td>
<td>19</td>
</tr>
<tr>
<td>P2 mutations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KT GUA</td>
<td>UUU change GUA</td>
<td>5.0</td>
<td>68</td>
<td>94</td>
<td>26</td>
</tr>
<tr>
<td>KT AAU</td>
<td>UUU change AAU</td>
<td>0.3</td>
<td>87</td>
<td>97</td>
<td>10</td>
</tr>
<tr>
<td>P2b U68C</td>
<td>G•U bp to G-C</td>
<td>4.9</td>
<td>16</td>
<td>98</td>
<td>82</td>
</tr>
<tr>
<td>P2a bp</td>
<td>C•C bp to G•A</td>
<td>17</td>
<td>70</td>
<td>88</td>
<td>18</td>
</tr>
<tr>
<td>P4 mutations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U168A</td>
<td>Stabilize PK</td>
<td>57</td>
<td>81</td>
<td>96</td>
<td>15</td>
</tr>
</tbody>
</table>

One more base pair of less importance and thus less conserved in the kink-turn architecture is the third base pair of mismatch most distant from the turn in P2a. This pair is usually a G•A mismatch, and in metE it appears as a C•C pair (Daldrop and Lilley, 2013) (Figure 2.9). Mutating the original sequence to contain a G•A pair (C55G and C74A), which is the well-conserved sequence in kink-turn architecture among RNAs, results in a riboswitch of poor regulatory activity. The T_{50} obtained is 17 µM and the regulation ranges from 70 to 87% termination (Figure 2.10, Table 2.3). This implies
that the regulatory activity of the riboswitch is tuned by the nucleotide composition of this region. It is interesting to note that almost all mutations performed in this section that revert the \textit{metE} sequence to a better-conserved kink-turn motif improve its regulatory activity. This suggests that perhaps the RNA has evolved to establish a weak gene control, which does not totally repress or activate gene expression but rather continuously allows a low-level expression of the protein. This mode of riboswitch regulation was first hypothesized by Ferre-D’Amarre and co-workers (Baird et al., 2010), this work however directly demonstrates their idea.

\textbf{2.2.5. Functionality of riboswitches is controlled by the co-transcriptional folding of P1 helix}

The importance of the length of the helical region 1 for ligand-dependent regulation has been demonstrated in recent reports (Blouin et al., 2011; Lemay et al., 2011). Adding or subtracting base pairs to this helix drastically changed the functionality of the adenine and lysine riboswitches (Blouin et al., 2011; Lemay et al., 2011). The addition of base pairs in this region favored the formation of the helix, while the deletion of base pairs impeded P1 stabilization. These results indicate that the competition between the formation of one state (P1), or its alternative state (P-AT) can be manipulated to favor one conformation over the other independent of ligand presence.

The fact that a sequence of RNA can form different secondary structures depending on the transcription conditions has been observed before in other RNAs (Xayaphoummine et al., 2007). One hypothesis that explains these results is the formation of alternative RNA secondary structures as the RNA synthesis occurs called
“encoded co-transcriptional folding” (Xayaphoummine et al., 2007). This model proposes that a combination of the 5’-to-3’ sequential order and relative stability of secondary structural elements directs the folding outcome. A recent study of synthetic riboswitches demonstrated that accounting for the thermodynamics and kinetics of RNA folding proved to be sufficient for the design of functional RNAs. These ribo-regulators fold into one of two mutually exclusive structures despite each being thermodynamically isoenergetic (Xayaphoummine et al., 2007). In application of this concept to transcriptional riboswitches, the competing P1 and antiterminator (P-AT) structures constitute the regulatory switch (Figure 2.1), while the terminator (P-T) is merely a readout of the P1/P-AT outcome. If this model is correct, then the information required to direct RNA folding along one of two folding trajectories is primarily encoded within the P1/P-AT switch of the expression platform. Any mechanism that changes the stability of the P1 helix relative to P-AT in response to a specific signal should therefore be able to act as an efficient effector for the riboswitch.

As a consequence of this model, the intrinsic stability of the P1 helix is a primary site for tuning the ligand-responsiveness of the secondary structural switch. To more precisely define how helix length/stability directs the switch to fold into one of the two regulatory states, a series of point mutations in the metE riboswitch were introduced at the 5’-terminus of its P1 helix (metE mutations shown in Figure 2.11). These mutations alter the stability of the P1 helix while sequence elements required for receptor binding and conformational switching are unaltered. To ensure the response was optimally monitored this study was performed in the context of metE Δ16, as it represents a construct fully capable of switching with no other sequence impediment. The metE
sequence exhibited a sharp boundary between predominantly folding in the “OFF” (terminated) and “ON” (read through) states in the absence of ligand (Figure 2.11). In contrast, the wild type lysC riboswitch is more tolerant to mutations, which will be discussed in detail in the next chapter.

![Diagram of helical region P1 of metE](image)

**Figure 2.11. Helical region P1 of metE controls the transcriptional switch.** a, The strategy to remove or add base pairs to the helix is represented as well as the state favored (P-AT or P1). b, PAGE of the metE Δ16 and mutants after *in vitro* transcription in the presence and absence of the cognate ligand. The transition of the RNA from mostly read-through to mostly terminated can occur independently of SAM presence.

### 2.2.6. Validation of computationally predicted P1a and P1b

Although the gene expression of the *metE* riboswitch varies as a result of sequence alterations in the aptamer, the functionality of this RNA is directed by the expression platform and changes in this domain present a pronounced effect. The regulatory activity variation that can be achieved by altering the secondary structural switch was first illustrated in section 2.2.3. In that segment the effect of the initial 16
residues was demonstrated by removing them from the sequence and yielded a riboswitch of greater regulation range, and a lower $T_{50}$ (Figure 2.1 and 2.4). However, the full characterization of the helical regions computationally predicted (P1b and P1a) is needed to assess general principles of riboswitching activity.

2.2.6.1 Roles of P1a and P1b

To properly test the influence of the full leader sequence in gene expression, a set of mutants systematically truncating the initial region of the riboswitch was tested. These mutations were designed to introduce deletions or sequence identity changes and assess whether the formation of P1a and P1b occur during transcription. These mutations would allow to monitor if the formation of these small helical regions are responsible for the high degree of termination observed during transcription in the absence of SAM. The first variation designed to test this hypothesis is to remove the residues responsible for the formation of P1b, which was obtained by deleting the initial 16 residues participating in the 5'-side of P1b (recall Figure 2.3). This construct changed the regulatory response of the riboswitch drastically by reducing the $T_{50}$ by approximately 43-fold to 1.7 $\mu$M, and the dynamic range varies from 15% terminated to 95%. This implies that the sequence deleted inhibits the formation of the P-AT helix in the absence of SAM. Further truncations of the sequence leading to P1 raised the $T_{50}$ by approximately 4-fold, but the dynamic range of the response only slightly decreased (Table 2.4). These results demonstrated that the formation of P1b is most deleterious for P-AT formation as evidenced by the metE Δ16 construct, whereas the formation of P1a only moderately changed the riboswitch activity as indicated by the regulatory response of the Δ24 and Δ30 sequence truncations.
Table 2.4. Truncations of metE varying the transcription start site change the regulation of the riboswitch.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Effect</th>
<th>$T_{50}$, µM</th>
<th>% Term – SAM</th>
<th>% Term + SAM</th>
<th>Dynamic Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>entire sequence</td>
<td>86</td>
<td>73</td>
<td>95</td>
<td>22</td>
</tr>
<tr>
<td>Δ16</td>
<td>start at +17</td>
<td>1.7</td>
<td>15</td>
<td>95</td>
<td>80</td>
</tr>
<tr>
<td>Δ24</td>
<td>start at +25</td>
<td>8.0</td>
<td>22</td>
<td>92</td>
<td>70</td>
</tr>
<tr>
<td>Δ30</td>
<td>start at +31</td>
<td>8.3</td>
<td>19</td>
<td>87</td>
<td>68</td>
</tr>
</tbody>
</table>

Alternatively, mutations capable of impeding the formation of this helix can probe the regulatory effect of a stable paired region formation. The deletion of the residues located in the 3’-side of the P1b helix (A213, A214, A215 and A216) may affect the regulatory activity of the riboswitch as the previously studied truncations. When two of these residues were deleted, the $T_{50}$ decreased to 51 µM, and the dynamic range varied between 22 and 98% termination in the absence and presence of SAM respectively (Figure 2.12 and Table 2.5). Removal of all four adenosines resulted in a range of termination increasing to 12-98%, while the $T_{50}$ improved to 6.7 µM, nearly recapitulating the metE Δ16 phenotype (Figure 2.12 and Table 2.5). The results obtained with these sequence variations clearly indicate that the formation of a stable P-AT is impeded in the wild type sequence. The mutations studied in this section may be responsible for the stabilization of an anti-terminator in the absence of ligand due to the lack of interaction with U13-U16, as indicated by the reaction favoring the read-through
transcript \((\text{Table 2.5})\). Furthermore, these mutants are capable of switching conformation to efficiently terminate transcription when SAM is present.

![Figure 2.12](image)

**Figure 2.12.** Sequence and secondary structure of the *metE* riboswitch illustrating the mutational strategy described in this section. Deletions mutants are represented in pink, and sequence variations mutants are represented in brown.

To more conclusively demonstrate that formation of helices P1a and P1b interferes with the formation of P-AT, more variations of the sequences involved were characterized. It should be noted that these interactions were not predicted by other software methods such as m-fold. Therefore a precise demonstration of the formation of these helices is needed. One, two or four of the uridines located in the 5’-side of P1b were mutated to adenosines (U14A, U14-15A, and U13-16A) \((\text{Figure 2.12})\), in order to
prevent base pair formation and resulted in a regulatory activity recapitulating that of the metE Δ16. The $T_{50}$s varied in value between 15 and 20 μM, but represented a significant improvement of 4-5 fold with respect to wild type (Figure 2.13 and Table 2.5), partially mimicking the Δ16 mutation. These data further demonstrated that the interaction between the two regions prevents full riboswitching activity. Similar regulation improvement was obtained when the formation of P1b was disrupted by the mutation of residues forming the 3'-side of this helix (positions 213-216 from adenosines to cytidines) (Figure 2.12). The variation of two of these positions (A213C and A214C) yielded a remarkable improvement in the dynamic range, 6-89% termination and a $T_{50}$ of 4.7 μM. The ability of the riboswitch to regulate was improved further by the mutation of the two remaining nucleotides in this location. The A213-6C mutant controls transcription with a $T_{50}$ of 3.7 μM and a dynamic range between 5 and 94% termination (Figure 2.13 and Table 2.5). This mutation yielded one of the best riboswitches in this study presenting a total control over transcription termination as well as a low $T_{50}$. 


Table 2.5. Quantitation of *metE* riboswitch and mutants during single round in vitro transcription. Sequence variations are localized to the P1b helix.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Effect</th>
<th>$T_{50}$, $\mu$M</th>
<th>% Term – SAM</th>
<th>% Term + SAM</th>
<th>Dynamic Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>entire sequence</td>
<td>86</td>
<td>73</td>
<td>95</td>
<td>22</td>
</tr>
<tr>
<td><em>metE</em> Δ16</td>
<td>eliminate P1b</td>
<td>1.7</td>
<td>15</td>
<td>92</td>
<td>77</td>
</tr>
<tr>
<td>U13-16A</td>
<td>mimic Δ16 impede P1b helix</td>
<td>15</td>
<td>29</td>
<td>95</td>
<td>66</td>
</tr>
<tr>
<td>U14A</td>
<td>partially destabilizes P1b helix</td>
<td>17</td>
<td>24</td>
<td>90</td>
<td>66</td>
</tr>
<tr>
<td>U14A/U15A</td>
<td>partially impedes P1b helix</td>
<td>20</td>
<td>25</td>
<td>89</td>
<td>64</td>
</tr>
<tr>
<td>U14C</td>
<td>partially impedes P1b helix</td>
<td>10</td>
<td>25</td>
<td>91</td>
<td>66</td>
</tr>
<tr>
<td>A215G</td>
<td>partially impedes P1b helix</td>
<td>4.7</td>
<td>6.0</td>
<td>89</td>
<td>83</td>
</tr>
<tr>
<td>U14C/A215G</td>
<td>restores P1b</td>
<td>26</td>
<td>73</td>
<td>90</td>
<td>17</td>
</tr>
<tr>
<td>AA213/4CC</td>
<td>partially impedes P1b helix</td>
<td>4.7</td>
<td>6</td>
<td>89</td>
<td>83</td>
</tr>
<tr>
<td>Δ213/4</td>
<td>partially impedes P1b</td>
<td>51</td>
<td>22</td>
<td>98</td>
<td>77</td>
</tr>
<tr>
<td>AAAA213/6CCCCC</td>
<td>disrupts P1b</td>
<td>3.7</td>
<td>5</td>
<td>94</td>
<td>89</td>
</tr>
<tr>
<td>AAAA213/6UUUU</td>
<td>disrupts P1b</td>
<td>12</td>
<td>17</td>
<td>95</td>
<td>88</td>
</tr>
<tr>
<td>Δ213/6</td>
<td>disrupts P1b</td>
<td>6.7</td>
<td>12</td>
<td>98</td>
<td>86</td>
</tr>
</tbody>
</table>

An alternative manner in which the P1b helix can be further tested is by performing a compensatory mutation on each side of the paired region, which is a standard practice when demonstrating RNA helical interactions. The U14C mutant resulted in a riboswitch capable of changing transcription levels from 25 to 91% termination with 0 and 1 mM SAM respectively and a $T_{50}$ of 10 $\mu$M. When the sequence alteration is performed on the 3’-side of P1b, A211G, a similar effect is observed in dynamic range, 6-89% termination, and in $T_{50}$ 4.7 $\mu$M. The transcription of the RNA
containing both mutations that would restore P1b showed a 73-90% termination in the dynamic range and a higher $T_{50}$ of 26 $\mu$M, consistent with a restored interaction and a stable formation of P1b (Figure 2.13).

Figure 2.13. Single round in vitro transcription data of the metE riboswitch and mutants. Data quantification illustrates the behavior different mutations of metE as a function of SAM. a, U14C mutant recapitulates the behavior of the $\Delta$16 truncation and the wild type behavior is rescued with the U14C/A215G mutations. b, Mutations disrupting the normal formation of P1b, on either the 5' or 3'-side, are presented.

The effects observed from the characterization of all these sequence variations demonstrate the interaction between the two regions of the expression platform of the riboswitch. The formation of the P1b helix is responsible for strongly repressing the anti-terminator and preventing the competition of the two structures regulating the expression of the gene (P1 and P-AT). Considered together, these data illustrate that the full sequence of this riboswitch may suppress the genetic control ability by favoring the terminated form in the absence of ligand, which is accomplished by providing an
energetic barrier to the anti-terminator formation. This unpredicted behavior correlates with the evolutionary tunability of these riboswitches to perform a function of regulating gene expression accordingly with the needs of the cell (Stoddard et al., 2013). These interactions may prevent the digital control of a gene by a riboswitch as was originally presumed and shown (Tomsic et al., 2008).

2.2.6.2. Stabilization of P5 can achieve anti-termination

It has been hypothesized that a nucleator stem can be formed outside the aptamer domain and assist the stabilization of an important helix (Blouin et al., 2011). In the case of metE a P5 could act as a nucleator for the formation of the P-AT. Although this helix is not predicted to form in the secondary structural model (Figure 2.1), it is not difficult to envision a set of mutations that would favor the formation of this arrangement. The native sequence can be easily substituted by four G-C or C-G pairs capped with a stable GAAA loop (Figure 2.14). When the mutations were incorporated into the sequence, the RNA was able to induce an anti-terminator fold in the absence of SAM and a structural rearrangement when the ligand was present corresponding to 16-87% termination switch and a T_{50} of 47 μM. When these mutations were performed in a more conservative manner, one residue variation per sequence studied, the results yielded similar regulatory activities (Figure 2.14 and Table 2.6). These results demonstrate that the disruption of P1a is advantageous for riboswitch activity, as well as the formation of a nucleator stem, or P5, that can favor the anti-terminator fold over other structures. It should be noted that these mutations do not alter the sequences responsible for P1b formation, indicating the importance of co-transcriptional folding for riboswitch based regulation since P5 and its formation precedes that of P1b.
Table 2.6. Data obtained in titrations of *metE* and mutations located in the P5 region.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Effect</th>
<th>T&lt;sub&gt;50&lt;/sub&gt;, µM</th>
<th>% Term – SAM</th>
<th>% Term + SAM</th>
<th>Dynamic Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>entire sequence</td>
<td>86</td>
<td>73</td>
<td>92</td>
<td>19</td>
</tr>
<tr>
<td>P5</td>
<td>Perfect P5 formed Impedes P1a, P1b</td>
<td>47</td>
<td>16</td>
<td>87</td>
<td>71</td>
</tr>
<tr>
<td>U200C</td>
<td>P5 improvement</td>
<td>14.5</td>
<td>38</td>
<td>83</td>
<td>45</td>
</tr>
<tr>
<td>U209C</td>
<td>P5 improvement</td>
<td>6.8</td>
<td>33</td>
<td>98</td>
<td>65</td>
</tr>
</tbody>
</table>

Figure 2.14. Mutation strategy of *metE* riboswitch to study the formation of the helix P5. *a*, Representation of mutations incorporated into the *metE* sequence. *b*, Graphical representation of the SAM titrations of the mutants is illustrated.

2.2.7. *In vivo* characterization of the *metE* riboswitch

The data presented in this section was obtained by John Zinder.
To determine whether the riboswitch regulatory activity is accurately described during the \textit{in vitro} transcription assays, I introduced the wild type and selected mutant sequences into a reporter system (Figure 2.15). The riboswitch of choice is located directly upstream of and can control the expression of the \textit{mCherry} gene. \textit{E.coli} cells are transformed with the vector containing the desired riboswitch sequence, which constitute a controllable model system. This reporter gene was chosen to easily monitor its expression by measuring the fluorescence difference as a function of SAM presence in the growth medium.

A challenge in this experimental approach is the fact that SAM and its precursor methionine are crucial metabolites for cell growth. In contrast to the \textit{in vitro} transcription assays where the metabolite concentrations can be controlled, the SAM concentration in the medium cannot be varied as it affects cell fitness. Introducing a mutation within the aptamer of \textit{metE}, U157A, which abolishes the ability of the aptamer to bind the metabolite, solved this problem. This result was validated in our \textit{in vitro} transcription assay and previously shown in section 2.2.3. Therefore, I constructed this variant in all the sequences of interest and their comparison was performed in the presence of methionine in the media. The unchanged sequence is indicative of the behavior of the riboswitch capable of binding SAM, while the U157A represents the riboswitch behavior in the absence of the ligand, as the mutation impedes ligand binding.

\textbf{2.2.7.1. P1b helix controls gene expression of \textit{metE in vivo}}

To demonstrate the importance of the leader sequence, and whether the
formation of P1b in the metE riboswitch controls genetic regulation, the entire sequence, the Δ16 metE as well as other mutants were cloned upstream of the reporter gene. The results obtained yielded expression levels of mCherry that generally resemble the trends observed during the in vitro transcription assays. The full-length riboswitch resulted in poor regulatory activity, which was enhanced by the deletion of the initially transcribed 16 residues (Figure 2.16). These results recapitulate those obtained during in vitro transcription and demonstrate that the sequence transcribed initially is critical for achieving the low level of expression of the metE gene. Thus, this short leader sequence allows only for low level of gene expression even in the absence of SAM in the cellular context. To probe the formation of the P1b helix, previously tested mutations were introduced in the riboswitch (see section 2.2.7.1.) The U14C mutation

Figure 2.15. Map of the parental riboswitch reporter plasmid (pRR1). The plasmid is derived from pBR322, from which the tetracycline resistance gene was removed and replaced with a gene encoding GFPuv (green) under control of a constitutively expressed tac promoter (yellow). The sequence immediately upstream of the tac promoter is an “insulator” sequence preceded by a strong rmB T₁T₂ terminator (red). All riboswitch leader sequences (orange) were cloned between the Nsil and HindIII sites.
was responsible for the disruption of the helical region, and in turn a compensatory mutation was capable of rescuing the wild type behavior (U14C, A215G). These two mutants yielded induction and repression of the reporter gene that recapitulated the behavior described by the *in vitro* data (Figure 2.16). Surprisingly, a mutation that robustly enhanced the activity of *metE in vitro*, Δ213-6, resulted in a regulation of *mCherry* similar to that of the wild type sequence. Although this result argues against the formation of the P1b helix, it is important to consider that the RNA folding mechanism in the cell may not mimic its behavior *in vitro*. In this experiment the fluorescence of an empty vector, and a plasmid capable of expressing the gene were employed as negative and positive controls respectively and they are represented in Figure 2.16.

![Figure 2.16](image)

**Figure 2.16.** *In vivo* genetic control of the *metE* riboswitch. Normalized fluorescence of reporter gene *mCherry* measured for the sequences represented on the x-axis. The red bars illustrate the fluorescence obtained with the binding competent sequence while the blue bars represent the U157A mutation responsible for impeding SAM binding. The data are the average of at least three measurements and the error bars represent the uncertainty of the value.
2.2.8. Do other riboswitches regulate gene expression similarly

To determine if the narrow genetic regulation range observed in metE is applicable to other riboswitches, another sequence was characterized in vitro and in vivo. To establish the generality of this effect, I chose to study the pbuE riboswitch from B. subtilis, which is outside of the SAM-I family, for several reasons. First, this RNA has been extensively studied and characterized using a variety of structural, biochemical, and biophysical techniques (Delfosse et al., 2010; Lemay et al., 2011; Lemay et al., 2009; Lemay et al., 2006; Mandal and Breaker, 2004). Second, it regulates gene expression at the level of transcription, which facilitated rapid comparison of the regulatory response between truncations using our in vitro transcription assay. Third, the pbuE riboswitch is capable of modulating gene expression when bound to 2-amino purine (2AP), an adenine analogue that is imported into the cell and is not susceptible to enzymatic modification. Furthermore, this ribo-regulator is composed of an expression platform that is predicted to have a simple fold and is relatively short in length. An in-depth characterization of the pbuE riboswitch is presented in chapter 4, whereas in this section its usage is solely for the purpose of comparison to metE.

The pbuE gene in B. subtilis encodes a purine base efflux pump (Johansen et al., 2003) that is capable of decreasing the concentration of adenine in the cell when a dangerous level has been reached, as adenine is not very soluble. The leader sequence of this gene forms a stable transcriptional rho-independent terminator capable of stopping mRNA synthesis. This effect can be reversed by the presence of the purine, which binds the RNA, altering its secondary structure into a conformation that promotes synthesis of the downstream gene (Figure 2.17). This represents a change from the
data shown up to this point. Instead of representing the “% termination” increase by the addition of ligand, the data will be presented as an increase of the “% read through” as a function of 2AP present in the reaction, or in the case of the in vivo studies, the data are plotted as fold induction of the reporter gene as a function of 2AP added to the growth medium.

Figure 2.17. Sequence and secondary structure of the pbuE riboswitch. The riboswitch forms a stable rho-independent terminator (P-T) in the absence of ligand. The RNA can bind adenine (Ade) or 2-aminopurine (2AP) and form the aptamer domain, and P1. The different transcription initiation sites are indicated on the unbound form.

