

Centre Armand-Frappier Santé Biotechnologie

ÉTUDE DE RIBOSWITCHS SAM-II DÉVIANTE DU CONSENSUS ET DE LEUR RÔLE DANS UNE RÉGULATION COORDONNÉE DE LA VOIE DE SYNTHÈSE DE LA MÉTHIONINE ET DE SAM

Par

Vesta Korniakova

Mémoire présenté pour l'obtention du grade de
Maître ès Sciences (M.Sc.)
en microbiologie appliquée

Jury d'évaluation

Président du jury et
examineur interne

Frédéric Veyrier
INRS-Institut Armand Frappier

Examineur externe

Daniel Lafontaine
Département de Biologie
Université de Sherbrooke

Examineur interne

Jonathan Perreault
INRS-Institut Armand Frappier

Directeur de recherche

Jonathan Perreault
INRS-Institut Armand Frappier

REMERCIEMENTS

J'aimerais remercier mon superviseur Jonathan Perreault de m'avoir donné l'opportunité de découvrir le monde des ARN régulateurs, de m'avoir encouragé à poursuivre toutes mes curiosités et partager mes recherches lors de nombreux congrès. J'aimerais aussi remercier tous mes collègues à l'INRS-IAF pour leur aide et conseils. C'était une belle expérience à travers laquelle j'ai découvert ma passion pour la recherche.

RÉSUMÉ

Ce travail présente trois chapitres liés à la caractérisation des riboswitchs.

Le premier explore l'utilisation du riboswitch fluor comme un outil pour réprimer l'activité des promoteurs forts et/ou non-caractérisés dans le but de mitiger la toxicité causée par la surexpression d'un gène rapporteur lors des clonages. Cette étude révèle le succès de cette stratégie et aussi ses limites. De plus, un phénomène inhabituel d'inversion de régulation lors de l'utilisation d'un riboswitch non-natif à l'hôte est rapporté.

Le deuxième chapitre est axé sur des riboswitchs SAM-II/V (S-adenosylmethionine) déviant du consensus par la présence d'éléments structurels distincts : une tige compétitrice et une tige accessoire. On confirme *in vitro* l'affinité pour SAM d'environ dix membres du nouveau consensus proposé. De plus, grâce au vecteur décrit au premier chapitre on confirme chez *Burkholderia thailandensis* la régulation *in vivo* de trois cas de riboswitchs SAM structurellement différent se retrouvant en amont de *metK*, *metX* ou *metZ*.

La dernière étude caractérise le riboswitch SAM en amont du gène *metK*, une structure modèle parmi les nouveaux riboswitchs découverts dans la deuxième étude. Une série de mutants étudié *in silico*, *in vitro*, et *in vivo* décortique le rôle des deux éléments structurels caractéristiques du nouveau consensus.

Mots-clés : riboswitch, fluor, gène rapporteur, plasmide, toxicité, fusion traductionnelle, SAM, *Burkholderia*, pseudo-nœud

ABSTRACT

This work presents three projects related to the characterization of riboswitches.

The first explores the fluoride riboswitch as a tool for inducing the activity of potent constitutive and/or un-characterized promoters for circumventing lethal toxicity caused by the overexpression of a reporter gene during cloning and transformation. This study reveals the success of such a strategy as well as its limits. Furthermore, an unusual phenomenon of inversion of regulation of a riboswitch in a non-native host is reported.

The second study focuses on SAM-II/V (S-adenosylmethionine) riboswitches deviating from the consensus due to the presence of one or two distinct structural elements: a competing stem and an accessory stem. We confirm *in vitro* SAM binding affinity of about ten members of this newly proposed consensus. Additionally, we verified *in vivo* regulation in *Burkholderia thailandensis* E264 of three instances of structurally different SAM riboswitches occurring upstream of either *metK*, *metX*, or *metZ*.