As in metE, a set of different transcription start sites was selected in the pbuE sequence to determine the impact of the leader region on regulatory function. The pbuE riboswitch was truncated by removing the initial 11 or 25 nucleotides (Δ11, Δ25) (Figure 2.17), and probed using both in vitro transcription and the cell-based reporter system. The wild type sequence presents an initial read through level of 22% in the
absence of 2AP, and only reaches a 10% increase of this value in the presence of saturating amounts of 2AP (1mM). This demonstrates the small dynamic range of riboswitch activity performed by the full wild type sequence. A greater dynamic range of regulation was obtained when deleting the initial 11 nucleotides (8-58% read through). The activity however, was not further enhanced by the deletion of 15 additional nucleotides (removing a total of 25). Instead, this truncation resulted in an intermediate riboswitch performance yielding a 15% read through transcript in the absence of ligand, which increased to 38% when the concentration of 2AP was 1 mM (Figure 2.18). These studies yielded a significant effect in the ability to regulate gene expression by the different truncations, demonstrating the repression of riboswitch activity by the full wild type sequence. The data showed a remarkable similarity with the studies characterizing the metE sequence, signifying that perhaps the importance of the transcription start site can be of general character for riboswitch function (compare Figure 2.18a containing pbuE data and Figures 2.3, and Table 2.4 for metE data).

Trends in regulatory activity for the pbuE riboswitch established by the in vitro transcription assay were mirrored in vivo using the cell-based reporter system. The wild type sequence induces a 3.5-fold induction of the reporter gene, whereas the induction of this gene can be increased to 10-fold by removing the initial 11 or 25 nucleotides of the sequence. This suggests that these sequences interact with the rest of the riboswitch sequence and prevent a robust response as was demonstrated for the metE regulator. It should be noted that the software employed to predict the co-transcriptional secondary structure of the metE (kinefold) accurately predicted the interaction of residues in the initial 11 residues with another region of the aptamer. This small helix is
likely responsible for the limitations imposed on the dynamic range, which can be expanded when these residues are deleted (for comparison see Figures 2.18 b and 2.16).

![Graphical representation of the response of the pbuE riboswitch and the alterations of the transcription start site.](image)

**Figure 2.18.** Graphical representation of the response of the *pbuE* riboswitch and the alterations of the transcription start site.  

**a,** Graphical representation of data obtained during *in vitro* transcription of the wild type sequence and truncations of the *pbuE* riboswitch as a function of 2AP are represented. The black data and fit represent the wild type riboswitch, blue represents the truncation of the 11 initial nucleotides and in red is the Δ25 initial residues removed when collecting these data. **b,** The same study as illustrated in a, however, the titrations are performed *in vivo.*

### 2.3 Discussion

In this chapter I have shown, using both *metE* and *pbuE* riboswitches, that short leader sequences have the potential to alter the folding landscape of an RNA and can also have a significant impact on regulatory function. Importantly, these interactions had not been predicted previously and have not been observed due to limitations in the experimental design and approach by several groups. The biological implications of our findings are significant in guiding how we understand the manner in which bacteria fine-
tune and maintain cellular homeostasis. Rather than these riboswitches operating in a regulatory regime where the levels of gene expression vary over a wide range, our data demonstrate that these switches likely allow for low levels of constitutive gene expression within a small regulatory window. Specifically, while helix formation in the leader region of *metE* was shown to inhibit assembly of the anti-terminator, the leader sequence in *pbuE* disrupted formation of the aptamer and prevented ligand binding. Although the mechanism by which the leader sequence influences gene expression differs in *metE* and *pbuE*, the end result is identical in that it diminishes the magnitude of regulatory activity in both riboswitches.

In this chapter I have demonstrated a form of regulation built into riboswitches that had not been illustrated before. The initially transcribed sequence can alter the function and the response range of these ribo-regulators by forming small helices. The two examples studied in this section illustrate the importance of these interactions. Although not initially predicted or involving many residues (helices of 5 nucleotides), these interactions present a profound effect in regulatory response. In the case of *metE*, the helical region involving the two portions (5’- and 3’-side) of the expression platform does not influence the aptamer structure, as both of these structures can form. This helix formation prevents the formation of the anti-terminator. However, in the case of *pbuE*, the interaction prevents the formation of a stable aptamer, demonstrated by the difficulty for ligand binding. Even though these two riboswitches may differ in the mechanism of action, the end result is identical, a prevention of riboswitch full-response control when the wild type sequence is allowed to co-transcriptionally fold.
This is a novel discovery involved in the functionality of two different riboswitches. In the past this effect has not been observed due to different limitations in the design of the experimental approach of various groups. The biological implications of these observations are indicative of the genetic regulation performed by these ribo-regulators. Cellular homeostasis in bacteria may be established by proteins functioning constantly. The production of genes in a cell is probably not a process that varies from total repression to total induction. However, a continuous production of an mRNA is likely to induce a small variation in the significant cellular metabolites and their concentration. Perhaps this mode of regulation is indeed how riboswitches control gene expression in bacteria. To fully determine whether this is a plausible general control mechanism of riboswitches, a greater survey is needed. However, to quickly predict secondary structure formation of a riboswitch a computational approach capable of verifying these interactions reliably would be of great assistance. The computational study could aid in the initial characterization of different expression platforms of riboswitches, which would encompass many RNA sequences without having to focus on transcriptional regulators.

A related study of the glycine riboswitch, which is composed of two aptamers that function in tandem, demonstrated the importance of the leader sequence in gene regulation. The initial sequence of this riboswitch is responsible for the interaction that abrogates the allostery monitored during the binding event of the ligand to the two tandem aptamer modules (Sherman et al., 2012). This cooperativity had been previously reported by two independent groups (Butler et al., 2011; Mandal et al., 2004).
In addition to the leader sequence effect in *metE*, it is interesting to note that this riboswitch has evolved to minimize its genetic regulatory activity in response to the presence of SAM. As it has been previously discussed, the *metE* riboswitch presents sequence variations in different regions not well represented in phylogeny (Grundy and Henkin, 1998; Winkler et al., 2003). All these variations demonstrate a diminishing effect in riboswitch response as illustrated by the increased activity of reversing the nucleotides to a better-represented sequence. Therefore, the repression of switching ability is acquired through evolution in two different manners: i) incorporating an inhibiting interaction in the expression platform and ii) allowing the aptamer to alter its sequence for less favorable nucleotides. This has probably occurred through evolution with a significantly important pressure for selection such as function.

A more complete description of other riboswitches is necessary. It is obvious that these findings need to be further investigated in other regulatory RNAs. Perhaps this general principle of function repression is important for transcriptional riboswitches. However, it points to the fact that evolution may incorporate small differences to accomplish a desired outcome. Other studies have also shown that leader sequences and sequence variations may be responsible for function alterations. These changes in nucleotide composition may be crucial for incorporating functional differences to expression platforms.

2.4 Materials and Methods

2.4.1. Preparation of sequence variations of the *metE* riboswitch.
The *metE* riboswitch was obtained from PCR amplification of the sequence from the genomic DNA of *B. subtilis*. This riboswitch was placed immediately downstream of the T7A1 promoter using recombinant PCR. This sequence was cloned into pUC19 (Invitrogen) in the multiple cloning site. This arrangement allowed to change (delete or add) the identity of any nucleotide by site-directed mutagenesis, using standard protocols. The sequence variations, deletions or additions were verified prior to use in transcription assays.

2.4.2. *In vitro* transcription assays

DNA templates containing a chimeric riboswitch were transcribed as previously described (Trausch et al. (2011). Briefly, 100 ng of DNA were incubated at 37 °C for 15 minutes in 25 µl of 2X transcription buffer (140 mM Tris-HCl, pH 8.0, 140 mM NaCl, 0.2 mM EDTA, 28 µM β-mercaptoethanol and 70 mg/mL BSA), 5 µl of 25 mM MgCl₂, 1 mCi of α-32P-ATP and 0.5 units of *E. coli* RNA polymerase 70 holoenzyme (Epicentre Biotechnologies) in a final volume of 35 µl. The reactions were initiated by the addition of 15 µl of NTP mix (165 µM each), 0.2 mg/ml heparin and the desired concentration of SAM. After incubation at 37 °C for 15 minutes the reactions were quenched with an equal volume of 8 M urea and 2 minutes incubation at 65 °C. The species were separated by denaturing PAGE, and the gel dried and exposed to a phosphorimager screen. Quantitation of radioactive counts in each band was performed with ImageQuant software (Molecular Biosystems) and data fit to a standard two-state model with nonlinear least squares analysis.
2.4.3. *In vivo* assay

A reporter plasmid for testing riboswitch function in *E. coli* with an *mCherry* reporter gene was constructed using pBR322 as the parental vector using standard molecular biological techniques. The map of the plasmid constituting the *mCherry* reporter is shown in **Figure 2.15**. The mutation U157A that inhibits SAM binding was incorporated into the desired sequences to establish the levels of gene expression when no SAM was bound. The resultant vectors were transformed into *E. coli* strain XL-Top. Single colonies were picked and grown overnight in 3 ml of CSB media + 100 µg/ml ampicillin. This saturated culture was used to inoculate 100 ml of fresh medium, and allowed to grow to early exponential phase (OD$_{600}$ = 0.1-0.5). Samples of these cultures were taken in triplicate (3 ml each sample). The cells were allowed to grow for 6 hours at 37 °C. At this point 300 µl of the cultures were used to measure the OD$_{600}$ and their fluorescence intensity in a plate reader (Tecan). Fluorescence measurements were taken at an excitation wavelength of 570 nm and the fluorescence was monitored at 612 nm, where the maximum emission for mCherry was observed. Optical density-normalized fluorescence values were plotted in the presence or absence of the mutation that allows SAM binding. The background fluorescence was obtained by performing a ligand titration into cells carrying the parental plasmid pBR322 and calculating the OD$_{600}$- normalized fluorescence. An average of the background was subtracted to the values of the cultures containing the plasmids with the chimeric riboswitch constructs.
Chapter 3

Modularity of select riboswitch expression platform enables facile engineering of novel genetic regulatory devices

RNA-based biosensors and regulatory devices have received significant attention for their potential in a broad array of synthetic biology applications. One of the primary difficulties in engineering these molecules is the lack of methods to facilely link sensory modules, or aptamers, to readout domains without extensive screening or selection of sequences that facilitate interdomain communication. In bacteria, a widespread form of gene regulation by riboswitches performs this task with sufficient fidelity to control expression of biosynthetic and transport proteins essential for normal cellular homeostasis. In this chapter, I demonstrate that select riboswitch readout domains, called expression platforms, are modular in that they can host a variety of natural and synthetic aptamers to create novel chimeric RNAs that regulate transcription both in vitro and in vivo. Importantly, this technique does not require selection of device-specific “communication modules” required to transmit ligand binding to the regulatory domain, enabling rapid engineering of novel functional RNAs.

3.1. Introduction to RNA biosensors

Devices composed of RNA are emerging as important tools in synthetic biology for applications including in vivo detection of small molecules with genetically encodable biosensors and engineered regulatory elements to control gene expression (Isaacs and Collins (2005). In part, this is because of the ability to select nucleic acids with a
desired activity using well-established *in vitro* evolution methods (SELEX) (Joyce, 2004). This method subjects a starting pool of random sequences to multiple rounds of a selection pressure and replication of the survivors. The result, when the selection pressure is for binding, is an aptamer capable of binding the target molecule with high affinity and specificity. In the 20 years since its development, SELEX has yielded RNAs capable of binding a chemically diverse assortment of compounds, revealing that aptamers for almost any desired molecule can be obtained (Lee et al., 2004). For many applications, the aptamer must be coupled to a readout module that transduces the binding event into a detectable signal, a significant hurdle in the engineering of practical devices.

Several distinct solutions have emerged that detect binding of a target molecule to an RNA (Famulok and Mayer, 2011). One example is the “aptazyme” in which an aptamer is linked to a hammerhead ribozyme such that ligand binding induces the RNA to cleave itself, an event that can be readily detected either *in vitro* or *in vivo* (Soukup and Breaker, 1999a). A striking example of the utility of such devices was a recent use of a theophylline-sensing aptazyme to monitor the directed evolution of a theophylline-degrading protein enzyme (Michener and Smolke, 2012). Engineering of an aptazyme often requires extensive optimization of a “communication module” between the two domains (Soukup and Breaker, 1999b). However, it has been observed that aptazymes display high background cleavage levels and slow allostery, which can limit their utility (de Silva and Walter, 2009; Suess et al., 2004). Other readout modules that have been explored are designed or evolved secondary structural switches (Lynch et al., 2007; Wachsmuth et al., 2012; Weigand et al., 2008) and fluorophore binding aptamers (Paige
et al., 2011; Stojanovic and Kolpashchikov, 2004). As with aptazymes, these devices can be difficult to engineer and suffer from high background or low dynamic ranges. Ideally, modular readout domains should be as “plug-and-play” as aptamers domains, which can be easily placed into different contexts with minimal or no re-design.

A potential source of modular readout domains is found in naturally occurring genetic control devices called riboswitches. These RNA elements are generally found in the 5’-leader sequence of mRNAs and are broadly distributed across bacteria regulating a wide assortment of metabolic processes (Barrick and Breaker, 2007). Riboswitches are generally composed of two distinct functional domains: the aptamer domain that serves as the receptor for a small molecule and an expression platform containing a secondary structural switch that directs the gene expression machinery (Figure 3.1a) (Roth and Breaker, 2009). Currently, over twenty riboswitches have been characterized whose aptamers bind a chemically diverse set of molecules including nucleobases and nucleosides, amino acids, protein cofactors, metal ions, second messenger signaling molecules, and aminoglycosides (Roth and Breaker, 2009). These naturally occurring aptamers have been exploited in engineered biosensors, reflecting their modularity and utility (Nomura and Yokobayashi, 2007; Wieland et al., 2009).
In contrast to the high degree of structural conservation of each aptamer class, expression platforms are highly variable. These domains regulate by a number of mechanisms, including transcription, translation, mRNA stability and splicing (Roth and Breaker, 2009), typically through their ability to form alternative structures in a ligand-directed fashion (Barrick and Breaker, 2007). For example, expression platforms that direct transcription generally contain two alternative hairpin structures: an intrinsic terminator (P-T, Figure 3.1a) that directs RNA polymerase (RNAP) to abort mRNA
synthesis and an antiterminator that enables the RNAP to proceed through the 5′-leader. Analogous structures comprising ribosome-binding site (RBS) sequester and antisequester hairpins control translation (Figure 3.1b). However, expression platforms are poorly characterized, both with respect to their structure beyond recognition of the binary switch and how they communicate with the aptamer domain. To date, the potential for the regulatory expression platforms as modular readout elements that can serve as building blocks for novel riboregulatory elements has not been explored.

In this study, I demonstrate the modularity of expression platforms of select riboswitches that control gene expression at the transcriptional level. These chimeric riboswitches comprising novel combinations of riboswitch aptamer domains and expression platforms retain the regulatory properties of biological sequences. Importantly for synthetic biology, we find that the secondary structural switch of these expression platforms can be directed by synthetic aptamers. The engineering strategy that I employ to couple aptamers with these modular expression platforms is conceptually simple and requires minimal screening of variants to find an effective riboswitch. These switches are functional in both a basic in vitro transcription assay and as a regulatory element controlling expression of a reporter gene in E. coli. Our data reveal a robust strategy for generating a wide variety of practical RNA devices.

3.2. Results and Discussion

3.2.1. Considerations in design of modular expression platforms.

The two primary criteria used for selecting modular expression platforms from biological sequences are: i) the two domains of the riboswitch must not overlap in their
sequence requirements and ii) easy and rapid assays for their in vitro and in vivo function of readout must exist. To determine whether the domains of a riboswitch are separable, the alternative secondary structures must be considered. For the archetypal "OFF" switch two distinct forms exist: a ligand-bound form comprising the intact ligand-bound aptamer and a rho-independent transcriptional terminator/ribosome binding site (RBS) sequester hairpin or an unbound aptamer that is partially disrupted by formation of the antiterminator/RBS antisequester hairpin (Figure 3.1a, b). In this model, the two domains intersect within a secondary structural element called the P1 helix. Within this helix are sequence elements crucial for productive ligand binding to the aptamer domain or part of the secondary structural switch of the expression platform.

For a large subset of riboswitches, the sequence requirements of the two domains do not overlap within the P1 helix, allowing for assignment of a functional boundary. For example, in one class of S-adenosylmethionine (SAM) binding riboswitches (known as SAM-I) only two A-U base pairs near the 3'-end of the P1 helix are required for ligand binding as revealed by the crystal structure of an aptamer-SAM complex (Figure 3.1c) (Montange and Batey, 2006). Of the eleven SAM-I riboswitches in the Bacillus subtilis genome (McDaniel et al., 2003; Winkler et al., 2003), seven present slight or no overlap between these A-U pairs and the switching sequence. This suggests the two domains can be decoupled in many riboswitches. Since crystal structures of most of the riboswitch aptamers have been determined, information about the ligand binding requirements for almost all riboswitches is available for this analysis (Montange and Batey, 2006). It is important to note, however, that a number of riboswitches cannot meet the above requirement for several reasons. For example,
riboswitches in which the penultimate helix is part of a pseudoknot generally require a sequence that is essential for both ligand recognition and for secondary structural switching (Batey, 2011, 2012). In the most extreme cases of the SAM-II and SAM-III classes, the two domains completely overlap (Batey, 2011). Well-characterized classes of riboswitches that are populated by suitable candidates include the purine, SAM-I, lysine, thiamine pyrophosphate and flavin mononucleotide (FMN) (Roth and Breaker, 2009).

The second consideration is that the regulatory domains function by a well-understood mechanism and whose activity can be rapidly assessed. While structural switches that control translation are most frequently used in designed riboswitches (Lynch et al., 2007; Sinha et al., 2010), I have concluded that the RNA sequence requirements of intrinsic (rho-independent) transcriptional termination are mechanistically better understood, and thus more amenable to design. In addition, well-established and simple in vitro single turnover transcriptional termination assays using commercially available E. coli RNAP can be performed to rapidly assess activity of novel riboswitches prior to their introduction into a cell (Artsimovitch and Henkin, 2009). In our implementation of this assay, a riboswitch is placed downstream of a strong T7A1 promoter in a DNA template and RNAP is allowed to transcribe only a single round (the presence of excess heparin allows no re-loading of the RNA polymerase) to ensure that the ligand and NTP concentrations remain constant. We have further strived to match physiological ionic conditions as closely as possible, but the NTP concentrations are low (50 µM) compared to intracellular concentrations during log-phase growth (>1 mM) (Buckstein et al., 2008), significantly slowing the rate of transcription a key parameter in
riboswitch function (Garst and Batey, 2009; Garst et al., 2012; Wickiser et al., 2005a; Wickiser et al., 2005b). Finally, *E. coli* RNAP holoenzyme was used rather than native RNAP (*B. subtilis* for the riboswitches that are used in this study). Previous work has shown that the difference in activity between riboswitches transcribed with *B. subtilis* and *E. coli* polymerases is moderate (Lemay et al., 2011; McDaniel et al., 2003; Wickiser et al., 2005b). Further, since *E. coli* is an organism of choice for many synthetic biology applications, it is important to assess the activity of artificial riboswitches using *E. coli* RNAP.

Using the above considerations, I focused upon three candidate riboswitches for further investigation: those that regulate the *B. subtilis* metE, yitJ, and lysC transcriptional units. All three riboswitches are “OFF” switches that terminate transcription in the presence of the effector ligand (SAM for metE and yitJ, lysine for lysC) and efficiently terminate transcription in a ligand-dependent fashion in a single turnover transcription assay (Artsimovitch and Henkin, 2009) ([Figure 3.2a](#)) consistent with previous reports (Blouin and Lafontaine, 2007; Epshtein et al., 2003; Garst et al., 2012; Grundy and Henkin, 1998, 2003; McDaniel et al., 2003; Sudarsan et al., 2003b; Winkler et al., 2003). In this assay, the ^32^P-labeled transcriptional products are visualized by separating the RNAs on a denaturing polyacrylamide gel. The amounts of RNA corresponding to the terminated (T) and fully transcribed read-through products (RT) are quantified (see Materials and Methods) and the data fit to a two-state binding equation ([Figure 3.2b](#)). This analysis yields three important parameters: $T_{50}$ (the amount of ligand required to elicit the half-maximal regulatory response), the maximal transcriptional termination at high ligand concentration ($T_{max}$), and the dynamic range
or DR (the amplitude of the response as the difference between percent terminated at low and high ligand concentrations) (Table 3.1). To directly compare $T_{50}$ and affinity of each aptamer for its ligand, the apparent equilibrium dissociation constant ($K_D$) was measured under transcription conditions (buffer matched and at 37 °C) by isothermal titration calorimetry (Table 3.1). While other candidate riboswitches, particularly of the purine class, besides the aforementioned three were considered they showed no riboswitching activity in the transcription assay (data not shown) and were not further considered.

![Figure 3.2](image)

**Figure 3.2. A chimeric xpt/metE riboswitch regulates transcriptional termination.** (a) Single round *in vitro* transcription assay of the wild type *B. subtilis* metE riboswitch as a function of S-adenosylmethionine (SAM) concentration. The two observed bands correspond to the read-through (RT) product in which the riboswitch allows transcription of the full mRNA and the terminated product (T), where transcription is stopped at the terminator. (b) *In vitro* transcription data of the wild type *B. subtilis* metE SAM-responsive riboswitch and the xpt/metE guanine-responsive chimera fitted to a two-state transition. The raw data for these two titrations are shown in panels (a) and (d). (c) Titration of the xpt/metE chimera with SAM. (d) Titration of the xpt/metE chimera with guanine.
Several characteristics can be noted about these three riboswitches. First, the SAM riboswitches are efficiently regulated by S-adenosylmethionine in this assay—both riboswitches show a high dynamic range and are almost completely "OFF" at high effector concentrations (95% terminated, Figure 3.3 and Table 3.1). The lysC riboswitch displays a lesser degree of efficiency both with respect to the dynamic range (52%) and the degree of termination at saturating ligand concentration (78%). The lower efficiency of transcriptional termination may indicate that this RNA may be less suitable for in vivo applications (see below) due to the potential for leaky expression. Second, both SAM riboswitches under our experimental conditions are regulated by ligand concentrations equivalent to the aptamer’s affinity for SAM. This “thermodynamic control” reflects the ability of the aptamer to fully equilibrate with respect to ligand binding prior to a regulatory decision being made (Wickiser et al., 2005a; Wickiser et al., 2005b). This likely reflects our choice to perform the assay at low NTP concentrations, which decreases the rate of transcription and increases the time the aptamer has to equilibrate with respect to ligand concentration. In contrast, the lysine riboswitch appears to be under weak “kinetic control,” which is defined by $T_{50}\gtrsim K_D$ (Wickiser et al., 2005a). This is the result of the aptamer having insufficient time to reach equilibrium prior to the RNAP reaching the intrinsic terminator and a regulatory decision being made.
Figure 3.3. Diverse chimeric riboswitches are functional. (a) Comparison of the in vitro transcription data quantified, plotted and fitted to a two-state transition as a function of ligand concentration. The *B. subtilis* metE riboswitch wild type (wt) in red circles, compared to different chimeras containing a different SAM-binding aptamer (*yitJ/metE*, cyan), a flavin mononucleotide (FMN) aptamer (*ribD/metE*, green), and a combination of these (*xpt C74U/metE*, blue).
green), a lysine binding aptamer (lysC/metE, pink), and the xpt (C74U) aptamer (blue). The values for $T_{50}$, dynamic range (DR) and % termination at saturating ligand concentrations are given in Table 3.1. In all experiments the riboswitch was titrated with the cognate effector of the aptamer. Error bars represent the standard deviation of at least three independent experiments. (b) Bar graph showing the percent termination at low (-) and high (+) ligand concentrations of the three riboswitches characterized and the 18 chimeras constructed. The dashed lines indicate the constructs with similar percent termination in the presence of ligand as the wild type riboswitches.

3.2.2. Chimeric riboswitches regulate transcription in vitro

The strategy for creating chimeric riboswitches is to splice a minimal aptamer domain into the metE, yitJ, and lysC expression platforms at the boundaries similar to those shown in Figure 3.1. For each aptamer, structural and biochemical analysis yielded sufficient information to determine the minimal sequence element required for ligand recognition. Each minimal aptamer includes all direct ligand-RNA contacts observed in the structure and all nucleotides involved in ligand-induced conformational changes as judged by in-line, RNase, or chemical probing. This sequence was substituted for the wild type aptamer in each expression platform such that the full switching sequence was preserved as well as the number of Watson-Crick base pairs of the P1 helix. It has been observed that the length of the P1 helix is important for efficient functioning of an adenine and a lysine riboswitch (Blouin et al., 2011; Lemay et al., 2011).

To determine whether expression platforms are capable of being modular, I created a fusion between the well-characterized B. subtilis xpt-pbuX guanine riboswitch aptamer domain(Mandal et al., 2003) and the metE expression platform. The resulting
chimera, called xpt/metE (aptamer domain/expression platform), showed no response to SAM (Figure 3.2c), but strongly regulates in the presence of guanine (Figure 3.2d) (the sequences of all riboswitches used in this work are listed in Appendix 2). The T50 of the guanine-responsive chimera correlates strongly with the K_D of the natural xpt aptamer for guanine under transcription conditions (0.03 and 0.024 nM, respectively; Table 1). Furthermore, a single point mutation (U51C) that is highly deleterious to guanine binding to the xpt aptamer (Gilbert et al., 2007) abrogates ligand-dependent regulation (data not shown), demonstrating that the chimera’s regulatory activity is due to the same specific interaction with guanine as the wild type riboswitch. Chimeras between the xpt aptamer domain and yitJ or lysC expression platform exhibit similar behavior (Table 3.1 and Figure 3.3b). These data, along with a fusion of the Streptococcus mutans tetrahydrofolate (THF) aptamer to the metE expression platform presented in a prior study (Trausch et al., 2011), demonstrate the modularity of some riboswitch expression platforms.

Assaying the regulatory activity of a diverse set of aptamer domain/expression platform chimeras reveals that the ability to reprogram regulation is both general and robust (Table 3.1 and Figure 3.3). In all, eighteen chimeras were created that used expression platforms of the metE, yitJ and lysC riboswitches in combination with aptamers from the guanine-sensing B. subtilis xpt riboswitch (Mandal et al., 2003), an xpt(C74U) aptamer that imparts adenine and 2-aminopurine (2AP) responsiveness (Gilbert et al., 2006a), and the B. subtilis flavin mononucleotide (FMN)-responsive ribD aptamer domain (Winkler et al., 2002b). For every chimera I was able to measure an effector-dependent regulatory response consistent with the aptamer’s affinity for the
effector, further supporting the generality of the mix-and-match strategy (examples of titration data for metE expression platform containing different aptamers are shown in Figure 3.3a).

Comparing the regulatory response of these chimeric RNAs to their wild type counterparts, I observed two general trends. First, there is a strong correlation between the $K_D$ of the isolated aptamer domain and the observed $T_{50}$ in the majority of riboswitches under in vitro transcription conditions reflecting the strong coupling of ligand binding to regulatory activity. Moderate deviation between these two values for some chimeras (metE/yitJ, for example) is likely the result of slow binding equilibration relative to the rate of transcription (Wickiser et al., 2005a; Wickiser et al., 2005b). Secondly, many of the chimeras are able to achieve maximal transcriptional termination levels ($\%T_{\text{max}}$, Table 3.1) that rival or exceed those of wild type riboswitches (red dashed lines, Figure 3.3b). Several aptamers appear to perform well, such as yitJ SAM-I aptamer that achieves $\geq 95\%$ transcriptional termination when coupled to the metE and lysC expression platforms—appear to perform as well as any of the wild type riboswitches. Conversely, several aptamers consistently underperform, such as the ribD FMN-responsive aptamer, which can only achieve 50-70% transcriptional termination at saturating FMN concentrations. This is consistent with observations of the wild type B. subtilis ribD riboswitch (Wickiser et al., 2005b). These observations indicate that the aptamer domain has a strong influence upon both $T_{50}$ and the maximal extent of transcriptional termination.

3.2.3. Synthetic aptamers can control the regulatory switch.
The above data reveal the modular nature of certain expression platforms whose regulatory switch can be directed by structurally diverse natural aptamers. However, these biological aptamers have presumably evolved in the context of a riboswitch and therefore may contain inconspicuous features that enable them to interface with the secondary structural switch of the expression platform. In fact, it has been proposed that the greater structural complexity of biological receptors as compared to their in vitro selected aptamer counterparts is important for their role in regulation (Sinha et al., 2011). Synthetic aptamers are selected in the absence of any pressure for regulatory function and thus may not be predisposed to interface with the expression platform. To determine whether biological RNA receptors have unique properties that enable them to direct secondary structural switching, I tested the theophylline (theo) (Jenison et al., 1994) and tetracycline (tet) (Berens et al., 2001) aptamers as chimeras with expression platforms.

As with biological aptamers, I observed that in vitro selected aptamers are capable of effector-dependent transcriptional regulation. Each of the three theophylline chimeras is functional to varying degrees in a single turnover transcription assay (Figure 3.4 and Table 3.1). While the theo/yitJ chimera showed a T_{50} close to K_D, the same aptamer in the context of metE and lysC showed considerably higher T_{50}s. Furthermore, the theo/metE and theo/yitJ chimeras showed a >90% repression at high theophylline concentrations, indicating that the in vitro selected aptamer can achieve a level of transcriptional termination rivaling that of wild type riboswitches. It should be reinforced that these chimeras were created without the need for a second selection
step to create an aptamer-specific expression platform as in other approaches (Lynch et al., 2007; Sinha et al., 2010; Soukup and Breaker, 1999a).

3.2.4. Chimeric riboswitches function in vivo.

The experiments shown in this section were performed by Joan Marcano-Velázquez.

While the above experiments reveal that artificial chimeras regulate transcriptional termination in vitro, these conditions are different from those encountered by natural riboswitches in the cell. Transcription conditions were chosen to match physiological ionic conditions as closely as possible including lowering the [Mg$^{2+}$], but the NTP concentrations are low (50 μM) in comparison to their intracellular concentration in rapidly growing bacteria (>1 mM) (Buckstein et al., 2008). Further, in the cell accessory transcription factors such as NusA influence the elongation rate (Grundy and Henkin, 2004; Yakhnin and Babitzke, 2002). Previous work has shown that the absence of NusA in the in vitro transcription assay can reduce riboswitch activity; for the B. subtilis FMN riboswitch this was approximately 2-fold (Wickiser et al., 2005b). Thus, in vitro activity of a chimera does not necessarily correlate with its performance in the cell.
To determine whether these artificial riboswitches function in a cellular environment, the ability of several chimeras to regulate expression of a *gfpuv* reporter gene was examined. We focused on two *metE* regulatory domain chimeras containing either the C74U variant of the *xpt* guanine riboswitch aptamer or the theophylline aptamer (Figure 3.5). These chimeras were chosen because *E. coli* does not rapidly
metabolize their effectors. Further, we used a strain of *E. coli* harboring a deletion of the purine efflux pump (*nep*) to prevent potential export of these molecules. As negative controls, a single point mutation that abrogates binding was introduced into the ligand binding pocket of each small molecule binding site (U51C for the *xpt* aptamer (Gilbert et al., 2007) and U24A for the theophylline aptamer (Zimmermann et al., 1997)). Both chimeric riboswitches display strong regulatory activity in *E. coli* (*Figure 3.6*) as indicated by reduction in fluorescence to a background level upon addition of 1 mM effector to a defined medium (*Figure 3.6a, c*). Despite the use of the strong constitutive promoter both chimeras achieve complete repression of reporter in comparison to background cellular fluorescence, underscoring the effectiveness of these regulatory RNA elements. Furthermore, aptamers containing a single point mutation preventing ligand binding fail to diminish *gfpuv* expression.

Titration of the effector into a defined liquid medium enabled measurement of an EC$_{50}$, defined as the concentration of ligand in the media that elicits half-maximal repression of reporter gene expression. For the *xpt*(C74U)/metE riboswitch an EC$_{50}$ of 730±80 µM was measured, with a 12-fold repression of expression (*Figure 3.6b*); this EC$_{50}$ is ~24-fold higher than the T$_{50}$ measured *in vitro* (*Table 3.1*). The theo/metE riboswitch displayed a response of 190±10 µM, similar to the observed T$_{50}$ (130±10 µM) and ~10-fold above the K$_D$ for the aptamer (20±6 µM) (*Table 3.1 and Figure 3.6d*). The observed EC$_{50}$ for this riboswitch is consistent with the performance of other theophylline-dependent synthetic riboswitches (Auslander et al., 2010; Desai and Gallivan, 2004; Feng et al., 2011; Ogawa, 2012; Suess et al., 2004; Win and Smolke, 2008). It must be noted that although I used a strain of *E. coli* deficient in purine efflux,
Figure 3.6. Chimeric riboswitches regulate gfpuv expression in E. coli. (a) 2-aminopurine (2AP) dependent regulation of gfpuv expression by the xpt(C74U)/metE chimera in E. coli. Cells transformed with either the control parental vector (pBR322) that does not contain gfpuv, the xpt(C74U)/metE riboswitch in the 5’-leader of an mRNA encoding gfpuv, and a mutant of this riboswitch (U51C) that specifically abrogates 2AP binding to the aptamer were plated on defined medium-agar in the absence (left) or presence (right) of 1 mM 2AP. pBR322 serves as a control for the background fluorescence of E. coli, while the mutant is a control for potential non-specific 2AP dependence of gfpuv regulation that is not directly tied to binding of this molecule to the xpt(C74U) aptamer of the chimera. (b) Quantification of the fluorescence of E. coli transformed with the xpt(C74U)/metE chimera (grey circles) or the mutant (U51C, black squares) grown in defined medium with increasing concentrations of 2AP. (c) Theophylline-dependent control of gfpuv expression in E. coli. Cells were grown under the same media conditions as in (a), but with a supplement of 1 mM theophylline in the right plate. A single point mutation in the theophylline aptamer (U24A) serves a control for potential theophylline-dependent effects on gfpuv expression that are not directly related to its binding to the aptamer. (d) Quantification of the fluorescence of E. coli transformed with the theo/metE chimera (grey circles) or the point mutation thereof (black squares).
the cells are still capable of purine degradation and thus the intracellular 2AP concentration does not necessarily tightly correlate with that in the medium, confounding making any conclusions regarding the lack of tight correlation between EC$_{50}$ and T$_{50}$. We also did not account for the small concentration of adenine in the cell (~1 µM (Bennett et al., 2009)), which further complicates this comparison. Nonetheless, these observations reinforce the validity of using the transcription assay to screen and optimize variants prior to introduction into a cell.

3.2.5. Optimizing riboswitch performance by altering P1 helix stability
Aspects of this section were originally discussed in section 2.2.5. for the metE riboswitch. However, the lysC and xpt/metE sequences are added.

It is not obvious why these expression platforms can be directed by a number of distinct aptamers. One hypothesis regarding co-transcriptional folding of alternative RNA secondary structures called “encoded co-transcriptional folding” (Xayaphoummine et al., 2007) may provide an explanation. This model proposes that a combination of the 5’-to-3’ sequential order and relative stability of secondary structural elements directs the folding outcome. A recent study of synthetic RNA switches demonstrated that accounting for the thermodynamics and kinetics of RNA folding proved to be sufficient for design of RNAs that fold almost exclusively into one of two mutually exclusive structures despite each being thermodynamically isoenergetic (Xayaphoummine et al., 2007). In application of this concept to transcriptional riboswitches, the competing P1 and antiterminator (P-AT) structures constitute the regulatory switch (Figure 3.1a), while the terminator (P-T) is merely a readout of the P1/P-AT outcome. If this model is correct, then the information required to direct RNA folding along one of two folding
trajectories is primarily encoded within the P1/P-AT switch of the expression platform. Any mechanism that changes the stability of the P1 helix relative to P-AT in response to a specific signal should therefore be able to act as an efficient effector for the riboswitch. This hypothesis is consistent with our observations that the regulatory switches of *B. subtilis* *metE*, *yitJ* and *lysC* expression platforms are not hard-wired to a specific receptor.

As a consequence of this model, the intrinsic stability of the P1 helix is a primary site for tuning the ligand-responsiveness of the secondary structural switch. A prior study demonstrated that weakening the P1 helix by removing base pairs forces the *lysC* riboswitch to become constitutively “ON”, whereas strengthening the helix by adding base pairs results in a constitutively “OFF” switch (Blouin et al., 2011). To more precisely define how helix length/stability directs the switch to fold into one of the two regulatory states, I introduced a series of point mutations in the wild-type *metE* and *lysC* riboswitches and the *xpt/metE* chimera at the 5'-terminus of their P1 helix (*metE* mutations shown in Figure 3.7a). These mutations alter the stability of the P1 helix while sequence elements required for receptor binding and conformational switching are unaltered. Both the *metE* and *xpt/metE* switches exhibit a sharp boundary between predominantly folding in the “OFF” (terminated) and “ON” (read through) states in the absence of ligand (Figure 3.7b). In the case of the *xpt/metE* riboswitch, this transition occurs over a single base pair (Figure 3.7c). In contrast, the wild type *lysC* riboswitch is more tolerant to mutations; lysine-dependent switching occurs at four different P1 lengths (Figure 3.8) in general agreement with a previous report (Blouin et al., 2011).
Figure 3.7. The length of the P1 helix can dictate the switch. (a) Variation of the length of the P1 helix was achieved by introducing mutations on the 5'-side of the helix that either removed Watson-Crick pairs (P1-1, P1-2) or added them (P1+1, P1+2). Variants are ordered from left to right as those expected to favor the antiterminator P-AT to those favoring formation of P1/P-T. (b) Separation of the RT and T products obtained in transcription reactions of the mutants represented (a) in a PAGE. In this case the aptamer used is the wild type receptor. The reactions are performed in the presence and absence of SAM, and the percentage of termination obtained after quantitation is presented underneath each reaction. (c) Identical representation as in (b), however, the aptamer used in this case is the guanine sensing B. subtilis xpt receptor fused to the metE regulatory domain. (d) Plot of the length of the P1 helix (number of successive Watson-Crick and G•U pairs) for the metE (red), lysC (green) and xpt/metE (blue) riboswitches as a function of percentage termination in the absence (solid lines) or presence (dashed lines) of ligand. Where the two values differ significantly (>20%) is the hallmark of ligand-dependent riboswitching.
The importance of P1 helix length variation in the design of new chimeras is illustrated in the tet/metE chimera (Figure 3.8). In the first design of this chimera, I took a conservative approach with the aptamer and retained three strong G-C base pairs in the aptamer portion of the P1 helix. This created a chimera whose P1 helix was twelve base pairs (Figure 3.8a) and showed very low tetracycline responsiveness (Figure 3.8b, c “wt”). The strong transcriptional termination in the absence of ligand was evidence for

Figure 3.8. P1 helix length can be used to optimize chimera performance. (a) The sequence of the P1 regulatory helix for a tetracycline aptamer (tet) fused to the metE expression platform. Mutations introduced into the 5'-side of the P1 helix to destabilize formation of the helix to favor formation of the competing antiterminator helix are shown below. (b) Transcription reactions at low and high tetracycline concentrations as a function of P1 helix length. Note that mutations that shorten the P1 helix by two or three base pairs enable a robust regulatory response. For each RNA, the T_{50} and percent termination at low and high effector concentrations are given as the average of three independent experiments. (c) Quantification of the regulatory response of the four chimeras shown in (b). The gel of a complete titration of tet/metE(P1-3) is shown in the inset of the graph.
an overly stable P1 helix. To design a better performing riboswitch, I introduced a set of mutations on the 5'-side of the P1 helix that do not disrupt ligand binding or P-AT formation. While the \textit{tet/metE} (P1-1) chimera still displayed a low dynamic range, the P1-2 and P1-3 switches showed substantial switching ability and lower T\textsubscript{50} (Figure 3.8b, c). Thus, modulation of the stability of the P1 helix is a simple and highly effective design strategy to optimize the activity of a new riboswitch. This approach is conceptually similar to the process for computationally designing riboswitches that terminate transcription by modulating the thermodynamic stability of P-T (Wachsmuth et al., 2012). However, in this case, I cannot calculate the strength of P1 because of the influence of the aptamer domain, and thus rely on an empirical approach.

3.3. Conclusion

A significant hurdle in the development novel RNA-based devices for synthetic biology is the ability to facilely connect sensory and readout domains with minimal experimental or computational design. Through a careful consideration of the sequence features of biological riboswitches, I have found three expression platforms that are capable of being directed by a diverse set of aptamers--both natural and synthetic. Since properties of co-transcriptional folding guide alternative secondary structure formation in these modules, altering the stability of one of the helices of the switch is an easy route to optimizing the performance of a designed chimera of aptamer and regulatory modules. This technique provides a simple route to "plug-and-play" readout modules that can be directly coupled to sensory modules and obtain practical RNA devices. We envision that the large repository of natural riboswitches can be prospected
for functional readouts in bacteria or eukarya to generate device libraries that regulate across a broad range of ligand concentrations, different levels of the information transfer process or in various combinations establish complex regulatory patterns.

3.4. Materials and Methods

3.4.1. Synthesis of DNA templates for transcription

The DNA templates used for single turnover transcription assays were synthesized by PCR amplification from *B. subtilis* genomic DNA in the case of the wild type sequences (listed in Appendix 2). These sequences were placed immediately downstream of the T7A1 promoter using recombinant PCR. Chimeric riboswitches were constructed in two different pieces. The first piece consisted of the T7A1 promoter, the 5'-side of the expression platform of choice (*metE, yitJ* or *lysC*), and the chosen aptamer domain. The second piece was composed of the 3'-side of the expression platform. Recombinant PCR was used to fuse these two portions. Mutant templates were obtained by altering the desired position/s in the DNA by standard site directed mutagenesis. The sequences of all DNAs were verified prior to use in transcription assays.

3.4.2. In vitro transcription assays

DNA templates containing a chimeric riboswitch were transcribed as previously described (Trausch et al., 2011). Briefly, 100 ng of DNA were incubated at 37 °C for 15 minutes in 25 µl of 2X transcription buffer (140 mM Tris-HCl, pH 8.0, 140 mM NaCl, 0.2 mM EDTA, 28 µM β-mercaptoethanol and 70 mg/mL BSA), 5 µl of 25 mM MgCl₂, 1 mCi of α-32P-ATP and 0.5 units of *E. coli* RNA polymerase 70 holoenzyme (Epicentre
Biotechnologies) to a final volume of 35 µl. The reactions were initiated with the addition of 15 µl of NTP mix (NTPs in equimolar ratios of 165 µM), 0.2 mg/ml heparin and the desired concentration of ligand. After incubation at 37 °C for 15 minutes the reactions were quenched with equal volume of 8 M urea and 2 minutes incubation at 65 °C. The species were separated by denaturing PAGE, dried and exposed to a phosphorimager screen. Quantitation of radioactive counts in each band was performed with ImageQuant software (Molecular Biosystems) and data fit to a two-state model with nonlinear least squares analysis.

3.4.3. In vivo assay

A reporter plasmid for testing riboswitch function in E. coli with a gfpuv reporter gene was constructed using pBR322 as the parental vector using standard molecular biological techniques. The map of the plasmid constituting the gfpuv reporter is shown in Figure 3.5. The sequence of the chimeric riboswitch insert for the xpt (C74U)/metE and theo/metE reporters are identical to those used for in vitro transcription assays. The resultant vectors were transformed into E. coli strain BW25113 (Δnep). Cells were transformed with pRR1 plasmid containing the different riboswitch chimeric constructs. Single colonies were picked and grown overnight in 3 ml of CSB media + 100 µg/ml ampicillin. This saturated culture was used to inoculate 100 ml of fresh media, and allowed to grow to early exponential phase (OD$_{600}$ = 0.1-0.5). Samples of these growths were taken in triplicate (3 ml each sample) and ligand was added to the media at the concentrations indicated in the titration. The cells were allowed to grow for 6 hours at 37 °C. At this point 300 µl of the cultures were used to measure the OD$_{600}$ and their
fluorescence intensity in a plate reader (Tecan). Fluorescent measurements were taken at an excitation wavelength of 395 nm and the average fluorescence was taken from 513-515 nm, where the maximum emission for GFP was observed. Optical density normalized fluorescent values were plotted as a function of ligand concentration and fitted to a two state binding equation to extract the EC$_{50}$ values. The background fluorescence was obtained by performing a ligand titration into cells carrying the parental plasmid pBR322 and calculating the OD$_{600}$ normalized fluorescence. An average of the background was subtracted to the values of the cultures containing the plasmids with the chimeric riboswitch constructs.
Table 3.1. Quantitation of *in vitro* transcription data and isothermal titration calorimetry.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>expression platform</th>
<th>B. sub metE</th>
<th>B. sub yitJ</th>
<th>B. sub lysC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>aptamer&lt;sup&gt;1&lt;/sup&gt;</td>
<td>K&lt;sub&gt;D&lt;/sub&gt;&lt;sup&gt;2&lt;/sup&gt; (µM)</td>
<td>T&lt;sub&gt;50&lt;/sub&gt; (µM)</td>
<td>D&lt;sub&gt;R&lt;/sub&gt;&lt;sup&gt;3&lt;/sup&gt;</td>
</tr>
<tr>
<td>SAM</td>
<td>metE</td>
<td>2.0 ± 0.5</td>
<td>2 ± 0.2</td>
<td>83</td>
</tr>
<tr>
<td>SAM</td>
<td>yitJ</td>
<td>1.5 ± 0.3</td>
<td>2.3 ± 0.5</td>
<td>38</td>
</tr>
<tr>
<td>Lysine</td>
<td>lysC</td>
<td>18 ± 2</td>
<td>160 ± 10</td>
<td>23</td>
</tr>
<tr>
<td>guanine</td>
<td>xpt</td>
<td>0.024 ± 0.003</td>
<td>0.03 ± 0.003</td>
<td>47</td>
</tr>
<tr>
<td>2-AP&lt;sup&gt;*&lt;/sup&gt;</td>
<td>xpt(C&lt;sub&gt;74U&lt;/sub&gt;)</td>
<td>8.8 ± 0.8</td>
<td>30 ± 5</td>
<td>34</td>
</tr>
<tr>
<td>FMN</td>
<td>nbD</td>
<td>0.91 ± 0.1</td>
<td>8.8 ± 1.9</td>
<td>46</td>
</tr>
<tr>
<td>theophylline</td>
<td>SELEX</td>
<td>20 ± 6</td>
<td>130 ± 10</td>
<td>59</td>
</tr>
</tbody>
</table>

<sup>1</sup>gene regulated by the riboswitch in *B. subtilis*, with the exception of the theophylline aptamer. All measurements were performed in triplicate with the error defined as the standard deviation obtained from three experiments.