The final study characterizes the *metK* SAM riboswitch, a model structure amongst the newly discovered riboswitches described in the second project. A series of mutants studied *in silico*, *in vitro*, and *in vivo* unravels the roles of the two structural elements characteristic of the new consensus.

Keywords : riboswitch, fluoride, reporter, plasmid, toxicity, translational fusion, SAM, *Burkholderia*, pseudoknot

TABLE DES MATIÈRES

REMERCIEMENTS	III
RÉSUMÉ	V
ABSTRACT	VII
TABLE DES MATIÈRES	IX
LISTE DES FIGURES	XI
LISTE DES TABLEAUX	XIII
LISTE DES ABRÉVIATIONS	XV
1 INTRODUCTION	1
1.1 DIVERSITE DES ARN REGULATEURS CHEZ LES BACTERIES.....	1
1.2 LES RIBOSWITCHS SAM.....	6
1.2.1 <i>Superfamille SAM-I</i>	8
1.2.2 <i>Superfamille SAM-II</i>	9
1.2.3 <i>Superfamille SAM-III</i>	12
1.2.4 <i>Riboswitchs SAM/SAH et SAH</i>	12
1.3 LES RIBOSWITCHS FLUOR.....	13
1.4 METHODES DE RECHERCHE ET DE VALIDATION DES RIBOSWITCHS	13
1.4.1 <i>Méthodes in silico</i>	13
1.4.2 <i>Méthodes in vitro</i>	15
1.4.3 <i>Méthodes in vivo</i>	16
1.5 VALIDATION D'UN NOUVEAU VARIANT DU RIBOSWITCH SAM	16
1.5.1 <i>Problématique</i>	16
1.5.2 <i>Hypothèse et objectif</i>	17
2 BIBLIOGRAPHIE DU CHAPITRE 1	19
3 RIBOSWITCH-MEDIATED EXPRESSION DAMPENING	21
3.1 ABSTRACT.....	22
3.2 INTRODUCTION	22
3.3 METHODS	25
3.3.1 <i>DNA amplification and reporter plasmid construct assembly</i>	25
3.3.2 <i>Bacterial strains and construction of reporter strains</i>	25
3.3.3 <i>Testing Gibson assembly-based cloning in the plasmid</i>	26
3.3.4 <i>lux reporter assay</i>	26
3.3.5 <i>Luciferase reporter data analysis</i>	27
3.4 RESULTS AND DISCUSSION	28

3.4.1	<i>The pVK-f-lux Plasmid Allows for Rapid Mixing and Matching Promoters and 5'UTRs ...</i>	28
3.4.2	<i>The Fluoride Riboswitch Dampens Reporter Gene Expression and Facilitates Potent Promoter Cloning.....</i>	31
3.4.3	<i>Portability of riboswitches from one species to another can be unexpected.....</i>	35
3.5	CONCLUSION.....	37
4	BIBLIOGRAPHIE DU CHAPITRE 3.....	39
5	STRUCTURAL DIVERSITY OF NEW SAM-II/V RIBOSWITCH VARIANTS.....	41
5.1	ABSTRACT.....	43
5.2	INTRODUCTION.....	43
5.3	RESULTS AND DISCUSSION.....	44
5.3.1	<i>Discovery of new, structurally diverse, riboswitch variants.....</i>	44
5.3.2	<i>Three SAM riboswitches regulating SAM and methionine synthesis in Burkholderia thailandensis.....</i>	46
5.4	CONCLUSION.....	51
5.5	ACKNOWLEDGEMENTS.....	52
6	BIBLIOGRAPHIE DU CHAPITRE 5.....	53
7	CARACTÉRISATION DES ÉLÉMENTS STRUCTURELS DU RIBOSWITCH SAM EN AMONT DU GÈNE <i>METK</i>.....	55
7.1	INTRODUCTION.....	55
7.2	MATERIELS ET METHODES.....	56
7.2.1	<i>Amplification de l'ADN et construction de plasmides.....</i>	56
7.2.2	<i>Souches bactériennes et transformations.....</i>	59
7.2.3	<i>Essais in-line et essais in vivo de luminescence.....</i>	61
7.3	RESULTATS ET DISCUSSION.....	61
8	BIBLIOGRAPHIE DU CHAPITRE 7.....	72
9	DISCUSSION GÉNÉRALE ET CONCLUSION.....	73
9.1	CARACTERISATION DE PVK-F-LUX COMME OUTIL POUR L'ETUDE DES SEQUENCES INTERGENIQUES.....	73
9.2	ÉTUDE DES RIBOSWITCHS SAM CHEZ <i>B. THAILANDENSIS</i> E264.....	74
9.3	CONCLUSION.....	76
10	BIBLIOGRAPHIE DU CHAPITRE 9.....	77
11	ANNEXE I : MATERIEL SUPPLÉMENTAIRE RELIÉ AU CHAPITRE 2.....	79
12	ANNEXE II : MATÉRIEL SUPPLÉMENTAIRE RELIÉ AU CHAPITRE 3.....	115