<sup>2</sup>K<sub>D</sub> obtained through ITC matching conditions of *in vitro* transcription assay.

<sup>3</sup>DR is the dynamic range (%T<sub>max</sub> − %T<sub>min</sub>), where %T<sub>max</sub> is the maximum fraction terminated at high ligand concentration and %T<sub>min</sub> is the minimum fraction terminated at low ligand concentration.
Chapter 4

Engineering of inducible modular “ON” riboswitches using biological components

In the previous chapter, I demonstrated that expression platforms found in different riboswitches can direct the expression of a gene in response to a variety of small molecules. This was achieved by fusing different aptamers (natural or SELEX-derived) in the place of the native aptamer. Demonstrating that the secondary structural switch of ribo-regulators can function independently facilitates their usage as biosensors. However, to be useful for synthetic biology the engineered riboswitch should activate a genetic response. To achieve this goal I pursued a similar strategy to that presented in chapter 3. In this case, however, the re-engineering of modular expression platforms was performed to activate gene expression responding to small molecules. In this work, I use straightforward strategies to engineer two different “on” riboswitches to obtain modular expression platforms where different aptamers can be fused. RNA sequences following this design provide a scaffold capable of inducing gene expression in vitro and in E. coli responding to a variety of small molecules.

4.1. Introduction

The aim of synthetic biology is to engineer an organism in order to produce a behavior modification or chemical substance. Although some researchers have focused their efforts on engineering strategies that use different biological components, RNA-based regulatory devices have demonstrated a great deal of success. These RNA
sensors are capable of detecting chemical change *in vitro* and *in vivo*, and in turn function by performing a specific variation of the normal cell behavior (Topp et al., 2010; Wang et al., 2013; Weigand et al., 2012; Wieland and Hartig, 2008; Wittmann and Suess, 2012). Some of these RNA devices take advantage of the ability to develop aptamers through *in vitro* selection (Joyce, 2004; Tuerk and Gold, 1990). Regulation of the engineered RNAs is performed by a variation of secondary structure rearrangements that occur when the molecule is bound to the aptamer, which interferes or enhances the communication with cellular processes.

A great amount of progress has been observed in these bio-sensory modulators. The major problems preventing efficient and predictable engineering of new functions in a cell are the lack of well-characterized biological parts of reliable, independent and tunable character (Liu et al., 2012; Lucks et al., 2011; Lucks et al., 2008). The methodology of re-engineering a transcriptional control mechanism embedded in natural riboswitches satisfies all requirements previously stated. This is due to the ability of these expression platforms to function by inducing gene expression independently of what aptamer is used, and more importantly, the response obtained can be tuned by rationally altering the transcriptional controls built into the secondary structural switch. Furthermore, fusing them to different bio-sensory receptors (aptamers), allows for the construction of novel functional bio-regulators that control gene expression in a predictable manner.

Different biological components, such as promoters or proteins have been grouped as BioBricks. It is possible to hypothesize that RNA biological components that can act in combination can expand the repertoire of the well-characterized BioBricks.
The examples shown in chapter 3 added to the two new expression platforms that will be presented here form a group of 5 modules capable of altering gene expression reliably. In addition different aptamers (natural and SELEX-derived) have been fused to these expression platforms, further demonstrating that RNA can form domains of independent function.

Another approach to mimic riboswitch function was designed by Arkin and co-workers. The strategy consisted of rationally engineering non-coding RNAs to sense ligands by fusing them to synthetic aptamers. These two modules interacted through a loop-loop interaction. Upon ligand binding, this contact was disrupted and the non-coding RNAs are capable of carrying out their function. The mechanism of action is reminiscent of riboswitch function (Qi et al., 2012). Other groups have combined rational with computer design to create synthetic transcriptional riboswitches capable of \textit{in vitro} and \textit{in vivo} functionality (Wachsmuth et al., 2013). The combination of these two strategies represents a reduction in the screening process because of the conditions that the construct must meet according to the computer algorithm. Most of the requirements used allow selection of energetic differences of competing secondary structures (Wachsmuth et al., 2013). The implementation of these conditions limit the screening process in contrast with the elevated number of sequences that must be monitored when SELEX is used to produce a construct library. This has been done to obtain aptamers of specific function or for communication modules capable of interacting with the cellular machinery (Lynch et al., 2007). The screening process is costly in resources and time.
The methodology described here consists of engineering and reprogramming a natural “on” riboswitch from two different organisms. A great candidate is \textit{pbuE}, which up-regulates the expression of a purine base efflux pump in response to a high intracellular concentration of adenine (Johansen et al., 2003). As a proof of principle, and to demonstrate that this strategy is applicable to other riboswitches, a similar methodology was designed with another expression platform from the riboswitch \textit{metH} from the organism \textit{Dechloromonas aromatica}. This ribo-regulator senses the presence of \textit{S}-adenosylhomocysteine (SAH) in the cell. If the concentration of this metabolite reaches toxic levels, the riboswitch activates the expression of 5-methyltetrahydrofolate-homocysteine methyltransferase, which breaks down SAH. Both of these genes prevent the toxic effect that high levels of a metabolite can produce in a cell and therefore must be efficiently regulated.

Recently our laboratory has focused on developing RNA sensors capable of controlling the expression of genes in response to small molecules. The development and characterization of different chimeric riboswitches by mixing and matching natural and artificial components have demonstrated that the engineered RNA sequences are capable of controlling transcription in a single round \textit{in vitro} transcription assay, as well as in \textit{E. coli}. This illustrates the relative ease of design of these chimeras and their portability into bacterial cells (Chapter 3).

In this work, I show that different aptamer domains obtained from either natural sources or \textit{in vitro} selection can be fused into the expression platform of a riboswitch that has been slightly altered. The re-construction design presented in this work does not require a selection step to create a communication module specific to the
aptamer/expression platform chimera, but rather a rational variation of the sequence to optimize response range. These ribo-regulators are functional in a simple in vitro transcription assay and as a regulatory element of a gfpuv reporter in E. coli, indicating that this strategy is capable of generating a wide variety of functional regulatory devices. In contrast with the previously discussed methodologies, this approach uses independent modules capable of interfacing with the cell machinery (Sinha et al., 2010). These RNA devices can induce the cellular machinery to express genes when bound to their cognate ligand.

4.2. Results

4.2.1. A survey of “ON” riboswitches

To determine if fusing different aptamers to various expression platforms could be performed with an “on” riboswitch, many ribo-regulators that are capable of activating gene expression were examined. A few families of riboswitches such as the SAH, and the c-di-GMP-I, primarily use expression platforms that activate gene expression when bound to their cognate ligand. Most families of riboswitches present a few examples of “ON” switches such as glycine, lysine and adenine (Barrick and Breaker, 2007). However, I encountered a fundamental difference between “ON” and “OFF” riboswitches. The secondary structure of the “OFF” riboswitches previously used to create chimeric biosensors consists of a competition between P1 and P-AT, where the terminator forms as a consequence of P1 stabilization (Chapter 3). However the “ON” switches present a different architecture where the competing secondary structures are P1 and P-T. In addition, in “OFF” switches the two domains are separable whereas the
“ON” controllers present two domains whose sequences overlap. Therefore, I proposed a design where the riboswitch could be re-structured in such a way that the expression platform no longer invaded into its aptamer. The first choice was the adenine-sensing *pbuE* from *B. subtilis* (Figure 4.1).

### 4.2.2. The *pbuE* riboswitch regulates gene expression *in vitro* and *in vivo*

This section was previously discussed in section 2.2.9. Here a more in depth description is given.

This riboswitch has been shown to bind and function in response to adenine and its analog 2-aminopurine (2AP) in a single round *in vitro* transcription assay (Lemay et al., 2011; Mandal and Breaker, 2004). We were able to recapitulate these results with small differences probably due to variations during the experimental procedure (mainly Mg²⁺ and NTP concentrations). The T₅₀ obtained during the titration with 2AP was 9.9 µM, which agrees with a previously reported value, although the measurement was performed with a different ligand, 2,6 di-aminopurine (Lemay et al., 2011). During transcription the percentage of read-through transcript increases as a function of 2AP concentration, confirming the direct effect of the ligand on the disruption of the transcriptional terminator (Figure 4.1b). To ensure that this result is due to the RNA binding the small molecule, a riboswitch incorporating the mutation U74A, which impairs the normal binding interaction, was transcribed in the absence and presence of adenine and 2AP. This sequence variation yielded no difference in the level of termination whether the ligand was present or absent during the reaction (Figure 4.1c). These results indicate the inability of binding by the mutant RNA and that the wild type response is truly due to the interaction of the aptamer with 2AP.
In addition to these results, our data demonstrate that the response of the riboswitch is highly dependent on the start site. The *in vitro* response of the wild type riboswitch is poor, but it is enhanced when the initial 11 nucleotides are deleted (Figure 2.19). This effect is observed during the *in vivo* characterization of this riboswitch as a function of 2AP concentration in the medium. For this study the sequence of *pbuE* will lack the initial 11 nucleotides, which enhances the dynamic range of the response obtained *in vitro* and *in vivo* (Figure 4.1).
4.2.3. Design of chimeric riboswitches containing the *pbuE* expression platform

In the *pbuE* riboswitch, the receptor domain and the expression platform are not separable because of overlapping sequence requirements of the two
domains. A significant part of the sequence (16 nucleotides) that participates in the formation of the terminator also contributes to the formation of the aptamer (P3) and ligand recognition (Figure 4.2 sequence highlighted in gray). This sequence arrangement contrasts with the riboswitches in which a clear separation of the two modules can be observed (Chapter 3). This means that the previous design of chimeric riboswitches, which substituted different aptamers for the wild type receptor, could not be directly employed.

To redesign the pbuE expression platform to be modular I explored the synthetic regulators designed by this laboratory to date. They consist of the competition between two helices, P1 and P-AT, during the co-transcriptional folding of these RNAs. In the absence of ligand the most stable structure is the P-T that represses the transcription of the gene; the folding landscape can be altered to favor P1 by either the presence of the cellular metabolite or by increasing the relative stability (base pairs) of this helix (Chapter 3). However, in the case of pbuE, the most energetically favored and predicted structure in the absence of ligand is the formation of a stable terminator hairpin (P-T) (Lemay et al., 2006; Mandal and Breaker, 2004) (Figure 4.1a).

The strategy employed to redesign the pbuE expression platform consisted of permitting the formation of the stable terminator (P-T) and allowing it to compete against the formation of P1. In this case the P1 helical region is conceptually analogous to the P-AT of “OFF” switches discussed previously. The competition of the secondary structural switch can be established between the formation of P-T or P1 by altering the switching mechanism of the RNA. The sequence of the ribo-regulator was altered by engineering a “decoupler” sequence consisting of 6 nucleotides (Figure 4.2b sequence...
represented in cyan). This added region is complementary to the 5'-side of the terminator hairpin (**Figure 4.2b, P1 or P-AT**).

In the new chimeric sequence, the absence of ligand would allow the RNA to form the stable terminator, which corresponds to transcription disruption. However, in the presence of ligand, the aptamer would adopt its highly structured fold, which correlates to the formation of the helical region P1 and therefore eliminates the terminator stem-loop. The aptamers used were kept to the minimum sequences capable of binding the ligand as was performed in chapter 3.
Figure 4.2. Sequence and secondary structure of the pbuE riboswitch and the engineering design of pbuE/pbuE. The co-transcriptional folding of the pbuE riboswitch favors the formation of a stable terminator in the absence of ligand. The ligand binding (adenine or 2-aminopurine) stimulates a conformational change that disrupts the terminator in a. The strategy to utilize the pbuE expression platform as a scaffold for transcription control by a small molecule is highlighted in b. The decoupler sequence capable of forming a stable P1 or P-AT is represented in cyan. The pbuE aptamer is placed as other receptors are when creating chimeric riboswitches. In both panels the sequence shared by the two domains in pbuE is highlighted in gray.
4.2.4. Chimeric riboswitches using the expression platform of \textit{pbuE} yields functional sensors \textit{in vitro}

Fusing of aptamers into the \textit{pbuE} expression platform yielded a collection of chimeric riboswitches. Natural sequences from \textit{B. subtilis}, as well as the \textit{in vitro} selected aptamer for theophylline, were incorporated into the \textit{pbuE} expression platform. The choice of aptamers and the usage of minimal sequences required to bind the ligand, reflect that the design principles of the constructs necessitate familiarity with the binding characteristics of the aptamer. All the aptamers used in this section were shown to function independent of expression platform in the previous chapter.

After creating these chimeric switches (\textit{xpt} (74U)/\textit{pbuE}, \textit{pbuE/pbuE} and \textit{yitJ/pbuE}), an evaluation of their performance reveals several properties of their behavior. The initial level of read-through transcript is extremely low in all sequences (1-5\%) (\textbf{Figure 4.3} and \textbf{Table 4.1}), which reflects the high stability of the P-T helix and the efficiency of transcription termination. This level or repression slightly improves that of the \textit{pbuE} wild type switch, which is approximately 8\% read-through in the absence of ligand (\textbf{Figure 4.1}). This is perhaps one of the critical features and most encouraging properties of these engineered constructs. For different applications in synthetic biology, minimizing background expression is crucial. Unfortunately, the induction of transcription in the presence of saturating amounts of ligand is low (approximately 10\%) for the \textit{xpt} (C74U)/\textit{pbuE} and for the \textit{pbuE/pbuE} used in the chimeric system. The T\textsubscript{50}s obtained were relatively high (21 and 13 \textmu M respectively, see \textbf{Figure 4.3} and \textbf{Table 4.1}). It should be noted that although a numerical value was assigned during the quantitation, the dynamic range of the response is the lowest of the chimeric riboswitches characterized to date and
the errors in $T_{50}$ are the largest obtained. This indicates that these data are perhaps not the most reliable and should be taken lightly. Other natural aptamers such as ribD, xpt, or the SELEX derived, theophylline, fused to the $pbuE$ expression platform in the same manner, resulted in an equally deficient functionality, indicating that perhaps this is not an optimum strategy (Table 4.1). These data are in contrast with the functional $pbuE$ wild type riboswitch, which presented lower $T_{50}$ (9.9 $\mu$M) and significantly greater dynamic range (8-57% read through). These results illustrate that perhaps some aptamers are not well suited for this strategy, or that the $pbuE$ expression platform is only functional in the context of its natural aptamer.
In contrast, the natural SAM-sensing aptamer \textit{yitJ} from \textit{B. subtilis} fused to the engineered construct achieves a 50\% in the level of read-through during a SAM titration, and a $T_{50}$ of 1.2 $\mu$M was obtained, in agreement with the $T_{50}$ obtained for the \textit{yitJ} aptamer in other expression platforms (Chapter 3) (Figure 4.3 and Table 4.1). Previously, isothermal titration calorimetry measurements matching transcription conditions were performed for these aptamers. The results yielded $K_{D}$s of 1.5 $\mu$M for \textit{yitJ}, 8.8 $\mu$M for \textit{xpt} (C74U) and 5.6 $\mu$M for the \textit{pbuE} aptamer. The data obtained when fusing the \textit{yitJ} aptamer to the \textit{pbuE} expression platform demonstrated that this secondary structural switch can be directed by a non-native aptamer, and that the de-coupler strategy is indeed effective.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure4.3}
\caption{\textit{In vitro} characterization of the chimeric riboswitch strategy to couple aptamers to the \textit{pbuE} expression platform. \textit{a}, Graphical representation of quantified data obtained from \textit{in vitro} transcription titrations showing the different behavior of chimeric riboswitches with the cognate ligand. The insets are PAGE showing a full titration of \textit{yitJ/pbuE} with SAM, on the left, and transcriptions of \textit{yitJ/pbuE} mutant U78A in the presence and absence of SAM, on the right. \textit{b}, Graph illustrating a close up view of the \textit{pbuE/pbuE} (black), and \textit{xpt(C74U)/pbuE} (blue) titrations, both chimeric constructs do not show a significant increase in the \% read through transcript obtained.}
\end{figure}
when creating chimeric riboswitches. To ensure that the response obtained with this chimeric riboswitch is due to the binding of the aptamer to SAM, a critical nucleotide within the binding pocket was mutated to inhibit the binding event. This variation in sequence is identical to the one performed to *metE* in chapter 2 (in the case of the *yitJ* aptamer the change in nucleotide identity is U78A). This mutation prevents binding of SAM and therefore no switching is observed (Figure 4.3a inset). This demonstrates that the SAM-dependent anti-termination requires the ligand to specifically bind the *yitJ* aptamer domain.

### Table 4.1. *In vitro* data characterization of initial trials for creating chimeric riboswitches using the *pbuE* expression platform.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Ligand</th>
<th>Kᵤ, μM</th>
<th>T₅₀, μM</th>
<th>% RT -Ligand</th>
<th>% RT +Ligand</th>
<th>Dynamic Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>yitJ</em>/pbuE</td>
<td>SAM</td>
<td>1.5 ± 0.3</td>
<td>1.2 ± 0.2</td>
<td>1</td>
<td>49</td>
<td>48</td>
</tr>
<tr>
<td>xpt (C74U)</td>
<td>2AP</td>
<td>8.8 ± 0.8</td>
<td>22 ± 10</td>
<td>5</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>/pbuE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pbuE/pbuE</td>
<td>2AP</td>
<td>5.6 ± 0.1</td>
<td>13 ± 11</td>
<td>5</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>xpt/pbuE</td>
<td>guanine</td>
<td>0.024 ± 0.003</td>
<td>n.d.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ribD/pubE</td>
<td>FMN</td>
<td>0.91 ± 0.1</td>
<td>n.d.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>theo/pbuE</td>
<td>theophylline</td>
<td>20 ± 6</td>
<td>n.d.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 4.2.5. *pbuE* engineered RNAs response can be tuned

Evaluating the problem that the majority of chimeras designed were virtually non-functional suggested that further changes to these riboswitches were required to restore function. We hypothesized that these changes needed to be focused on the expression platform, as this domain is the response controller and additionally, the fusion of these aptamers to other secondary structural switches had proven successful (Chapter 3). The observation that all of these nonfunctional chimeras terminated transcription
regardless of effector concentration suggested that the P-T helix was too stable relative to the stability of P1. To improve read-through levels at high ligand concentration two options were considered. The first strategy was to increase the stability of the competing structure, which in this case is P1. This approach proved effective in the functional alteration of natural and chimeric riboswitches (Chapter 3). The second option consisted of systematically reducing the number of uridines composing the poly-uridine track, which would reduce the intrinsic pausing of the polymerase at this site and therefore allow the transcription reaction to continue (Nudler and Gottesman, 2002).

The first element of the chimeric riboswitches capable of influencing the co-transcriptional folding of these RNAs is the helix P1, which is in competition with the terminator helix. The increased stability of P1 was accomplished by adding base pairs to this helix, favoring its formation over other secondary structures (Chapter 3). This was shown to alter the ability of the secondary switch to function in the pbuE and lysine-sensing lysC riboswitches (Blouin et al., 2011; Lemay et al., 2011). However, in the case of the chimeric riboswitches created in this study, the addition of one base pair to P1 resulted in an improvement of dynamic range for the xpt (C74U) aptamer, but not in yitJ (Figure 4.4). Furthermore, the levels of read-through transcript in the absence of ligand were drastically changed from 1-5% with the original design to 10 or 45% with the additional base pair in P1, which makes the levels of gene regulation less than optimal for other possible applications (Figure 4.4). This indicates that the mutations added to P1 shift the secondary structural switch to favor the formation of this helical region. This conformation of the RNA, forming P1 and inducing gene transcription, was virtually
never sampled in the absence of ligand under the original strategy probably due to its low stability.

Figure 4.4. *In vitro* characterization of P1 mutants of *pbuE* expression platform. a, Representation of *pbuE* expression platform responsible for secondary structural switch, highlighting the mutation incorporated to stabilize P1 and adding one base pair (P1+1). b, Graphical representation of the quantified data obtained with *yitJ/pbuE*, which is represented in red, and *xpt* (C74U) in blue. The original data are presented with solid lines and the P1+1 mutants are fitted with dashed lines. The data plotted are the average of three independent experiments and the error bars represent the uncertainty in the measurement.

Since increasing the stability of P1 did not alter the levels of transcription read-through as was intended another set of mutations was designed. In this case, the new strategy consisted of systematically weakening the terminator element. Two possibilities were considered: the composition of the hairpin and the poly-uridine stretch. We decided to first focus on the poly-uridine tract, which constitutes the pause site identified by the RNA polymerase, immediately downstream of the terminator
hairpin. The enzyme will stall at the uridine-rich sequence, and if given enough pausing time, the terminator hairpin will form and cause the polymerase to disengage from the template DNA strand and stop transcription. Canonical pause sites are composed of 7-9 uridines (Nudler and Gottesman, 2002). In the case of \textit{pbuE}, and under the transcription conditions of the assay, the efficiency of the terminator does not allow a switch in secondary structure, as demonstrated in Figure 4.3. To test whether this pause site can be used to tune the efficiency of transcription termination, I systematically deleted uridines from this region of the RNA.

The results of the deletions of uridines to decrease pausing of the polymerase yielded greatly improved response of all the chimeric riboswitches. The \textit{yitJ/pbuE} response improved to 76 and 86\% dynamic range (from the original 50\% change) with the deletion of 1 and 2 uridines respectively. These mutations influenced the read-through levels obtained in the presence of high concentrations of SAM, but most importantly, the transcription in the absence of ligand remained unchanged (Figure 4.5 and Table 4.2). Therefore, these alterations in the intrinsic pause site of the RNA create riboswitches capable of virtually an-all-or none response during \textit{in vitro} transcription. Significantly, the $T_{50S}$ increased only slightly (1.6 $\mu$M, for 7 uridines and 2.1 $\mu$M for 6 uridines compared to 1.2 $\mu$M for the original construct), demonstrating that the properties of the aptamer were unaltered, and in agreement with other chimeric riboswitches containing the same aptamer (Chapter 3). Removing one more uridine (5 uridines) did not result in a gain of functionality, but rather a greater amount of the read-through transcript in the absence of ligand, comparable to that of the P1 stabilizing mutant (P1+1 presented in Figure 4.4). Importantly, this construct is not capable of fully
inducing gene expression (Figure 4.5). This illustrates the necessity of a stable terminator stem followed by a canonical pause site (7-9 uridines) when engineering an “ON” riboswitch.

Table 4.2. Data quantification of yitJ/pbuE and mutants located in the poly-uridine track.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Ligand</th>
<th>T$_{50}$, µM</th>
<th>% RT -Ligand</th>
<th>% RT +Ligand</th>
<th>Dynamic Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>yitJ/pbuE</td>
<td>SAM</td>
<td>1.2 ± 0.2</td>
<td>1</td>
<td>49</td>
<td>48</td>
</tr>
<tr>
<td>yitJ/pbuE 7 Us</td>
<td>SAM</td>
<td>1.6 ± 0.1</td>
<td>1</td>
<td>77</td>
<td>76</td>
</tr>
<tr>
<td>yitJ/pbuE 6 Us</td>
<td>SAM</td>
<td>2.1 ± 0.1</td>
<td>1</td>
<td>87</td>
<td>86</td>
</tr>
<tr>
<td>yitJ/pbuE 5 Us</td>
<td>SAM</td>
<td>2.9 ± 0.7</td>
<td>19</td>
<td>50</td>
<td>31</td>
</tr>
</tbody>
</table>
Implementing the deletion strategy in the poly-uridine stretch required the deletion of 3 residues for the xpt (C74U)/pbuE and pbuE/pbuE in order to create a riboswitch capable of inducing at least 40% change in regulation (Figure 4.6 and Table 4.3). These two constructs yielded a 46 and 49% read-through change, respectively, during the in vitro transcription assay when a high concentration of 2AP (1 mM) was present in the reaction. The T₅₀s obtained are 20 µM for xpt (C74U)/pbuE, in agreement with the fusion of the same aptamer to other expression platforms (Chaper 3), and 27 µM for pbuE/pbuE. As noted earlier, the deletion of the third uridine changed the level of read-through transcription in the absence of ligand from 1% obtained with the wild type terminator sequence to 10% obtained with the poly-uridine variants.

Figure 4.5. The response of yitJ/pbue chimeric riboswitch can be tuned by mutating the poly-uridine region. By deleting uridines that form the pause identified by the RNA polymerase, the response of the chimeric riboswitch improves its dynamic range, but presents similar T₅₀s. The data represents in vitro transcriptions of yitJ/pbue with different pause sites, 8 uridines (black), 7 uridines (blue), 6 uridines (red) and 5 uridines (green). The values plotted represent the average of three independent experiments with the uncertainty in the measurement shown as the error bars.
(Figure 4.6 and Table 4.3). Depending on the application for which these bio-sensory devices may be used the deletion of one, two or three uridines may be ideal. If the highest dynamic range is sought, then the deletion of 3 uridines is appropriate, whereas if the level of repression in the absence of ligand is crucial, then perhaps the removal of two uridines in the poly-uridine tract may be a better choice. The effect of these mutations on the levels of gene induction in vivo will be shown in a later section.

Table 4.3. xpt (C74U)/pbuE and pbuE/pbuE chimeric riboswitches response can be tuned by mutating the poly-uridine stretch in the expression platform.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Ligand</th>
<th>$T_{50}$, µM</th>
<th>% RT -Ligand</th>
<th>% RT +Ligand</th>
<th>Dynamic Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>xpt (C74U) aptamer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>xpt (C74U)/pbuE</td>
<td>2AP</td>
<td>22 ± 10</td>
<td>5</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>xpt (C74U)/pbuE 7 Us</td>
<td>2AP</td>
<td>12 ± 3</td>
<td>4</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td>xpt C74U/pbuE 6 Us</td>
<td>2AP</td>
<td>18 ± 3</td>
<td>5</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>xpt C74U/pbuE 5 Us</td>
<td>2AP</td>
<td>28 ± 2</td>
<td>8</td>
<td>55</td>
<td>47</td>
</tr>
<tr>
<td>pbuE aptamer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pbuE/pbuE</td>
<td>2AP</td>
<td>13 ± 11</td>
<td>5</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>pbuE/pbuE 7 Us</td>
<td>2AP</td>
<td>90 ± 32</td>
<td>5</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td>pbuE/pbuE 6 Us</td>
<td>2AP</td>
<td>24 ± 7.0</td>
<td>6</td>
<td>42</td>
<td>36</td>
</tr>
<tr>
<td>pbuE/pbuE 5 Us</td>
<td>2AP</td>
<td>20 ± 4</td>
<td>10</td>
<td>58</td>
<td>48</td>
</tr>
</tbody>
</table>

Other chimeras also improved functionality without drastically changing $T_{50}$s when these deletions were applied to the expression platform. The xpt/pbuE construct was capable of inducing a 58 and 70% increase in the level of read-through transcription obtained by removing one and two uridines. The $T_{50}$s obtained with these constructs are 39 and 41 nM, values similar to the ones attained with other chimeras with the same aptamer presented in chapter 3 (Figure 4.7 and Table 4.4). This illustrates that these engineered RNA molecules function in a regime similar to $K_D$ under
these transcription conditions as previously reported (Chapter 3). The ribD/pbuE construct was capable of inducing the read-through transcript in vitro by 50% when one uridine was deleted, but the removal of a second nucleotide in the poly-uridine tract only added an additional 5% read-through to total 55%, with T_{50}s of 0.9 and 1.23 \mu M respectively (Figure 4.7 and Table 4.4). Most importantly, although this aptamer is capable of controlling the transcription in vitro, it is clear that the ability of disrupting the intrinsic terminator at high concentrations of ligand is not as efficient as what was obtained by other aptamers (only a maximum 50 and 55% RT is obtained). This illustrates that not all aptamer/pbuE combinations behave similarly. However, the dynamic ranges of these chimeras are similar to the responses obtained with the three “OFF” biosensors. Taken together, these data demonstrate that the secondary switch

![Figure 4.6](image)

**Figure 4.6.** Different aptamers can direct the expression platform of pbuE. The a and b panels show data obtained during titrations of cognate ligands with the different constructs, which are labeled aptamer/pbuE containing 7 uridines (7 Us) in a and 6 uridines (6 Us) in b (top left corner of graph). Both graphs contain identical color labels: xpt C74U in green, pbuE in red, ribD in black, theophylline (theo) in blue, yitJ in cyan and xpt in pink. The data plotted are the average of at least of three independent experiments and the standard deviations of these measurements are represented by the error bars.
of \textit{pbuE} can be directed by different natural aptamers, which control the binding specificity (ligand) and thermodynamics ($T_{50}$). In contrast, the amplitude of the response (dynamic range) is largely controlled by the expression platform and its co-transcriptional folding mechanism, as shown by the variation in the level of read-through achieved with the mutations in the poly-uridine stretch (Chapter 3).

Table 4.4. Data quantification of \textit{in vitro} titrations of chimeric riboswitches utilizing the expression platform of \textit{pbuE}.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Ligand</th>
<th>$T_{50}$, $\mu$M</th>
<th>% RT -Ligand</th>
<th>% RT +Ligand</th>
<th>Dynamic Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 uridines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{yitJ/pbuE} 7 Us</td>
<td>SAM</td>
<td>1.6 ± 0.1</td>
<td>1</td>
<td>77</td>
<td>76</td>
</tr>
<tr>
<td>\textit{xpt/pbuE} 7 Us</td>
<td>guanine</td>
<td>0.039 ± 0.001</td>
<td>1</td>
<td>59</td>
<td>58</td>
</tr>
<tr>
<td>\textit{ribD/pbuE} 7 Us</td>
<td>FMN</td>
<td>0.91 ± 0.06</td>
<td>1</td>
<td>51</td>
<td>50</td>
</tr>
<tr>
<td>\textit{xpt C74U/pbuE} 7 Us</td>
<td>2AP</td>
<td>12 ± 3</td>
<td>4</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td>\textit{pbuE/pbuE} 7 Us</td>
<td>2AP</td>
<td>90 ± 32</td>
<td>5</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td>\textit{theo/pbuE} 7 Us</td>
<td>theophylline</td>
<td>1.2 ± 0.1</td>
<td>12</td>
<td>62</td>
<td>50</td>
</tr>
<tr>
<td>6 uridines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{yitJ/pbuE} 6 Us</td>
<td>SAM</td>
<td>2.1 ± 0.1</td>
<td>1</td>
<td>87</td>
<td>86</td>
</tr>
<tr>
<td>\textit{xpt/pbuE} 6 Us</td>
<td>guanine</td>
<td>0.041 ± 0.003</td>
<td>1</td>
<td>71</td>
<td>70</td>
</tr>
<tr>
<td>\textit{ribD/pbuE} 6 Us</td>
<td>FMN</td>
<td>1.2 ± 0.1</td>
<td>3</td>
<td>58</td>
<td>55</td>
</tr>
<tr>
<td>\textit{xpt C74U/pbuE} 6 Us</td>
<td>2AP</td>
<td>18 ± 3</td>
<td>5</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>\textit{pbuE/pbuE} 6 Us</td>
<td>2AP</td>
<td>24 ± 7.0</td>
<td>6</td>
<td>42</td>
<td>36</td>
</tr>
<tr>
<td>\textit{theo/pbuE} 6 Us</td>
<td>theophylline</td>
<td>7.5 ± 2.0</td>
<td>9</td>
<td>72</td>
<td>63</td>
</tr>
</tbody>
</table>
4.2.6. *pbuE* expression platform can be directed by a synthetic aptamer

To ensure that this strategy can be generalized to different applications in synthetic biology, the Batey laboratory has shown that SELEX-derived aptamers can induce a change in the secondary switch of an expression platform (Chapter 3). To further test this hypothesis, I fused the theophylline aptamer to the *pbuE* expression platform. These experiments resulted in a response of 50 and 62% induction of read-through product in transcription assays for the constructs containing the deletion of one or two uridines in the transcription pause. A differential factor observed in these experiments in comparison to switches composed of natural aptamers, is that the variation in the terminator composition significantly changed the value of $T_{50}$ from 1.2 to 7.5 $\mu$M (Figure 4.8 and Table 4.4). These $T_{50}$s represent an improvement from the response of other chimeras previously reported by 1 and 2 orders of magnitude (Chapter 3). The great discrepancy of these values when the theophylline aptamer is fused to different expression platforms may demonstrate the importance of having different modules that are capable of inducing or repressing gene expression as desired by the user.

4.2.7. Engineering of chimeric ribo-regulators using the SAH riboswitch

The data presented in this section were obtained by Jeremy Trausch.

Because the above design and the genetic control obtained may not satisfy all different future applications of this technique, another secondary structural switch was tested. The goal of this new strategy is to improve the control over the response obtained, the ability of using a different expression platform and the reduction of
construct screening. In this methodology we use the expression platform of an S-adenosylhomocysteine (SAH)-sensing riboswitch from the organism *Dechloromonas aromatica* (*D. aromatica*). This riboswitch controls the expression of 5-methyltetrahydrofolate-homocysteine methyltransferase (*metH*), capable of degrading SAH (Old et al., 1988). The crystal structure of the aptamer was solved in the Batey laboratory and showed a folded structure forming a pseudoknot (Edwards et al., 2010). This motif has not been previously employed by this laboratory in the design of synthetic riboswitches and its unique secondary structure requires an alternative engineering strategy not utilized to this point (**Figure 4.9**).

**Figure 4.7.** SELEX derived aptamers are capable of directing gene expression when fused to *pbuE*. **a**, graph of data obtained after quantification of the *in vitro* transcription titrations of theophylline/pbuE with their ligand. Values are average of at least three independent experiments and error bars are the variation in the measurement. **b**, PAGE of the titrations obtained, showing the raw data.
Figure 4.8. Sequence and secondary structure of the *metH* riboswitch and the engineering strategy to create chimeric riboregulators. a, The sequence and secondary structure of the switching mechanism of *sahH/metH* is represented in the absence (left) and presence (right) of SAH. The poly-uridine tract mutation is shown in a box in the structure on the left. This variation was kept for subsequent chimeric riboswitches. The structure on the right side shows a brown box, which illustrates the sequence that was discarded to produce a secondary structural switch capable of accommodating different aptamers. b, The strategy to fuse natural or artificial aptamers such as *xpt* or *xpt* (C74U), which infers different selectivity, is presented. Other aptamers were fused in identical manner.
The natural *D. aromatica* riboswitch was tested in a single round *in vitro* transcription assay yielding a poor response during transcription (Figure 4.10). Rather than re-engineering and optimizing this uncharacterized *metH* aptamer, the sequence from the aptamer domain of *sahH* from *Pseudomonas aeruginosa* (PAO1) was employed by fusing it to the *metH* expression platform (Edwards 2010). This choice ensured that the construct was able to bind the aptamer with high affinity (570 ± 0.01 nM) and selectivity, minimizing the error due to improper binding and optimization in this module of the chimera. This sequence was not capable of terminating transcription robustly in the absence of SAH with approximately 50% read through product obtained, where a value close to 0 is desirable (Figure 4.10).

Since this effect in transcription was probably not due to improper binding of the ligand, the deficient module must be the expression platform. After exploring the sequence and secondary structure of this riboswitch, we observed that the poly-U track, responsible for transcription termination, contained the insertion of two nucleotides (CG) that could reduce pause time of the RNAP and therefore not allowing efficient termination of the transcript (Figure 4.9) (Nudler and Gottesman, 2002; Nudler and Gusarov, 2003). Mutating the two residues to UU increased the poly-uridine tract to 9 consecutive uridines, which showed greater termination efficiency in the absence of SAH (25% read through). Meanwhile, the RNA was also capable to undergo the proper switching mechanism and yielded high levels of read through, transcript in the presence of SAH (89%) (Figure 4.10 and Table 4.5). These rationally designed mutations, which improve functionality, were maintained in the sequence of the expression platform and employed in all constructs shown in this study.
In previous studies the fusion site between the expression platforms and aptamer had been relatively straightforward. Fusion of the desired aptamer at the 3'-side of the P1 helix commonly resulted in a functional chimeric riboswitch (chapter 3). In the metH riboswitch a pseudoknot is formed between the aptamer and the 5'-side of the expression platform. Maintaining this pseudoknot while alternating aptamers would be overly challenging. For this reason, the new strategy consisted on using the switching sequence and mechanism of the metH expression platform, while discarding the nucleotides responsible for SAH recognition or formation of the binding pocket (Figure 4.9a brown box). Therefore, the switching sequence forms the 3'-side of the new P1 when the sensory domain binds the ligand (Figure 4.9b sequence highlighted in cyan). This design results in one of the shortest sequences and simplest secondary structural

Figure 4.9. Engineering of the sahH/metH chimeric riboswitch. Data representation obtained with different sequences to engineer an SAH chimeric riboswitch. metH/metH is represented in red, sahH/metH is plotted in blue and the mutation incorporating two uridines in the terminator in black.
switch of all chimeras produced to date and with notable similarity to the constructs engineered and presented in chapter 3.

The initial trials employing this design yielded a modest \textit{in vitro} transcription regulation. The response obtained when the \textit{xpt} aptamer was fused to the \textit{metH} expression platform varied from 30 to 66% read through in the absence and presence of guanine (\textbf{Figure 4.11}). However, following the engineering principles that have demonstrated the ability to favor the formation of one confirmation over its alternative, the number of base pairs that formed the P1 helix was reduced. This reduction was obtained by introducing mutations in the 5’-side of P1, as previously done (\textbf{Figure 4.11}). The incorporation of mutations of 3 nucleotides was necessary to obtain a chimeric riboswitch with a low level of read through transcript of 14% in the absence of guanine and 65% in its presence. These mutations, which were highly effective to produce a responsive sequence, were kept in the design of other chimeras.
4.2.7.1. Functionality of chimeric riboswitches created with the *metH* expression platform

Following these principles we obtained functional chimeric constructs that regulate in response to: i) guanine by using the *xpt*-*pbuX* aptamer, ii) 2AP by varying the pyrimidine at position 74 from a C to a U of the previous construct, and iii) tetracycline by using the *in vitro* selected aptamer (Berens et al., 2001). These
aptamers were fused to the *metH* expression platform, in addition to the SAH aptamer from *sahH* (Figure 4.9 and Table 4.5). These chimeric riboswitches yielded $T_{50}$s similar to the values obtained using the *pbuE* expression platform, as well as results observed in a previous study (Chapter 3). Once again these data demonstrate that the aptamer used in the chimeric construct controls specificity of ligand as well as binding thermodynamics of these biosensors. The dynamic range of the responses, however, exhibited slight differences, illustrating that the expression platform employed in these experiments controls gene expression differently than *pbuE* does (compare Figure 4.7, Figure 4.12).

![Figure 4.11](image.png)

**Figure 4.11.** *In vitro* characterization of chimeric riboswitches constructed with the expression platform of the *metH* riboswitch. a, Graphical representation of data obtained and fitted to a two state model during titrations with cognate ligands of the *xpt/metH* (blue), *sahH/metH* (black), tetracycline *tet/metH* (green), and *xpt* (C74U)/*metH* (red). b and c, are PAGE of the transcriptions of two chimeric riboswitches as function of cognate ligands *sahH/metH* titration with SAH in b and *xpt* (C74U)/*metH* titration with 2-aminopurine (2AP) in c.
Table 4.5. Data quantification of chimeric riboswitches created by fusing aptamers to the metH expression platform.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Ligand</th>
<th>T$_{50}$, µM</th>
<th>% RT -Ligand</th>
<th>% RT +Ligand</th>
<th>Dynamic Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>sahH/metH</td>
<td>SAH</td>
<td>0.10 ± 0.014</td>
<td>25</td>
<td>89</td>
<td>64</td>
</tr>
<tr>
<td>xpt/metH</td>
<td>guanine</td>
<td>0.020 ± 0.002</td>
<td>12</td>
<td>66</td>
<td>54</td>
</tr>
<tr>
<td>xpt (C74U)/metH</td>
<td>2AP</td>
<td>20 ± 2.3</td>
<td>11</td>
<td>56</td>
<td>45</td>
</tr>
<tr>
<td>tet/metH</td>
<td>tetracycline</td>
<td>0.260 ± 0.021</td>
<td>40</td>
<td>81</td>
<td>41</td>
</tr>
</tbody>
</table>

4.2.8. *In vivo* functional chimeras created with the expression platform from *pbuE* and *metH*

In order to identify whether these constructs are functional in *E. coli*, two chimeric riboswitches (*pbuE* and *xpt* (C74U)) were incorporated immediately upstream of the *gfpuv* gene in a vector previously described (Chapter 3). The change in expression of GFP was monitored as a function of the ligand concentration added to the medium (see Materials and Methods). We tested two different aptamers, *pbuE* and *xpt* (C74U) fused to the *pbuE* expression platform, in addition to the *pbuE* wild type riboswitch. These constructs yielded different levels of gene induction during 2AP titrations (Figure 4.11 and Table 4.6).
Figure 4.12. *In vivo* characterization of the *pbuE* riboswitch and the functional chimeras that allow gene expression. Titrations of 4 different riboswitches with varying concentrations of 2AP in the growth medium (see Materials and Methods) are represented. Each graph represents the GFP gene induction factor data quantified, fitted to a two state model and plotted as a function of 2AP. The color scheme for a,b and c panels is as follows: blue is 8 Us in the terminator poly-uridine track, 7 Us is represented in black, red corresponds to 6Us and green is 5Us. The pink data points represent the titrations performed with the non-binder constructs used as controls (U74A for the *pbuE* aptamer and U51C for the *xpt* (C74U) aptamer). In the d panel the same plot is obtained with the *xpt* (C74U) aptamer in *methH*. The original design is represented in black and the re-engineered variation of the construct is represented in red. The non-binder control is represented in cyan. The data presented is the average of at least 3 independent experiments and the standard deviation of the average value is represented as the error bars.
The induction level of GFP obtained varied when using these chimeric constructs, as did the EC$_{50}$ (ligand concentration added to the medium to reach half-maximal response). Of all the sequences employed, the pbuE wild type riboswitch is activated by the lowest amount of ligand illustrated by an EC$_{50}$ value of 38 µM. The induction factor is quantified to be greater than 9, after normalizing the fluorescence obtained. In contrast, all of the chimeric riboswitches were virtually non-functional when using the un-altered expression platform (Figure 4.11). To test whether the rational re-design of these constructs would result in an improvement of the functionality in vivo, I systematically deleted uridines residues from the poly-uridine track. The pbuE riboswitch containing 7 uridines presented no significant difference in either the EC$_{50}$ or the induction factor (Figure 4.11a and Table 4.6). However, the ability to control gene expression of all chimeric riboswitches was greatly improved after deletion of one uridine from the poly-uridine tract of the terminator (Figure 4.11b and c). The regulatory activity of these RNAs showed induction factors of 11 for the pbuE/pbuE and 7 for the xpt (C74U)/pbuE constructs. The removal of an additional uridine resulted in a reduction of functionality for the pbuE/pbuE with a fold induction of 8, while performing identical change in the xpt (C74U)/pbuE chimeric riboswitch resulted in a gene expression of 14-fold, significantly increasing the control function of these riboswitches (Figure 4.11b and c and Table 4.6).
Table 4.6. *In vivo* titration data of *pbuE*, mutants and chimeric riboswitches. The EC$_{50}$ is the concentration of ligand that displays half the maximum activity. The fluorescence obtained is normalized to the amount of cells measured.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Ligand</th>
<th>EC$_{50}$, µM</th>
<th>Induction Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>pbuE</em></td>
<td>2AP</td>
<td>38 ± 7</td>
<td>9.8</td>
</tr>
<tr>
<td><em>pbuE</em> 7U</td>
<td>2AP</td>
<td>46 ± 9</td>
<td>9.4</td>
</tr>
<tr>
<td>xpt (C74U)/<em>pbuE</em></td>
<td>2AP</td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
<tr>
<td>xpt (C74U)/<em>pbuE</em> 7 Us</td>
<td>2AP</td>
<td>220 ± 40</td>
<td>7</td>
</tr>
<tr>
<td>xpt (C74U)/<em>pbuE</em> 6 Us</td>
<td>2AP</td>
<td>150 ± 10</td>
<td>14.5</td>
</tr>
<tr>
<td>xpt (C74U)/<em>pbuE</em> 5 Us</td>
<td>2AP</td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
<tr>
<td><em>pbuE</em>/<em>pbuE</em></td>
<td>2AP</td>
<td>24 ± 5</td>
<td>4.5</td>
</tr>
<tr>
<td><em>pbuE</em>/<em>pbuE</em> 7 Us</td>
<td>2AP</td>
<td>57 ± 10</td>
<td>12</td>
</tr>
<tr>
<td><em>pbuE</em>/<em>pbuE</em> 6 Us</td>
<td>2AP</td>
<td>240 ± 70</td>
<td>8.0</td>
</tr>
<tr>
<td><em>pbuE</em>/<em>pbuE</em> 5 Us</td>
<td>2AP</td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
<tr>
<td>xpt (C74U)/<em>metH</em> “wt”</td>
<td>2AP</td>
<td>36 ± 17</td>
<td>2.0</td>
</tr>
<tr>
<td>xpt (C74U)/<em>metH</em> CTS</td>
<td>2AP</td>
<td>89 ± 10</td>
<td>12.5</td>
</tr>
</tbody>
</table>

In contrast, I found that removing a third uridine from the *pbuE* expression platform (reducing the number to 5 uridines) did not stimulate the desirable response, but resulted in a lost of function probably due to the leaky expression in the absence of ligand (*Figure 4.11*). The pattern observed in the induction factor was not followed in the values obtained for EC$_{50}$s. The *pbuE* riboswitch presented similar EC$_{50}$s when the terminator was composed of 8 or 7 uridines (38 or 46 µM respectively), while the EC$_{50}$ of *pbuE*/*pbuE* containing 7 uridines was 57 µM and increased considerably (240 µM) when the sequence was changed to 6 uridines following the terminator hairpin. The EC$_{50}$s for *xpt* (C74U)/*pbuE constructs* were similar, 220 µM for the terminator containing
7 uridines and 150 μM for the terminator containing 6 uridines (Figure 4.11). These data demonstrate that controlling transcription with chimeric riboswitches responding to small molecules can be utilized to control gene expression in *E. coli* and may be utilized for different applications in biology and synthetic biology.