LISTE DES FIGURES

FIGURE 1.1	MECANISMES DE REGULATION DE RIBOSWITCHS.....	2
FIGURE 1.2	MECANISMES DE REGULATION DE RIBOSWITCHS.....	5
FIGURE 1.3	STRUCTURES DE SAM ET SAH.....	7
FIGURE 1.4	REGULATION DES VOIES METABOLIQUES IMPLIQUEES DANS LA SYNTHESE DE METHIONINE CHEZ LES PROTEOBACTERIES.....	8
FIGURE 1.5	LA SUPERFAMILLE SAM-I.....	9
FIGURE 1.6	LES PSEUDO-NŒUDS FAISANT PARTIE DES RIBOSWITCHS SAM.....	10
FIGURE 1.7	RIBOSWITCH SAM-II DE LA SUPERFAMILLE SAM-II.....	11
FIGURE 1.8	RIBOSWITCH SAM-III.....	12
FIGURE 1.9	CLIVAGE SPONTANE DE L'ARN.....	16
FIGURE 3.1	RIBOSWITCH-MEDIATED REPORTER EXPRESSION UNDER THE CONTROL OF A POTENT PROMOTER.	30
FIGURE 3.2	TIME-COURSE LUMINESCENCE INDUCTION CURVES OF <i>B. THAILANDENSIS</i> E264 CLONES CONTAINING THE <i>B. THAILANDENSIS</i> FLUORIDE RIBOSWITCH AND REPRESSION EFFECT.....	32
FIGURE 3.3	THE <i>B. THAILANDENSIS</i> E264 FLUORIDE RIBOSWITCH UPREGULATES IN ORIGINAL HOST AND DOWNREGULATES IN <i>E. COLI</i> DH5A AND <i>E. COLI</i> SM10 Δ PIR.....	36
FIGURE 5.1	A NOVEL RIBOSWITCH FAMILY WITHIN THE SAM-II CLAN.....	45
FIGURE 5.2	DIFFERENTIAL REGULATION OF THE METHIONINE/SAM BIOSYNTHESIS PATHWAYS BY DIFFERENT SAM RIBOSWITCHES IN <i>BURKHOLDERIA</i>	49
FIGURE 7.1	DIVERS MUTANTS DE L'APTAMÈRE DU RIBOSWITCH SAM EN AMONT DE <i>METK</i> CHEZ <i>B.</i> <i>THAILANDENSIS</i> E264.	64
FIGURE 7.2	RESTRUCTURATION DU MUTANT P1 BULGE.....	66
FIGURE 7.3	COURBE DE K_D GÉNÉRÉE À PARTIR D'UNE EXPÉRIENCE DE <i>IN LINE PROBING</i> POUR LE MUTANT <i>P1</i> <i>BULGE</i>	67
FIGURE 7.4	DISTRIBUTION DES STRUCTURES DES MUTANTS ET DU WT SELON LEUR POPULATION STATISTIQUE, LEUR STABILITÉ THERMODYNAMIQUE, ET DE LEUR AFFINITÉ POUR SAM.	68
FIGURE 7.5	REPRÉSENTATION D'UN RÉSEAU D'INTRA/INTERACTION DE LA DISTRIBUTION DE LA STABILITÉ DES STRUCTURES D'ARN RÉGULATEURS POUR UNE SÉQUENCE DONNÉE.	70