The construct *xpt* (C74U)/*metH* was a perfect candidate to test the utility of the new strategy to create a chimeric biosensor in *E. coli*. The chimeric riboswitch was cloned directly upstream of the *gfpuv* gene, as described above. The results obtained show that after 10 hours of growth in the presence of 2AP, the induction of GFP was approximately 2-fold, which was judged as only moderately successful in comparison with the performance of other chimeric riboswitches. The EC$_{50}$ obtained was 35 μM, in significant contrast with previously reported values for the same aptamer fused to the *metE* expression platform, where the EC$_{50}$ was 730 μM (Chapter 3). In these data, I noted that the riboswitch showed substantial expression in the absence of 2AP. We hypothesized that by altering the stability and/or folding mechanism of the terminator, this leaky expression could be suppressed. The stem of the terminator was capped with a stable and fast-folding UUCG tetra loop (Figure 4.11). This mutation was introduced into a region that does not take part in the switching mechanism. The results yielded an improved induction factor of approximately 10-fold, demonstrating again that rational design variation of a functional chimera can be used to develop biosensors of desired response. The EC$_{50}$ of the re-engineered construct is 89 μM, versus the initial 35 μM (Figure 4.11 and Table 4.6).
4.3. Discussion

In this work I have further demonstrated the modular character of aptamers, natural and artificial, and expression platforms from two different riboswitches. The possibility of joining them to create cellular RNA-based genetic controls increases the real-world applications for riboswitches. The two different strategies presented here, as well as those previously reported (Chapter 3), illustrate that synthetic biology may be able to utilize naturally evolved “building blocks”, as well as SELEX-derived aptamers. This fusion of independently functional modules allows monitoring a signal and the induction of a desired genetic response. The sensing building block can be an aptamer (natural or artificial), that can detect virtually any molecule. The reaction to the binding of the ligand of choice may be performed by the expression platform of a riboswitch, which can activate or repress a desired gene. The two models presented here are of special interest because inducing gene expression is highly desirable when building cellular circuits. These sensors present the ability to tune their response and achieve the desired level of control by rationally changing a transcription pause and enhance the switching ability.

We focused our efforts in transcriptional regulators for different reasons. For example, the ease of developing new aptamers through \textit{in vitro} selection methods, their well-characterized affinity, specificity and independence of function enables the design of biosensors responsive to virtually any molecule. The fact that transcriptional riboswitch control is very efficient \textit{in vitro} facilitates the quick assessment of the function of the new chimeric regulators. Different studies have demonstrated that transcriptional control dependent upon rho-independent terminator formation is even more robust \textit{in}
vivo (Nudler and Gottesman, 2002; Nudler and Gusarov, 2003), which indicates that our approach is likely to be transferable to different bacteria. It is not difficult to envision a tandem riboswitch (a combination of two different riboswitches to ensure proper response) functioning in vivo and creating different logic mechanisms, which have been created before and are present in nature (Sudarsan et al., 2006). Another factor that favors ribo-sensor development is the fact that RNA folding has been studied in depth and the bi-stability that RNA is capable of achieving is known. Therefore, strategies that can vary the outcome response are easily achieved. All these reasons make transcriptional RNA chimeric riboswitches an outstanding strategy for control of gene expression for any synthetic biology application. In addition, the different manners in which I have manipulated the RNA make excellent tools for varying the response of the regulation to match the conditions desired by the user.

In the presented strategies for the design of chimeric riboswitches I sought to implement, as effectively as possible, engineering design principles that have been applied in the past towards other synthetic riboswitch systems (Chapter 3). As pointed earlier in the chapter, the approaches use biological parts to engineer efficient and predictable cellular circuits such as reliability, independence and tunability. These qualities were recently identified as the major condition for efficient, predictable engineering of new function in cells (Lucks et al., 2008). The chimeric riboswitches created in this study meet these criteria.

Reliability, or the application of a sensory platform that can accommodate different ligand-sensory modules, has been demonstrated through the ability of two different expression platforms to respond to a variety of natural and artificial aptamers in
a ligand-dependent fashion. Portability, the independence of the riboswitch to cell-specific machinery or regulatory mechanisms, has been established. The ability of expression platforms derived from B. subtilis and D. aromatica to direct E. coli RNAP both in vitro and in vivo signifies that these biosensors can probably be used across bacteria. Since bacterial RNA polymerases in general respond to intrinsic rho-independent terminators found in these expression platforms, the expectation is that these chimeras will function in a broad spectrum of bacteria. Utility, or the coupling of the regulatory device to a functional output, is evidenced by their ability to control expression of gfpuv in vivo. In our design, the chimeras could be used to control expression of any conceivable downstream gene to elicit a desired output. Composibility, or a strategy to modularly couple aptamers and expression platforms is a way that preserves the activity of each domain. This is clearly demonstrated through the ability to vary genetic response by the variation of the pause site composition. Importantly, combining aptamers and expression platforms does not require selection or extensive screening of “communication modules” that are highly specific for the aptamer/expression platform pair. In general, a means of standardizing the communication between the liganded state of the aptamer to the expression platform, is the greatest strength of this strategy for creating novel riboswitches.

Another strength of these strategies for creation of artificial riboswitches is flexibility such that it can be extended to other forms of regulation or regulatory responses. For example, there are many expression platforms that control translation by exposing or occluding the ribosome binding site (RBS) using a similar secondary structural switch as described above. Since translational regulation has been shown to
be a viable means of generating cellular biosensors in other contexts, this may be a preferred regulatory mechanism for some applications. Furthermore, in synthetic biology, the opportunity for a combination of different levels of gene control (transcription and translation) may be desirable. This may allow a total control over the presence or absence of a gene product.

The possible uses for “ON” switches in different applications are considered to be extremely desirable. The levels of induction achieved with these ribo-regulators are acceptable and comparable to other strategies employed to date. Finally, expression platforms found in fungi and plants could be adapted to this strategy to create novel regulatory devices for use in eukaryotic cells. The vast number and diversity of riboswitches and their associated expression platforms that have been identified provides a substantial reservoir of potential candidates for even more robust regulatory switches than the few characterized in this study. This combined with the advances in the SELEX process to obtain aptamers capable of binding almost any molecule illustrates the endless assortment of possibilities for these RNA devices.

4.4. Materials and methods

4.4.1. DNA template construction

The DNA templates encoding riboswitches used for the transcription assays were obtained by standard PCR amplification from *B. subtilis* genomic DNA in the case of the wild type sequence. The T7A1 promoter was placed immediately upstream of the templates using standard recombinant PCR methodology. Chimeric riboswitches were constructed by overlapping oligonucleotides composing the aptameric region, and then recombinantly jointed to the expression platform of *pbuE*, which was amplified from *B.*
subtilis genomic DNA. Mutation or deletion of the sequence was obtained by standard site directed mutagenesis. The templates were sequenced prior to use in transcription assays to ensure verification.

4.4.2. In vitro transcription assays

DNA templates encoding wild type and chimeric riboswitches were transcribed in a single turnover assay. E. coli RNA polymerase holoenzyme (0.25 units) (Epicentre Biotechnologies) was incubated with 50 ng of DNA at 37 °C for 10 minutes in 12.5 µL of 2X transcription buffer (140 mM Tris-HCl, pH 8.0, 140 mM NaCl, 0.2 mM EDTA, 28 mM β-mercaptoethanol and 70 mg/mL BSA), 2.5 µL of 25 mM MgCl₂, 0.5 mCi of α-³²P-ATP in a final volume of 17.5 µL. The reactions were initiated with the addition of 7.5 µL of NTP mix (165 µM each), 0.2 mg/mL heparin and the ligand of choice at the desired concentration. After transcription was incubated for 10 minutes at 37 °C, the reactions were quenched with 25 µL of 8 M urea containing traces of bromophenol blue and xyanel xynol and incubated for 2 minutes at 65 °C. The transcripts were separated by denaturing PAGE, dried and exposed to a phosphorimager screen. Quantitation of radioactive intensity in each band was performed with ImageQuant software (Molecular Biosystems). The data obtained were fit to a two-state model with nonlinear least squares analysis.

4.4.3. In vivo assay

A reporter plasmid for testing riboswitch function in E. coli with a gfpuv reporter gene was constructed using pBR322 as the parental vector using standard molecular biological techniques. The sequence of the chimeric riboswitches inserted into the plasmid between the two restriction enzyme sites for the pbuE/pbue and xpt
(C74U)/pbuE reporters and mutants are identical to those used for *in vitro* transcription assays ([Appendix 3](#)). The resultant vectors were transformed into *E. coli* strain BW25113 (Δnep). Single colonies were picked and grown overnight in 3 mL of CSB media + 100 µg/ml ampicillin. This saturated culture was used to inoculate 100 mL of fresh media, and allowed to grow to early exponential phase (OD$_{600}$ = 0.1-0.5). Samples of were taken in triplicate (3 mL each sample) and ligand was added to the media at the concentrations indicated in the titration graphs. The cells were allowed to grow for 6 hours at 37 °C. At this point 300 µL of the cultures were used to measure the OD$_{600}$ of their fluorescent intensity in a plate reader (Tecan). Fluorescent measurements were taken at an excitation wavelength of 395 nm and the average fluorescence was taken from 513-515 nm, where the maximum emission for GFP was observed. Optical density normalized fluorescent values were plotted as a function of the ligand concentration where a curve was generated and fitted to a two state binding equation to extract the EC$_{50}$ values. The background fluorescence was obtained by performing a ligand titration into cells carrying the parental plasmid pBR322 and calculating the OD$_{600}$ normalized fluorescence. An average of the background was subtracted to the values of the cultures containing the plasmids with the riboswitch constructs.
Chapter 5

5.1. Dissertation Summary

The work presented in this thesis illustrates that by fully characterizing the functionally of a riboswitch, several details about the mechanism of action of these riboregulators have been demonstrated. Most of these intricacies are discrete in nature (do not involve a great number of base pairs), but relevant in function. Most importantly, these interactions may have been overlooked in the past. These small paired regions and their function represent a major and novel effort in building a complete description of the genetic regulation performed by riboswitches to date. Most important, these variations in function are probably not specific to a particular regulator, but probably to many sequences as exemplified by the two model systems characterized in this work (metE and pbuE).

The studies performed in this dissertation characterizing the role of the expression platforms of two riboswitches contrast with the study of these regulatory elements to date. Many research laboratories have characterized the binding event of a ligand to an aptamer domain using different techniques (Delfosse et al., 2010; Heppell and Lafontaine, 2008; Heppell et al., 2009; Jose et al., 2001; Lemay et al., 2006; Mandal et al., 2004). These studies have illustrated important aspects of ligand recognition by a natural aptamer. The discoveries have revealed the plasticity of RNA and its ability to recognize essentially every functional group of the ligand, which in turn facilitates the discrimination of other resembling metabolites. It is remarkable that most, if not all, of these studies have utilized the aptamer in isolation and that its function is
entirely retained independently of the expression platform. It is also noteworthy that the
ligands bound by natural aptamers present such diversity (from small anions such as F\(^-\),
to coenzymes of greater size such as B12). The specificity of the aptamers to
recognize their cognate ligand and differentiate compounds of similar composition must
arrive from high evolutionary pressures, which are probably most evidenced by the high
degree of conservation of these domains.

However, the expression platforms of riboswitches have been mostly ignored for
different reasons. The first reason is the lack of conservation patterns through
phylogeny (except the rho-independent terminator). The second factor, and as a
consequence of the first reason, is the possibility to perform co-variation studies does
not exist. Third, the lack of predicted structures prevents the study of function, as many
times the structure of a molecule may provide an explanation of functionality. The only
available structural information for different expression platforms is the secondary
structural switch and the terminator formation, which is often not validated and merely a
computational prediction. Finally, for transcriptional riboswitches the mode of regulation
is highly dependent on the co-transcriptional folding of the RNA and increases the
difficulty of this study. The solution to this difficulty is increasingly becoming simpler as
new predicting software methods are available. One of these methods, Kinefold,
demonstrated its usefulness during the study of the \textit{metE} riboswitch from \textit{B. subtilis}
(Chapter 2).

Although the characterization of an expression platform from a riboswitch may
require the relative difficulty mentioned, the few examples available predict that small
changes in this domain are capable of greatly influencing the genetic control of a
riboswitch (Blouin et al., 2011; Sherman et al., 2012). The work presented in chapter 2 is an example of how secondary structural switches may tune the activity of these ribo-regulators independent of the binding characteristics of the aptamer domain to the cellular metabolite. There exists no report that experimentally examines the secondary structural switch of a riboswitch. The functional analysis of the metE sequence yielded novel results. The interaction between the 5’ and 3’-side of the expression platform, forming a helix termed P1b, was shown to reduce the dynamic range of the riboswitch. The full switching ability of the ribo-regulator was obtained when different sequence mutations were used to impair the stable formation of the P1b helix (Chapter 2). Perhaps the most convincing demonstration of this interaction is the compensatory mutation in the middle of this helix. Incorporating mutations in the 5’- or 3’- sequence responsible for the formation of this interaction resulted in an increased dynamic range of riboswitch activity, while performing these variations simultaneously rescued the original activity of the wild type.

The higher dynamic range of the riboswitch was also demonstrated by the formation of P5 helical region in the expression platform. It has been hypothesized before that the formation of a short stem loop outside of the aptamer domain may assist in the formation of either the terminator or anti-terminator (Blouin et al., 2011). In the case of metE, the P5 stabilization performed by adding one, two or more base pairs was functionally important to regain full response range of the riboswitch. These sequence variations were another avenue to demonstrate that the formation of P1b was deleterious for functionality; however, the formation of P5 was not directly demonstrated (see Future Directions section).
The small dynamic range of *metE* and the fact that it is due to an interaction occurring between the two sequences flanking the aptamer domain is a novel discovery, and may be a general feature of riboswitches. To identify the generality of this deleterious interaction, another riboswitch was analyzed in the same manner. Studies of the *pbuE* riboswitch demonstrated that the performance of the full sequence was very poor and that the construct lacking the initial 11 nucleotides (Δ11) was capable of an increased range of gene regulation. This suggests that the activity of riboswitches may be altered by the initially transcribed sequences and the possible interactions formed. Therefore, to fully characterize the function of a ribo-regulator, the original transcription start site must be employed. This inhibitory effect is also observed *in vivo*, where the removal of the initial sequences enhances the expression of a reporter gene by either the *metE* or *pbuE* riboswitch.

These examples point towards the importance of the full characterization of different secondary structural switching elements of these RNAs. In this study, the focus has been several transcriptional regulators. However, the number could be extended further to investigate other controllers that act in the same manner or even translational riboswitches. Performing this analysis could assist in the identification of common elements in RNA-based regulation. The ability of RNA to adopt different structures that correlate to a variation in function is crucial for regulation by riboswitches and other nc-RNAs. The investigation of the switching mechanisms of other RNAs could further expand our understanding of bi-stable functionality. In addition, the function of the sequence transcribed initially may demonstrate if the effect observed in *metE* and in *pbuE* is general in RNA regulatory processes.
Taken together, these data may change the view of the ability of riboswitches to regulate the expression of genes. As pointed in chapter 2, perhaps the optimum mechanism to regulate cellular homeostasis is to steadily induce or repress the transcription of genes, and only varies the control slightly following ligand binding to the riboswitch. This is in contrast to the originally, proposed control that riboswitches were capable, which consisted of a mechanism allowing expression of approximately 100% of all transcripts and repressing this value to nearly 0% in the presence of the ligand for an “OFF” switch. The data presented in this work must be understood in their original context. During in vitro transcription assays, the RNA polymerase is allowed to catalyze a single turnover event, which does not entirely recapitulate what occurs in a cell. However, the in vivo data collected further illustrated that the behavior observed in vitro is representative of the genetic control of the riboswitch in the cell. These interactions may have been acquired through evolution and represent a valuable advantage for the genetic economy of a cell.

In chapters 3 and 4 an in-depth demonstration of the independence and reliable function of expression platform modules from five different riboswitches was presented. Approximately fifty chimeric ribo-regulators, recognizing a variety of ligands were created and their functionality evaluated. This is the first demonstration that these domains contained within riboswitches can be independent functional modules. The ability to adopt two functionally different structures is intrinsically encoded in the sequence of an expression platform. This highlights the importance of characterizing the regulatory domains of riboswitches as they may be applicable for re-engineering new cellular pathways.
Some common features in the behavior of the chimeric riboswitches are observed. First, the aptamer domain determines the affinity and specificity. These receptors are capable of discriminating between the cognate ligand and similar molecules present in the cellular environment. Second, the aptamer domain in a chimeric construct behaves similarly as in its original riboswitch. This is illustrated by the different mutations performed that abrogate binding (U51C in the case of xpt, U74A in the pbuE sequence, or U78A in the yitJ aptamer). Third, under the in vitro transcription conditions chosen the $T_{50}$s obtained correlate with the $K_D$ measured, indicating that these riboswitches function under a thermodynamic control and/or that the transcription reaction proceeds at speeds allowing this kind of control. Fourth, the amplitude of the response of the chimeric riboswitch is dependent on the expression platform. This module may be tunable by controlling the competing structures that determine its function (P1 versus P-AT in “OFF” switches and P1 versus P-T in “ON” regulators). The data collected indicate that in the case of transcriptional “OFF” switches, the major control element is the stability of the P1 helix. Over-stabilizing this region by adding base pairs, favors the formation of this helical region and termination of transcription is observed. In the contrary, de-stabilizing the helix by removing base pairs favors the formation of the anti-terminator and thus gene transcription. For “ON” switches the P1 stabilization rational was not as effective. However, the alterations in the transcriptional pause responsible for termination (systematically removing uridines forming the poly-U repeat) were efficient in altering the response of the chimeric riboswitches. It is worth noting that the strategies developed to vary the response of these RNAs demonstrate similar effects in vitro (with all of the caveats of this method).
as they do in vivo. The fact that these combinations of modules are functional in a cell demonstrates their general applicability and efficiency of engineering principles, which may enable employment of these RNA devices in synthetic biology. The terminator composition recognized by RNA polymerases in bacteria is similar throughout this domain of life. Therefore these chimeric riboswitches could be implemented in different organisms. This is demonstrated by the fact that riboswitch sequences of B. subtilis are capable of inducing or repressing gene expression in E. coli. The successful usage of SELEX-derived aptamers to control gene expression also ensures the utility of these RNA-based devices.

5.2. Future directions

Even though the completion of these studies has efficiently demonstrated novel attributes of riboswitches uncharacterized to date, further investigation may complement these findings. The fact that a natural riboswitch may have evolved to inhibit its dynamic range is biologically relevant (Chapter 2). This discovery deserves attention and the demonstration of whether the phenomenon is of general character in riboswitches or specific to a sub-set of them.

5.2.1. Investigation of different and varied expression platforms from riboswitches

To further test the hypothesis that these ribo-regulators may control gene expression with a small dynamic range of response a greater number of expression platforms need characterization. This can be accomplished employing different techniques. However, the amount of resources that are necessary to characterize a
representative number of these sequences may require a great amount of labor. To solve this problem, it is necessary to use and develop new bio-informatic tools to guide these efforts. When these algorithms identify possible RNA sequences that may form interactions detrimental for full riboswitch activity, then the \textit{in vitro} and \textit{in vivo} characterization may take place. In the present study, the employment of a software package such as Kinefold, was proven successful. However, due to the discrete nature of some of these interactions, a more reliable and tested software may be beneficial.

This new tool may be important as a new predictor of structure. It should be focused on the RNA molecules that function in a co-transcriptional manner, and the possibility to adopt different conformations that may be energetically similar. The method should consider bi-stability of RNA and the implications for function of a particular transcript. This new algorithm may be important for structure prediction of small functional RNA controllers, but also could be instrumental in the early stages of investigating long nc-RNAs function and their involvement in human diseases (Clark and Mattick, 2011). These long nc-RNA sequences present the added difficulties of not being as well represented in different species and their extensive length, which are greater than the average riboswitch. A reliable method capable of predicting secondary structures of any RNA in different organism may be important for developing testable methods of function.

5.2.2. \textbf{Chemical probing of the metE riboswitch can demonstrate the formation of P1b helix.}

The data collected in chapter 2 investigated the functionality of the \textit{metE} riboswitch. This study illustrated the interaction of the two sides of the expression
platform flanking the aptamer domain. This small helix is responsible for the repression of anti-terminator. Chemical probing is a method that interrogates the RNA flexibility and monitors the formation of secondary and tertiary interactions. This technique may directly demonstrate the formation of P1b in the metE riboswitch and perhaps, when the appropriate mutations are incorporated into the sequence, the formation of P5. There exist certain technical difficulties in the design of this experiment that can be overcome. Some of these issues are centered on the differences in the transcription reactions used to observe the effect of co-transcriptional folding in comparison to the standard transcription of RNA employed for chemical probing.

*E. coli* RNA polymerase has been employed for the *in vitro* transcription assays. However, T7 RNA polymerase is usually employed to quickly and reliably transcribe the RNA that will be subsequently chemically probed. This enzyme has been demonstrated to ignore transcription pause sites and therefore does not respond to terminator hairpins. Therefore the usage of T7 polymerases containing mutations that slow down the rate of transcription, and therefore are more susceptible to transcription termination, is necessary. These mutants should recognize the terminator structure and stop the RNA synthesis allowing for identical co-transcriptional folding pattern observed in previous reactions. This methodology will allow to visualize the reactivity pattern of the helical region (P1b) of the wild type sequence, which can be compared to the probing pattern obtained with the mutants that abolish its formation. This should be done in the future to further demonstrate the stabilization of this helix.

5.2.3. Possible application of the novel bio-sensory devices
Perhaps one of the most significant discoveries demonstrated in this dissertation is the independence of activity exhibited by expression platforms derived from natural riboswitches (Chapter 2 and 3). This is a novel discovery and merits the attention of the synthetic biology field that searches the development of new tools for implementing new functions in cells (Keasling, 2010). The different strategies illustrated in this dissertation used to induce or repress the expression of genes by re-engineering riboswitches, may be employed in virtually any bacterial cell and in combination with any aptamer. This will allow to highly impacting the normal cellular environment in response to virtually any molecule.

A possible application where these bio-sensory devices may become helpful is in the production of biofuels. During this process, the production of undesired intermediates may interfere with the final product concentration, as well as with the cell fitness. The engineered riboswitches could signal the production of an enzyme after the binding of an inhibiting species to a previously developed aptamer. The enzyme then, could be responsible for breaking down the dangerous species.

It can also be easily envisioned that the combination of these RNA devices may result in the construction of logic gates such as AND, OR or NOR. These combinations of chimeric biosensors produced to date, could couple the induction or repression of genes in response to the same or different molecules. Thus, the presence of a small molecule could turn on or off different genes, or a combination of them.

One more application where the fusion of an aptamer to an expression platform can be beneficial has been illustrated in the Batey laboratory (Trausch et al., 2011). This strategy can be employed to characterize aptamer domains from riboswitches that
remain unresponsive during *in vitro* transcription assays despite our efforts. This signifies that not all riboswitches behave identically and that some may need the assistance of different protein factors important for transcriptional regulation such as NusA. In the case of the tetrahydrofolate (THF) riboswitch, the wild type sequence rendered no alteration of transcriptional termination as a function of ligand concentration. To measure the regulatory activity induced by two different ligands capable of binding the aptamer with similar affinity monitored by ITC, the THF aptamer was fused to the *metE* expression platform. This module allows easily monitoring of gene regulation *in vitro*. The results illustrated that the ability to bind with similar affinity does not correlate to a similar control in gene expression. This was again a novel result obtained with this strategy, and introduced a new consideration for the development of small molecule binders to RNA molecules (Trausch et al., 2011).

Another example where this strategy has proven its efficiency is in the comparison of different aptamers, within the same family, by fusing them to the *metE* expression platform. In this case the influence of the differences present in the expression platforms of these genes, which may have evolved different genetic control mechanisms, are eliminated. Thus, the change in regulation obtained from these aptamer/*metE* chimeric riboswitches is due solely to the sequence alterations found in the aptamer. These sensory domains vary in sequence only slightly, but demonstrated different ligand affinity in a previous report (Mulhbacher and Lafontaine, 2007). Furthermore, these differences may be responsible for tuning the response of gene regulation by the purine riboswitches (Stoddard et al., 2013).
Taken together, the data presented in this dissertation represent novel discoveries about the regulation of riboswitches. These findings are significant, new to the field, and perhaps even unexpected. It is important to point out that although the assay used to quickly determine the response of a riboswitch is an *in vitro* technique, and may not exactly recapitulate the cellular environment, the behavior of the majority of the sequences tested *in vivo* mimic the response obtained *in vitro*. Perhaps this reflects the importance of allowing the RNA to fold as its synthesis proceeds. This co-transcriptional folding ability of RNA may be determinant of function.
REFERENCES


Blouin, S., Lafontaine, D.A., 2007. A loop loop interaction and a K-turn motif located in the lysine aptamer domain are important for the riboswitch gene regulation control. RNA 13, 1256-1267.


Grissa, I., Vergnaud, G., Pourcel, C., 2007. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC bioinformatics 8, 172.


### Appendix

Appendix 1. Sequences of the *metE* riboswitch and mutants utilized to understand the switching mechanism of a riboswitch. The expression platform is represented in white, the aptamer is highlighted in yellow, the translation start site is represented in red and specific mutations are highlighted in pink.