LISTE DES TABLEAUX

TABLEAU 1.1	EXEMPLES DES FONCTIONS DES ARNnc DU TABLEAU	2
TABLEAU 3.1	CLONING SUCCESS OF THE PROMOTERS UPSTREAM OF THE LUX OPERON DEPENDS ON THE CHOICE OF 5'UTR.	34
TABLEAU 7.1	LISTE DES OLIGONUCLEOTIDES UTILISES POUR CE CHAPITRE	56
TABLEAU 7.2	LISTE DES PLASMIDES UTILISEES POUR CE CHAPITRE.....	58
TABLEAU 7.3	LISTE DES SOUCHES UTILISEES POUR CE CHAPITRE	59
TABLEAU 7.4	SOMMAIRE DE LA CARACTERISATION PAR ETUDE MUTATIONNELLE DU RIBOSWITCH SAM EN AMONT DE <i>METK</i> CHEZ <i>B. THAILANDENSIS</i> E264.....	62

LISTE DES ABRÉVIATIONS

5' UnTranslated Regions	
(UTRs).....	24
adénosine triphosphate	
(ATP).....	6
ARN non-codants	
(ARNnc)	1
<i>B. thailandensis</i> E264 fluoride riboswitch	
(B. thai F RS).....	29
<i>B. thailandensis</i> E264 <i>metX</i> SAM riboswitch	
(B. thai <i>metX</i> SAM RS).....	31
Clustered Regularly Interspaced Short Palindromic Repeats	
(CRISPR)	2
fluoride ions	
(F-)	27
fluoride riboswitch	
(F RS).....	24
fold induction	
(FI)	24
Gibson assembly	
(GA)	25
la polymérase d'ARN	
(RNAP).....	72
la S-adénosylméthionine	
(SAM).....	6
Les petits ARN	
(sRNA)	2

Luria Broth	
(LB)	25
M9 Minimal Media	
(M9-MM).....	26
M9 Minimal Media with no sulfur	
(0.5X M9-MM-S)	69
Multiple Cloning Site	
(MCS).....	31
Multiple Cloning Site 1	
(MCS1).....	28
Multiple Cloning Site 2	
(MCS2).....	28
N-methylisatoic	
(NMIA).....	15
<i>P. syringae</i> fluoride riboswitch	
(P. syr F RS).....	28
réactif sulfate de diméthyle	
(DMS).....	15
régions 5' non-traduite	
(5'UTR).....	3
régions intergéniques	
(IGR)	13
Relative Luminescence Units	
(RLU).....	28
restriction enzyme	
(RE).....	25
S7 ribosomal protein gene promoter	
(PS7).....	24
S-adénosyl-L-homocystéine	

(SAH)	6
S-adenosylmethionine-V	
(SAM-V)	43
SAM-I	
(S-box)	8
<i>Selective 2'-hydroxyl acylation analyzed by primer extension</i>	
(SHAPE).....	15
Shine-Dalgarno	
(SD).....	4
sites de clonages multiples	
(MCS).....	71
Watson-Crick	
(WC).....	9

1 INTRODUCTION

L'évolution a permis aux bactéries ont d'assurer leur survie dans une multitude d'environnements. C'est grâce à une diversité de mécanismes de régulations géniques que cette adaptation se manifeste au cours de la durée de vie d'une bactérie. La régulation peut se faire à plusieurs niveaux: au niveau transcriptionnel, traductionnel, post-transcriptionnel ou post-traductionnel (Ramírez-Clavijo & Montoya-Ortíz, 2013).

Bien que la majorité des gènes soient régulés à l'aide des protéines, la régulation par l'action des ARN est de plus en plus étudiée. Des ARN non-codants (ARNnc) , tels que les petits ARN et les longs ARNnc, les riboswitchs, les ribozymes et les thermomètres d'ARN, régulent une grande variété de processus cellulaires et ont été découverts dans la nature chez les archées, les bactéries et les eucaryotes (Lotz & Suess, 2018).

1.1 Diversité des ARN régulateurs chez les bactéries

Les ARNnc peuvent nécessiter des complexes d'ARN-protéines pour exécuter leurs fonctions, comme réguler l'expression génique, protéger la bactérie contre des virus, participer à la synthèse d'ADN, ou réarranger les acides nucléiques lors de l'épissage (Cech & Steitz, 2014). Des exemples des ARNnc couplés aux protéines incluent les ribosomes, la télomérase, les microARN, et les longs ARNnc (Cech & Steitz, 2014). D'autres ARN régulateurs comme les riboswitchs fonctionnent sans l'aide des complexes protéiques. Les diverses fonctions des ARNnc chez une bactérie sont représentées schématiquement dans la Figure 1.1. Quelques exemples précis des ARNnc et de leurs fonctions sont énumérés dans le Tableau 1.1.

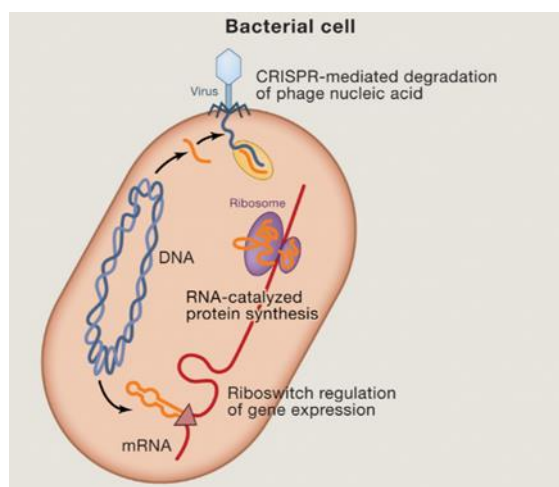


Figure 1.1 Mécanismes de régulation de riboswitchs

Chez les bactéries, les ARNnc peuvent prendre le rôle de régulateurs géniques ou de stabilisateurs et organisateurs structurales du génome ou de certaines nucléoprotéines (Cech & Steitz, 2014). Les ARNnc régulateurs, en orange, agissent sur l'ARNm, en rouge. Les ovales sont des unités protéiques et l'ADN est en bleu. Le triangle est un métabolite lié à un riboswitch. (Image tirée avec permission de (Cech & Steitz, 2014) Copyright © 2014 Elsevier Inc. All rights reserved.)

Tableau 1.1 Exemples des fonctions des ARNnc du tableau

ARNnc	Fonction
ARN CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)	ARN régulateur contenant des séquences répétées qui sert pour diriger la nucléase Cas à cliver un ADN exogène.
ARN CsrB	ARN qui interagit avec la protéine CsrA pour la régulation et qui, entre autres joue un rôle dans la gestion de carbone.
ARNr	ARN qui, associés aux protéines du ribosome, orchestre la traduction.
sRNA (petit ARN, <i>small RNA</i>)	ARN qui s'hybride sur un ou des ARNm cibles pour réguler certains gènes.
ARN 6S	ARN qui régule la transcription en inhibant l'ARN polymérase.
riboswitch	ARN structuré qui change de conformation suite à la liaison d'un ligand pour réguler l'expression d'un gène au niveau de la transcription ou de la traduction.