<table>
<thead>
<tr>
<th>Construct name</th>
<th>sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Wild type</strong></td>
<td>AAATGATTCAACTTTTCAAAAAATTAATAACATTTTCTTTATCGAGATTTGGGGCGAGGATTTGGGGCGAGGGATTGGCCTTTTGACCACCACAGCAACCAGCAGCTGTTACCATTTGACATTTTGACTGTTACGATTTGTCTGACGAGATGAGAGAGGCAGTGTTTTACGTAGAAAAGCCTCTTTCTCTCATGGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGTATCCGCGGGTGAAAGAGAGTGTTTTACATATAAAGGAGGAGAAACATG</td>
</tr>
<tr>
<td><strong>Δ16</strong></td>
<td>CAAAATAATAACATTTTCTTTATCGAGAGTTGGGGCGAGGGATTGGCCTTTTGACCCCAACAGCAACCAGCAGCTGTTACCATTTGACATTTTGACTGTTACGATTTGTCTGACGAGATGAGAGAGGCAGTGTTTTACGTAGAAAAGCCTCTTTCTCTCATGGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGTATCCGCGGGTGAAAGAGAGTGTTTTACATATAAAGGAGGAGAAACATG</td>
</tr>
<tr>
<td><strong>Δ24</strong></td>
<td>TAAATAACATTTTCTTTATCGAGAGTTGGGGCGAGGGATTGGCCTTTTGACCCCAACAGCAACCAGCAGCTGTTACCATTTGACATTTTGACTGTTACGATTTGTCTGACGAGATGAGAGAGGCAGTGTTTTACGTAGAAAAGCCTCTTTCTCTCATGGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGTATCCGCGGGTGAAAGAGAGTGTTTTACATATAAAGGAGGAGAAACATG</td>
</tr>
<tr>
<td><strong>Δ30</strong></td>
<td>ACATTTTCTTTATCGAGAGTTGGGGCGAGGGATTGGCCTTTTGACCACCACAGCAACCAGCAGCTGTTACCATTTGACATTTTGACTGTTACGATTTGTCTGACGAGATGAGAGAGGCAGTGTTTTACGTAGAAAAGCCTCTTTCTCTCATGGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGTATCCGCGGGTGAAAGAGAGTGTTTTACATATAAAGGAGGAGAAACATG</td>
</tr>
<tr>
<td><strong>Δ35</strong></td>
<td>CTTTATCGAGAGTTGGGGCGAGGGATTGGCCTTTTGACCACCACAGCAACCAGCAGCTGTTACCATTTGACATTTTGACTGTTACGATTTGTCTGACGAGATGAGAGAGGCAGTGTTTTACGTAGAAAAGCCTCTTTCTCTCATGGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGTATCCGCGGGTGAAAGAGAGTGTTTTACATATAAAGGAGGAGAAACATG</td>
</tr>
<tr>
<td><strong>T13-16A</strong></td>
<td>AAATGATTCAACAAAAAATAATAACATTTTCTTTATCGAGAGTTGGGGCGAGGGATTGGCCTTTTGACCACCACAGCAACCAGCAGCTGTTACCATTTGACATTTTGACTGTTACGATTTGTCTGACGAGATGAGAGAGGCAGTGTTTTACGTAGAAAAGCCTCTTTCTCTCATGGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGTATCCGCGGGTGAAAGAGAGTGTTTTACATATAAAGGAGGAGAAACATG</td>
</tr>
<tr>
<td></td>
<td>Sequence</td>
</tr>
<tr>
<td>-----</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>T14A</td>
<td>AAATGATTTCAACTATTTCAAAAAATTTAATAACATTCTTTATCTCTCGAGAGCTGCTTTCTGCTTCTGTCGAGAAGCATGGGCGGAGGATTGGCCTTTTGACCCCAACAGCAACCGGACCCTGATGTCTGCTATAAGGCACGGTGCTAATTCCATCAGGAGAGGTGGAAGGCCGAGTGTTTTACGTAGAAAAGCCTCTTTCTCTCATGGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGTATC</td>
</tr>
<tr>
<td>T14-5A</td>
<td>AAATGATTTCAACTATTTCAAAAAATTTAATAACATTCTTTATCTCTCGAGAGCTGCTTTCTGCTTCTGTCGAGAAGCATGGGCGGAGGATTGGCCTTTTGACCCCAACAGCAACCGGACCCTGATGTCTGCTATAAGGCACGGTGCTAATTCCATCAGGAGAGGTGGAAGGCCGAGTGTTTTACGTAGAAAAGCCTCTTTCTCTCATGGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGTATC</td>
</tr>
<tr>
<td>T14C</td>
<td>AAATGATTTCAACTATTTCAAAAAATTTAATAACATTCTTTATCTCTCGAGAGCTGCTTTCTGCTTCTGTCGAGAAGCATGGGCGGAGGATTGGCCTTTTGACCCCAACAGCAACCGGACCCTGATGTCTGCTATAAGGCACGGTGCTAATTCCATCAGGAGAGGTGGAAGGCCGAGTGTTTTACGTAGAAAAGCCTCTTTCTCTCATGGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGTATC</td>
</tr>
<tr>
<td>A215G</td>
<td>AAATGATTTCAACTTTTCAAAAATTAAATAACATTCTCTTATCTCTCGAGAGCTGCTTTCTGCTTCTGTCGAGAAGCATGGGCGGAGGATTGGCCTTTTGACCCCAACAGCAACCGGACCCTGATGTCTGCTATAAGGCACGGTGCTAATTCCATCAGGAGAGGTGGAAGGCCGAGTGTTTTACGTAGAAAAGCCTCTTTCTCTCATGGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGTATC</td>
</tr>
<tr>
<td>T14C/ A215G</td>
<td>AAATGATTTCAACTTTTCAAAAATTAAATAACATTCTCTTATCTCTCGAGAGCTGCTTTCTGCTTCTGTCGAGAAGCATGGGCGGAGGATTGGCCTTTTGACCCCAACAGCAACCGGACCCTGATGTCTGCTATAAGGCACGGTGCTAATTCCATCAGGAGAGGTGGAAGGCCGAGTGTTTTACGTAGAAAAGCCTCTTTCTCTCATGGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGTATC</td>
</tr>
<tr>
<td>P5 formation</td>
<td>AAATGATTTCAACTTTTCAAAAATTAAATAACATTCTTTATCTCTCGAGAGCTGCTTTCTGCTTCTGTCGAGAAGCATGGGCGGAGGATTGGCCTTTTGACCCCAACAGCAACCGGACCCTGATGTCTGCTATAAGGCACGGTGCTAATTCCATCAGGAGAGGTGGAAGGCCGAGTGTTTTACGTAGAAAAGCCTCTTTCTCTCATGGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGTATC</td>
</tr>
<tr>
<td>Protein Name</td>
<td>DNA Sequence</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------</td>
</tr>
</tbody>
</table>
| T200C        | GGGAAAGAGGCTTTTTTGTGAGAAAACCTCTTTAGCAGGCTGTTAAT
               | ATGGTCGAAAGGATTGCCCTTTGACCCCAACAGCAACGACGGCT
               | GATGAGAGAGGGTTTACGTAGAAAAGACCTCCTTTTCGCTCAGT
               | GCAAGGAGGCTTTTGGTGAGAAACCTCTTTGACCGCTGTAC
               | CCGGGGCTGAAGAGGAGGTGTTTACATATAAAGGAGGAACAAATG |
| T211C        | AAATGATTCAACTTTTTCAAAATTAATAACATTTCCTTTATCGAGA
               | GGGGGAGGATTGCGCTTTTGACCCCAACAGCAACGACGGCT
               | GATGAGAGAGGGTTTACGTAGAAAAGACCTCCTTTTCGCTCAGT
               | GCAAGGAGGCTTTTGGTGAGAAACCTCTTTGACCGCTGTAC
               | CCGGGGCTGAAGAGGAGGTGTTTACATATAAAGGAGGAACAAATG |
| G200T        | AAATGATTCAACTTTTTCAAAATTAATAACATTTCCTTTATCGAGA
               | GGGGGAGGATTGCGCTTTTGACCCCAACAGCAACGACGGCT
               | GATGAGAGAGGGTTTACGTAGAAAAGACCTCCTTTTCGCTCAGT
               | GCAAGGAGGCTTTTGGTGAGAAACCTCTTTGACCGCTGTAC
               | CCGGGGCTGAAGAGGAGGTGTTTACATATAAAGGAGGAACAAATG |
| AA213-4      | AAATGATTCAACTTTTTCAAAATTAATAACATTTCCTTTATCGAGA
               | GGGGGAGGATTGCGCTTTTGACCCCAACAGCAACGACGGCT
               | GATGAGAGAGGGTTTACGTAGAAAAGACCTCCTTTTCGCTCAGT
               | GCAAGGAGGCTTTTGGTGAGAAACCTCTTTGACCGCTGTAC
               | CCGGGGCTGAAGAGGAGGTGTTTACATATAAAGGAGGAACAAATG |
| Δ213-4       | AAATGATTCAACTTTTTCAAAATTAATAACATTTCCTTTATCGAGA
               | GGGGGAGGATTGCGCTTTTGACCCCAACAGCAACGACGGCT
               | GATGAGAGAGGGTTTACGTAGAAAAGACCTCCTTTTCGCTCAGT
               | GCAAGGAGGCTTTTGGTGAGAAACCTCTTTGACCGCTGTAC
               | CCGGGGCTGAAGAGGAGGTGTTTACATATAAAGGAGGAACAAATG |
| A213-6C      | AAATGATTCAACTTTTTCAAAATTAATAACATTTCCTTTATCGAGA
               | GGGGGAGGATTGCGCTTTTGACCCCAACAGCAACGACGGCT
               | GATGAGAGAGGGTTTACGTAGAAAAGACCTCCTTTTCGCTCAGT
               | GCAAGGAGGCTTTTGGTGAGAAACCTCTTTGACCGCTGTAC
<pre><code>           | CCGGGGCTGAAGAGGAGGTGTTTACATATAAAGGAGGAACAAATG |
</code></pre>
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGCGGGTGAAAAGAGAGATGTATTTTACATATAAAGGAGGAGAAACAAATG</td>
<td>A213-6T</td>
</tr>
<tr>
<td>AAATGATTCAACTTTTTCAAAAAATTAATAACATTTTCTCTATTCGAGA</td>
<td>Δ213-6</td>
</tr>
<tr>
<td>GTTGGGCGAGGGATGGGCTTTTGACCACAAGCAACCGGAGCGT</td>
<td>Δ211-16</td>
</tr>
<tr>
<td>AAATGATTCAACTTTTTTCAAAAAAATTAATAACATTTTCTCTATTCGAGA</td>
<td>P2aΔ1bp</td>
</tr>
<tr>
<td>GTTGGGCGAGGGATGGGCTTTTGACCGCAAAGCAAACCGGACCGTAA</td>
<td>P2aΔ2bp</td>
</tr>
<tr>
<td>AACATTTTGTGGAATGCGCTTTTGACCGCAAAGCAAACCGGACCGTAA</td>
<td>KT GTA</td>
</tr>
<tr>
<td>AAATGATTCAACTTTTTTCAAAAAAATTAATAACATTTTCTCTATTCGAGA</td>
<td>KT AAT</td>
</tr>
<tr>
<td>GTTGGGCGAGGGATGGGCTTTTGACCGCAAAGCAAACCGGACCGTAA</td>
<td></td>
</tr>
<tr>
<td>AACATTTTGTGGAATGCGCTTTTGACCGCAAAGCAAACCGGACCGTAA</td>
<td></td>
</tr>
<tr>
<td>AAATGATTCAACTTTTTTCAAAAAAATTAATAACATTTTCTCTATTCGAGA</td>
<td></td>
</tr>
<tr>
<td>GTTGGGCGAGGGATGGGCTTTTGACCGCAAAGCAAACCGGACCGTAA</td>
<td></td>
</tr>
<tr>
<td>AACATTTTGTGGAATGCGCTTTTGACCGCAAAGCAAACCGGACCGTAA</td>
<td></td>
</tr>
<tr>
<td>Position</td>
<td>Mutation</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>P2b T68C</td>
<td>AAATGATTCAACCTTTTCAAATAAACATTTTTCTTTATCGAGA</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>P2a bp</td>
<td>AAATGATTCAACCTTTTCAAATAAACATTTTTCTTTATCGAGA</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>T157A</td>
<td>AAATGATTCAACCTTTTCAAATAAACATTTTTCTTTATCGAGA</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>T168A</td>
<td>AAATGATTCAACCTTTTCAAATAAACATTTTTCTTTATCGAGA</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 2. Sequences of wild type riboswitches unresponsive during *in vitro* transcription assay. All the sequences are riboswitches found in *B. subtilis* responsive to guanine. Expression platform sequences are represented in gray, aptamer domain in cyan, and translation start site in red.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>B. subtilis yxjA</strong></td>
<td>ATCATCTTAGAAAAAGACATTCTTTGTATATGATCGAATATATGGTC TGGATTGTTTCTACCTAGTAAACCGTAACAAAAACTAGACTA CAAAGAAAGTTTGAAATTTGAAGGACAAGTTTCTTGTT TGCTCTTTATTTCACACTTTTGCACTTTCCAAGAATTGGGAAGGA TAAGAGCTTTTTTTTTCCATAATAACCCCTCATAGGAGTTGCTAT CAGTC</td>
</tr>
<tr>
<td><strong>B. subtilis purE</strong></td>
<td>ATGAAAAA CAACAAACAGACCTCATATAATCTTTGGGAATATGGCCCAT ATAGTTTCTACCCCGGCAACCGTAACATTTCGGGAACTATGCAAGAAAG TGTATCATGAAAATCTGACATGATATACGCAGAAGCGAAGCAGACTG ACGATACATGTACCCTGGCGTTTTGGATTTGCTCTCTCATAGTG TCATGCGACAGCGGTATTTTTTATTTTTCTGAAAAACAAAGCATTAG AAGGTGGGGAACAGGAAATC</td>
</tr>
<tr>
<td><strong>B. subtilis xpt</strong></td>
<td>TTGCTAGGAACACACTCATATAATCGCGTGGAATATGGGACAGTTT TCTACCCCGGCAACCGTAACATTTCGGGAACACATTGCGCAGAAATGGATATATGGTCGAAGCGAAGCAGACTGACGATACATGTACCATGCCCGGTTTGTATTGCTTCCTCATAAGTGCAATGCAGAGCGGGTATTTTTTATTTTCGTAATACAAAAGCATTAGAAGGTGGGGAACAGAATG</td>
</tr>
<tr>
<td><strong>B. subtilis pbuG</strong></td>
<td>ATAGACACAGAAATCAAATAAGATGAATTCCGATAATCTGCTACCAGCGCTATAACCTGTAGTGTTGTGTTGATATATGGACAGAAATGGATATATGGTCGAAGCGAAGCAGACTGACGATACATGTACCATGCCCGGTTTGTATTGCTTCCTCATAAGTGCAATGCAGAGCGGGTATTTTTTATTTTGTGTAATACAAAAGCATTAGAAGGTGGGGAACAGAATG</td>
</tr>
<tr>
<td><strong>B. subtilis yebG</strong></td>
<td>GTTGTTAAGATAAACATGAAATCAAACACGACCTCATATAATCTTT GGGAATATGGCCCATATAAATTTCTACCCCGGCAACCGTAACATTTCGG GACTATGCAAGGAAATGATGATGACTAAACTGACATGATATATGCG AGAAGCGAACGACGTGACCATTGACCCTGGCGTTTTGTATTGCTTCCTCATAGTGTGCAATGCGAGACGGGTATTTTTTATTTTTCTGAAAAACAAAGCATTGAGATGCGGGAACAGGAAATC</td>
</tr>
</tbody>
</table>
Appendix 3. Sequences of natural or chimeric riboswitches presented in Chapter 3. The aptamer is presented in green, expression platform in cyan, the starting translation site in red, and the mutations introduced in pink. The “wild type” `metE` sequence used in this paper starts at the +15 nucleotide of the predicted start site of the `metE` transcript.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>lysC expression platform chimeras</strong></td>
<td></td>
</tr>
<tr>
<td><strong>lysC wild type</strong></td>
<td>AATTTCATAGTTAGATCGTGTTATATGGTGAAAGATAAGGTGC GAACTTCAAGAGATGCCTTTGGAGAAAGATGGATCTTGGA AAAAGGCTGAAGGGGAGCGTCGCGAAGCAAATAAACC CATTGGATATTTGCGCTGCGATTGAATAATTGATAATG AGATA GAGGTGC</td>
</tr>
<tr>
<td><strong>lysC (P1-3)</strong></td>
<td>AATTTCATAGTTAGATCGTGTTATATGGTGATCTTATAGAGGTGC GAACTTCAAGAGATGCCTTTGGAGAAAGATGGATG AATTGGCGAAGGAGCTTTGGAGAAAGATGGATTCTGGA AAAAGGCTGAAGGGGAGCGTCGCGAAGCAAATAAACC CATTGGATATTTGCGCTGCGATTGAATAATTGATAATG AGATA GAGGTGC</td>
</tr>
<tr>
<td><strong>lysC (P1-2)</strong></td>
<td>AATTTCATAGTTAGATCGTGTTATATGGTGATCTAAGATAGGTC GAACTTCAAGAGATGCCTTTGGAGAAAGATGGATG AATTGGCGAAGGAGCTTTGGAGAAAGATGGATTCTGGA AAAAGGCTGAAGGGGAGCGTCGCGAAGCAAATAAACC CATTGGATATTTGCGCTGCGATTGAATAATTGATAATG AGATA GAGGTGC</td>
</tr>
<tr>
<td><strong>lysC (P1-1)</strong></td>
<td>AATTTCATAGTTAGATCGTGTTATATGGTGATCTAAGATAGGTC GAACTTCAAGAGATGCCTTTGGAGAAAGATGGATG AATTGGCGAAGGAGCTTTGGAGAAAGATGGATTCTGGA AAAAGGCTGAAGGGGAGCGTCGCGAAGCAAATAAACC CATTGGATATTTGCGCTGCGATTGAATAATTGATAATG AGATA GAGGTGC</td>
</tr>
<tr>
<td>lysC (P1+1)</td>
<td>AATTTCATAGTTAGATCGTGTTATATGGTGCAGATAAGGTTGCAACCTTCAAAGATGATGTTTCTTGAGAAAGATGGAAATGGGTACCATCTGGATTAGAATTGCTGGCTGCCGAAGCAAAATATACGCCCATCGGTATTATTTGCTGGCCGTGCATTGAATAAATGTAAGGCTGTCAAGAATACTTTTCTTGAGGGCTATCTCGGTGTTCATCTCATTTGATAGTTTATGCTGGATAAATGTAAGGCATCTGTCAAGAAATCATTTTCTTGGAGGGCTATCTCGGTGTTCAT</td>
</tr>
<tr>
<td>Chimeras</td>
<td>Sequence</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>xpt(C74U)/lysC</strong></td>
<td>AATTTTCATAGTTAGATCGTGTTATATGGTGAAAGATAATATAATC</td>
</tr>
<tr>
<td></td>
<td>GCCGGATATGGCAGCCAGATTTTCTACCGGGGACCGTAAATGT</td>
</tr>
<tr>
<td></td>
<td>GTCCGATTATTATCTGTGTTGTCATAATCATTATATGAGATTAA</td>
</tr>
<tr>
<td></td>
<td>TTGATAAGGAAATGAGAGTATTTCTCTCTATTTGCGTTTTTTATTGT</td>
</tr>
<tr>
<td></td>
<td>GGAACAAAGGCCTCTTTTCTCCTCACCCGACGAACAAATGTAAGGGGTTGTAATAC</td>
</tr>
<tr>
<td><strong>ribD/lysC</strong></td>
<td>AATTTTCATAGTTAGATCGTGTTATATGGTGAAAGATAATCGGGGCA</td>
</tr>
<tr>
<td></td>
<td>AGGGTGGAAGTCGCCGACCGCCGTAGTAAAGCAGACATTGCTT</td>
</tr>
<tr>
<td></td>
<td>TAGAGCAGCCGGAGCCGCTGCTGTAAGGACGCGCGGTGGATTT</td>
</tr>
<tr>
<td></td>
<td>GCTTAAAGGCGGCTGCTGTAAGGACGCGCGGTGGATTT</td>
</tr>
<tr>
<td></td>
<td>CTTCTTCCTCACCCGACGAACAAATGTAAGGGGTTGTAATAC</td>
</tr>
<tr>
<td><strong>theo/lysC</strong></td>
<td>AATTTCATAGTTAGATCGTGTTATATGGTGAAAGATAATTACCAG</td>
</tr>
<tr>
<td></td>
<td>GTCGAAAGCCATTCCTTTGGAGTTATTCTCMGTGTTATGTAATACCATAT</td>
</tr>
<tr>
<td></td>
<td>TTATTGAGTTAATTGATAAGGAGATGATTATTTCTCTCTATT</td>
</tr>
<tr>
<td></td>
<td>GCTTTTTATTGTTGGAACAAAGGGCTCTTTTCTCCTCACCCGCA</td>
</tr>
<tr>
<td></td>
<td>CGAACAAATGTAAGGGGTTGTAATAC</td>
</tr>
<tr>
<td><strong>yitJ expression platform chimeras</strong></td>
<td>AATTTTCATAGTTAGATCGTGTTATATGGTGAAAGATAATCCAGAGAC</td>
</tr>
<tr>
<td></td>
<td>CTTCGAAAGAGCCCTTTGGAGTTATTCTCMGTGTTATGTAATACCATAT</td>
</tr>
<tr>
<td></td>
<td>TTATTGAGTTAATTGATAAGGAGATGATTATTTCTCTCTATT</td>
</tr>
<tr>
<td></td>
<td>GCTTTTTATTGTTGGAACAAAGGGCTCTTTTCTCCTCACCCGCA</td>
</tr>
<tr>
<td></td>
<td>CGAACAAATGTAAGGGGTTGTAATAC</td>
</tr>
<tr>
<td><strong>yitJ wild type</strong></td>
<td>CTTTCTGACACGAAAAATTTCATATCCGTTTTATCAAGAGAAAG</td>
</tr>
<tr>
<td></td>
<td>CAGAGGGACTGGGCCGACGAGCTTCTAGCAACCCGTTGATAT</td>
</tr>
<tr>
<td></td>
<td>GGGCGATCGCCATGACAAAGTTGCTAATCCAGCAACGCGGCTC</td>
</tr>
<tr>
<td></td>
<td>AACAGCTTGGAAGTAGAAGAGACAAAATCAGCTGACAAGAT</td>
</tr>
<tr>
<td></td>
<td>CTGGCTTTTATTGTTGGAACAAAGGGCTCTTTTCTCCTCACCCGCA</td>
</tr>
<tr>
<td></td>
<td>GGATGTGAATAAGGGGAGGCGAGACAACATG</td>
</tr>
<tr>
<td><strong>metE/yitJ</strong></td>
<td>CTTTCTGACACGAAAAATTTCATATCCGTTTTATCAAGAGAGTT</td>
</tr>
<tr>
<td></td>
<td>GGGCGAGGCCAGTTTGTGCTTTTGCACCGGGGCCAGCTGTATATTTCTCCTCACCGGGCAGCTGTAACAG</td>
</tr>
<tr>
<td></td>
<td>TAAATCCATTTGGGGAATTGGGCGCACTGCTTTTGCACGGCCAGAT</td>
</tr>
<tr>
<td></td>
<td>ACTGATGCTTCTCATAAGGCACCGTTGCTAATTCCAGCAGATTT</td>
</tr>
<tr>
<td></td>
<td>GTCTGAGGATTAAAGAGGACAAAATCAGCTGACAAAGAT</td>
</tr>
<tr>
<td></td>
<td>CTGGCTTTTATTGTTGGAACAAAGGGCTCTTTTCTCCTCACCCGCA</td>
</tr>
<tr>
<td></td>
<td>GGATGTGAATAAGGGGAGGCGAGACAACATG</td>
</tr>
<tr>
<td><strong>xpt/yitJ</strong></td>
<td>CTTTCTGACACGAAAAATTTCATATCCGTTTTATCAAGAGAGTT</td>
</tr>
<tr>
<td></td>
<td>GGGCGATCGCCATGACAAAGTTGCTAATCCAGCAACGCGGCTC</td>
</tr>
<tr>
<td></td>
<td>AACAGCTTGGAAGTAGAAGAGACAAAATCAGCTGACAAGAT</td>
</tr>
<tr>
<td></td>
<td>CTGGCTTTTATTGTTGGAACAAAGGGCTCTTTTCTCCTCACCCGCA</td>
</tr>
<tr>
<td></td>
<td>GGATGTGAATAAGGGGAGGCGAGACAACATG</td>
</tr>
<tr>
<td><strong>xpt(C74U)/yitJ</strong></td>
<td>CTTTCTGACACGAAAAATTTCATATCCGTTTTATCAAGAGAGTT</td>
</tr>
<tr>
<td></td>
<td>GGGCGATCGCCATGACAAAGTTGCTAATCCAGCAACGCGGCTC</td>
</tr>
<tr>
<td></td>
<td>AACAGCTTGGAAGTAGAAGAGACAAAATCAGCTGACAAGAT</td>
</tr>
<tr>
<td></td>
<td>CTGGCTTTTATTGTTGGAACAAAGGGCTCTTTTCTCCTCACCCGCA</td>
</tr>
<tr>
<td></td>
<td>GGATGTGAATAAGGGGAGGCGAGACAACATG</td>
</tr>
<tr>
<td>lysC/yitJ</td>
<td>CTTCTGACACGAAATTTTCATATCCGTTCTTAAGATAGAGGTGCGAACTTCAAGAGTATGCCGTCGAGATAGGATTCTGTGGAAAAGGCTGGAGGAGGCGTGCCTTGGAGAAAGATGGATTCTGTGCTGTCAGAGAATACTTCTTGAGGCTATTTCTTAAAGAAGAGCAGAACAAAATTCATCTACGTTGTGCCGAGAAGCAGTTATGTAAGGCTGTCAAAATCTTCTTCTTTCTTGAGGACTTTTTTTATTCTCTTTTCTTCTTGCTGATGTGAATAAAGGAGGCAGACGAAATG</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>ribD/yitJ</td>
<td>CTTCTGACACGAAATTTTCATATCCGTTCTTAAGGCAACGCCAGCTGGAGTAGTAAGCACTTTGCTTATAAGGCCTCCCGGAGGTGGAAATCCCGACCGGCGGTAGTAAAGCACATTTGCTTTAAGCCCGTGACCCGTGTGCATAAGCACGCGGTGGATTCAGTTTAAGCTGAAGCCGACAGTGAAAGTCTGGATGGGAGATAAGAAGACAAAATCACTGACAAAGTCTTCTTCTTAAGAGGACTTTTTTATTTCTCTTTTCTTCTTGCTGATGTGAATAAAGGAGGCAGACGAAATG</td>
</tr>
<tr>
<td>theo/yitJ</td>
<td>CTTCTGACACGAAATTTTCATATCCGTTCTTAACTACCGCTTGACGAGTAAAGACAGACTTTCCGAAAGAGGCGCTGGTATCCGGACTTTCTTCTTTCTTTCTTGCTGATGTGAATAAAGGAGGCAGACGAAATG</td>
</tr>
<tr>
<td><strong>metE</strong> expression platform chimeras</td>
<td><strong>metE</strong> wild type</td>
</tr>
<tr>
<td><strong>metE (P1-2)</strong></td>
<td>CAAAAAAATTAATAACATGCTCTTTATCGAGAGTGGGCGAGGGATTGGCCTTTTGACCCCAACAGCAACCGGACGGTAAATTACCACTTTTCTTTCTTTCTTTCTTGCTGATGTGAATAAAGGAGGCAGACGAAATG</td>
</tr>
<tr>
<td>Gene</td>
<td>Sequence</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>metE (P1-1)</td>
<td>CAAAAAAATTAAATACATGGTCTCTTATCGAGAGTTGGGCGAGGATTGCCCTTTTACCCCAACAGCAACCGACCCTGAAATTACATGTGTGGAATGAGGCTTTTTACCTTTTGGGCACTGACTGATGCTCATAGCGACACGTGCTAATCCACGATTGTGTCTGAGAGATGAGAGAGAGGCAGTGTTTTACGTAGAAAAGCCTCTTTCTCTCATGGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGTATCCGGGCTTTGAAAGAGAGAGGAGAAACAATG</td>
</tr>
<tr>
<td>metE (P1+1)</td>
<td>CAAAAAAATTAAATAAACGTGTCTCTTATCGAGAGTTGGGCGAGGATTGCCCTTTTACCCCAACAGCAACCGACCCTGAAATTACATGTGTGGAATGAGGCTTTTTACCTTTTGGGCACTGACTGATGCTCATAGCGACACGTGCTAATCCACGATTGTGTCTGAGAGATGAGAGAGAGGCAGTGTTTTACGTAGAAAAGCCTCTTTCTCTCATGGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGTATCCGGGCTTTGAAAGAGAGAGGAGAAACAATG</td>
</tr>
<tr>
<td>metE (P1+2)</td>
<td>CAAAAAAATTAAATAAACATGTTCTCTTATCGAGAGTTGGGCGAGGATTGCCCTTTTACCCCAACAGCAACCGACCCTGAAATTACATGTGTGGAATGAGGCTTTTTACCTTTTGGGCACTGACTGATGCTCATAGCGACACGTGCTAATCCACGATTGTGTCTGAGAGATGAGAGAGAGGCAGTGTTTTACGTAGAAAAGCCTCTTTCTCTCATGGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGTATCCGGGCTTTGAAAGAGAGAGGAGAAACAATG</td>
</tr>
<tr>
<td>yitJ/metE</td>
<td>CAAAAAAATTAAATAAACATTTCCTTCTTTACAGAGAGACAGAGGCACTGGCCCGACGAACTCTCATGGAAGAAGAGGGGCAGTGTTTTACGTAGAAAAGCCTCTTTCTCTCATGGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGTATCCGGGCTTTGAAAGAGAGAGGAGAAACAATG</td>
</tr>
<tr>
<td>xpt/metE</td>
<td>CAAAAAAATTAAATAAACATTTCCTTCTTTATATAATCAGGCTGGGTGAAATATGGCAGCAATTCCTACCGGACGCTGAAATGTCCGACTATAGGAGAGGAGTCTTTAGTAGAAGAGCCTCTTCTCTGTGGGAGAAACCTTTGAGAGGAGGTCTTACCATATAAAGGAGGAGAAACAATG</td>
</tr>
<tr>
<td>xpt/metE (P1-2)</td>
<td>CAAAAAAATTAAATAACATGGTCTCTTATATAATCAGGCTGGGTGAAATATGGCAGCAATTCCTACCGGACGCTGAAATGTCCGACTATAGGAGAGGAGTCTTTAGTAGAAGAGCCTCTTCTCTGTGGGAGAAACCTTTGAGAGGAGGTCTTACCATATAAAGGAGGAGAAACAATG</td>
</tr>
<tr>
<td>xpt/metE (P1-1)</td>
<td>CAAAAAATTAATAACATGTTCTCTTATATAATCGCGTGGATATGACCGAAGCTTCTACCGGCGACCTGAATTGTCGACTATGGAGAGGCCAGTGTGTTTACGTAAGAAAGGCCTTTCTCTCTCATGAGGAAAGAGGTGCTTTTTACGTAGAAAAGCCTCTTTCTCTCATGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGATCCGCGGGTGAAAGAGAGTGTTTTACATATAAAGGAGGAGAAACAATG</td>
</tr>
<tr>
<td>xpt/metE (P1+1)</td>
<td>CAAAAAATTAATAAACATTCTCTCTTTATATAATCGCGTGGATATGAGACCGAAGCTTCTACCGGCGACCTGAATTGTCGACTATGGAGAGGCCAGTGTGTTTACGTAAGAAAGGCCTTTCTCTCTCATGAGGAAAGAGGTGCTTTTTACGTAGAAAAGCCTCTTTCTCTCATGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGATCCGCGGGTGAAAGAGAGTGTTTTACATATAAAGGAGGAGAAACAATG</td>
</tr>
<tr>
<td>xpt/metE (P1+2)</td>
<td>CAAAAAATTAATATACATTCTCTCTTTATATAATCGCGTGGATATGAGACCGAAGCTTCTACCGGCGACCTGAATTGTCGACTATGGAGAGGCCAGTGTGTTTACGTAAGAAAGGCCTTTCTCTCTCATGAGGAAAGAGGTGCTTTTTACGTAGAAAAGCCTCTTTCTCTCATGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGATCCGCGGGTGAAAGAGAGTGTTTTACATATAAAGGAGGAGAAACAATG</td>
</tr>
<tr>
<td>xpt(C74U)/metE</td>
<td>CAAAAAATTAATACATTCTCTCTTTATATAATCGCGTGGATATGAGACCGAAGCTTCTACCGGCGACCTGAATTGTCGACTATGGAGAGGCCAGTGTGTTTACGTAAGAAAGGCCTTTCTCTCTCATGAGGAAAGAGGTGCTTTTTACGTAGAAAAGCCTCTTTCTCTCATGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGATCCGCGGGTGAAAGAGAGTGTTTTACATATAAAGGAGGAGAAACAATG</td>
</tr>
<tr>
<td>lysC/metE</td>
<td>CAAAAAATTAATACATTCTCTCTTTATATAATCGCGTGGATATGAGACCGAAGCTTCTACCGGCGACCTGAATTGTCGACTATGGAGAGGCCAGTGTGTTTACGTAAGAAAGGCCTTTCTCTCTCATGAGGAAAGAGGTGCTTTTTACGTAGAAAAGCCTCTTTCTCTCATGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGATCCGCGGGTGAAAGAGAGTGTTTTACATATAAAGGAGGAGAAACAATG</td>
</tr>
<tr>
<td>ribD/metE</td>
<td>CAAAAAATTAATACATTCTCTCTTTATATAATCGCGTGGATATGAGACCGAAGCTTCTACCGGCGACCTGAATTGTCGACTATGGAGAGGCCAGTGTGTTTACGTAAGAAAGGCCTTTCTCTCTCATGAGGAAAGAGGTGCTTTTTACGTAGAAAAGCCTCTTTCTCTCATGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGATCCGCGGGTGAAAGAGAGTGTTTTACATATAAAGGAGGAGAAACAATG</td>
</tr>
<tr>
<td>theo/metE</td>
<td>CAAAAAATTAATACATTCTCTCTTTATATAATCGCGTGGATATGAGACCGAAGCTTCTACCGGCGACCTGAATTGTCGACTATGGAGAGGCCAGTGTGTTTACGTAAGAAAGGCCTTTCTCTCTCATGAGGAAAGAGGTGCTTTTTACGTAGAAAAGCCTCTTTCTCTCATGGAAAGAGGCTTTTTGTTGTGAGAAAACCTCTTAGCAGCCTGATCCGCGGGTGAAAGAGAGTGTTTTACATATAAAGGAGGAGAAACAATG</td>
</tr>
<tr>
<td><strong>tet/metE</strong> (P1-1)</td>
<td>CAAAAATTAACATTTGCTTTGGGAGAGGTGAAGAATACGAGGCACCTAGGTAAGATACCTCTTAATAACATTTTCTCTTGGGAGAGGTGAAGAATACGAGGCACCTAGGTAAGATACCTCTTAATAACATTT</td>
</tr>
<tr>
<td>-----------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>tet/metE</strong> (P1-2)</td>
<td>CAAAAATTAACATTTGCTTTGGGAGAGGTGAAGAATACGAGGCACCTAGGTAAGATACCTCTTAATAACATTT</td>
</tr>
<tr>
<td><strong>tet/metE</strong> (P1-3)</td>
<td>CAAAAATTAACATTTGCTTTGGGAGAGGTGAAGAATACGAGGCACCTAGGTAAGATACCTCTTAATAACATTT</td>
</tr>
</tbody>
</table>
Appendix 4. Sequences of chimeric riboswitches referred to in chapter 4. Sequences of the different constructs characterized. All sequences are given in DNA format for clarity. In the main text and figures the construct mutations located at the poly U repeat are referred to as 7, 6 or 5 Us, while in this table they are represented as 7, 6 or 5 Ts. The aptamer used in each DNA template is represented in different colors (pink for *pbuE*, yellow for *yitJ*, green for *xpt*, gray for *ribD*, marron for theophylline, blue for *xpt* C74U, green for SAH and another shade of green for tetracycline). The engineered sequence utilized in the chimeric riboswitches containing the *pbuE* expression platform is represented in cyan. The translation start site is represented in red.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>pbuE</strong></td>
<td>AATTAAATAGCTTTATCACTTG TATAACCTCAATAATATGGTTTG AGGGTGCTTACCAGGAACCCGTTAAAATCTGGATTAACAGGTTTTATG TATGACATTTTTTGTATACAGGATTTTTTTATTATCAGAAAACATTTA ATAGTAAAGGAGTTTTGTTATG</td>
</tr>
<tr>
<td><strong>pbuE 7T</strong></td>
<td>AATTAAATAGCTATTATCACCTTG TATAACCTCAATAATATGGTTTG AGGGGTCTTACCAGGAACCCGTTAAAATCTGGATTAACAGGTTTTATG TATGACATTTTTTGTATACAGGATTTTTTTATTATCAGAAAACATTTA ATAGTAAAGGAGTTTTGTTATG</td>
</tr>
<tr>
<td><strong>pbuE 6T</strong></td>
<td>AATTAAATAGCTTTATCACTTG TATAACCTCAATAATATGGTTTG AGGGTGCTTACCAGGAACCCGTTAAAATCTGGATTAACAGGTTTTATG TATGACATTTTTTGTATACAGGATTTTTTTATTATCAGAAAACATTTA ATAGTAAAGGAGTTTTGTTATG</td>
</tr>
<tr>
<td><strong>pbuE/pbuE</strong></td>
<td>AATTAAATAGCTTTATCACTTG TATAACCTCAATAATATGGTTTG AGGGGTCTTACCAGGAACCCGTTAAAATCTGGATTAACAGGTTTTATG TATGACATTTTTTGTATACAGGATTTTTTTATTATCAGAAAACATTTA ATAGTAAAGGAGTTTTGTTATG</td>
</tr>
<tr>
<td><strong>pbuE/pbuE 7T</strong></td>
<td>AATTAAATAGCTTTATCACTTG TATAACCTCAATAATATGGTTTG AGGGGTCTTACCAGGAACCCGTTAAAATCTGGATTAACAGGTTTTATG TATGACATTTTTTGTATACAGGATTTTTTTATTATCAGAAAACATTTA ATAGTAAAGGAGTTTTGTTATG</td>
</tr>
<tr>
<td><strong>pbuE/pbuE 6T</strong></td>
<td>AATTAAATAGCTTTATCACTTG TATAACCTCAATAATATGGTTTG AGGGGTCTTACCAGGAACCCGTTAAAATCTGGATTAACAGGTTTTATG TATGACATTTTTTGTATACAGGATTTTTTTATTATCAGAAAACATTTA ATAGTAAAGGAGTTTTGTTATG</td>
</tr>
<tr>
<td><strong>pbuE/pbuE 5T</strong></td>
<td>AATTAAATAGCTTTATCACTTG TATAACCTCAATAATATGGTTTG AGGGGTCTTACCAGGAACCCGTTAAAATCTGGATTAACAGGTTTTATG TATGACATTTTTTGTATACAGGATTTTTTTATTATCAGAAAACATTTA ATAGTAAAGGAGTTTTGTTATG</td>
</tr>
<tr>
<td><strong>yitJ/pbuE</strong></td>
<td>AATTAAATAGCTTTATCACTTG TATAACCTCAATAATATGGTTTG AGGGGTCTTACCAGGAACCCGTTAAAATCTGGATTAACAGGTTTTATG TATGACATTTTTTGTATACAGGATTTTTTTATTATCAGAAAACATTTA ATAGTAAAGGAGTTTTGTTATG</td>
</tr>
<tr>
<td>yitJ/pbuE 7T</td>
<td>ATAAAAATAGCTATTATCACGATTTTATCAAGAGAAGCAGAGGGAGCTGGGCCCAGCAAGCTGGGATGAAAATGCTGAATGGGCGTGAACGATGACAAAGCTCAAGAACTTTATGCATTTTTATATTTATCAAAACATTTAAGTAAAGGAGTTATG</td>
</tr>
<tr>
<td>yitJ/pbuE 6T</td>
<td>ATAAAAATAGCTATTATCACGATTTTATCAAGAGAAGCAGAGGGAGCTGGGCCCAGCAAGCTGGGATGAAAATGCTGAATGGGCGTGAACGATGACAAAGCTCAAGAACTTTATGCATTTTTATATTTATCAAAACATTTAAGTAAAGGAGTTATG</td>
</tr>
<tr>
<td>xpt/pbuE 7T</td>
<td>ATAAAAATAGCTATTATCACGATTTTATCAAGAGAAGCAGAGGGAGCTGGGCCCAGCAAGCTGGGATGAAAATGCTGAATGGGCGTGAACGATGACAAAGCTCAAGAACTTTATGCATTTTTATATTTATCAAAACATTTAAGTAAAGGAGTTATG</td>
</tr>
<tr>
<td>xpt/pbuE 6T</td>
<td>ATAAAAATAGCTATTATCACGATTTTATCAAGAGAAGCAGAGGGAGCTGGGCCCAGCAAGCTGGGATGAAAATGCTGAATGGGCGTGAACGATGACAAAGCTCAAGAACTTTATGCATTTTTATATTTATCAAAACATTTAAGTAAAGGAGTTATG</td>
</tr>
<tr>
<td>ribD/pbuE 7T</td>
<td>ATAAAAATAGCTATTATCACGATTTTATCAAGAGAAGCAGAGGGAGCTGGGCCCAGCAAGCTGGGATGAAAATGCTGAATGGGCGTGAACGATGACAAAGCTCAAGAACTTTATGCATTTTTATATTTATCAAAACATTTAAGTAAAGGAGTTATG</td>
</tr>
<tr>
<td>ribD/pbuE 6T</td>
<td>ATAAAAATAGCTATTATCACGATTTTATCAAGAGAAGCAGAGGGAGCTGGGCCCAGCAAGCTGGGATGAAAATGCTGAATGGGCGTGAACGATGACAAAGCTCAAGAACTTTATGCATTTTTATATTTATCAAAACATTTAAGTAAAGGAGTTATG</td>
</tr>
<tr>
<td></td>
<td>TGCCCCCGGTTTGC[...]CAATTTCGCTTTGCCGAGGTGATCCATG</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>tet/metH</td>
<td>CTCAGGCGGCCCTTCCGGGTCGCTCCGCCTTCCGGTCTTCGAGCUCCGCTGGAGAGGTTGAGAGAATACGACCACCTAGGTTAGAAAATACCTAAAACATACCGGGCGCTCGCAAAACCCGCGTTTTCCCTTGCCCCGGGTGTCGCGCCTCTTTTTTTTGCGGCGTCCCCGGTGCAATTTCGCTTTGCCGAGGTGATCCATG</td>
</tr>
</tbody>
</table>