(Cech & Steitz, 2014)

Les petits ARN (sRNA), qui sont typiquement entre 40 et 400 nt de longueur (Leistra *et al.*, 2019), peuvent réguler la transcription, la traduction, ou la stabilité de l'ARNm (Leistra *et al.*, 2019) des gènes en trans, en se liant, grâce à une certaine complémentarité de bases, à une cible précise

d'un ARNm (Cech & Steitz, 2014). Chez les bactéries, les ARN antisens régulent un gène sur le même locus, mais sur le brin d'ADN opposé, duquel ils sont transcrits (Cech & Steitz, 2014). Les sRNA peuvent aussi contrôler l'expression génique en *cis*, agissant sur la transcription ou la traduction d'un gène adjacent (Leistra *et al.*, 2019).

Les sRNA peuvent réprimer l'expression d'un gène en s'hybridant proche d'un site de liaison du ribosome pour ne pas permettre au ribosome de se lier ou en introduisant un site de dégradation d'ARNm par l'ARNase III (Cech & Steitz, 2014). L'activité des sRNA est souvent dépendante de la protéine chaperonne, Hfq (Cech & Steitz, 2014). De plus, les sRNA peuvent avoir un effet régulateur en se liant aux protéines d'une façon à inhiber leur activité (Leistra *et al.*, 2019).

Parmi les différents sRNA, plusieurs nécessitent l'interaction avec la protéine chaperonne Hfq pour pouvoir effectuer leur rôle de régulateur. Cette protéine a une structure quaternaire qui ressemble à un anneau et qui a deux faces qui peuvent interagir avec l'ARN (Vogel & Luisi, 2011) pour le stabiliser et aider la liaison à l'ARNm (Santiago-Frangos *et al.*, 2016). Il est estimé que les sRNA associés à la protéine Hfq régulent jusqu'à 1% de la totalité des gènes chez certaines espèces tel que *E. coli* et *Salmonella* (Hor *et al.*, 2018). La classe des ARN qu'on appelle les ARN CsrB régulent la protéine CsrA qui est clé dans les voies métaboliques du glycogène (Hor *et al.*, 2018). Finalement, les ARN 6S qui dirigent l'ARN polymérase (RNAP) ainsi que les ARN CRISPR qui participent dans le système immunitaire des bactéries font aussi partie des classes majeures des petits ARN régulateurs (Hor *et al.*, 2018).

Finalement, les riboswitchs sont une autre catégorie d'ARN régulateurs qui permettent la régulation génique, grâce à leur capacité d'adopter des structures distinctes selon la présence ou l'absence d'un ligand spécifique. Les ligands pour les riboswitchs connus incluent des vitamines (ou plutôt leurs co-facteurs), des acides aminés, des nucléotides, des seconds messagers et des ions, pour un total d'environ quarante classes de riboswitchs découvertes depuis 2002, chacune différant par sa structure de base (Lotz & Suess, 2018).

Se retrouvant principalement dans les régions 5' non-traduite (5'UTR) des ARNm chez les bactéries, ces régulateurs fonctionnent majoritairement en *cis* par action directe sur un gène adjacent. Un riboswitch comprend un domaine récepteur, qu'on appelle l'aptamère, et qui permet la liaison d'un ligand spécifique, et d'un domaine transducteur, qu'on appelle la plateforme d'expression, et qui, selon l'état de liaison de l'aptamère, traduit ce signal en modulant le niveau d'expression génique. Cette modulation est possible grâce à différents mécanismes qui contrôlent typiquement soit le niveau de transcription soit le niveau de traduction d'un gène. L'aptamère d'un riboswitch adopte différentes structures en présence et en absence de son ligand spécifique, ce

qui cause un réarrangement dans la plateforme d'expression qui est nécessaire pour la régulation.

Selon le riboswitch, la répression d'un gène peut s'exercer soit lorsque la présence ou l'absence d'un ligand cause une séquestration de la séquence de Shine-Dalgarno (SD) bloquant ainsi la liaison du ribosome lors de la traduction (Cech & Steitz, 2014) (Figure 1.2 a), ou cause la formation d'un terminateur de transcription, souvent de type rho-indépendant (Leyn *et al.*, 2014), qui ne permet pas au RNAP de procéder (Figure 1.2 b).

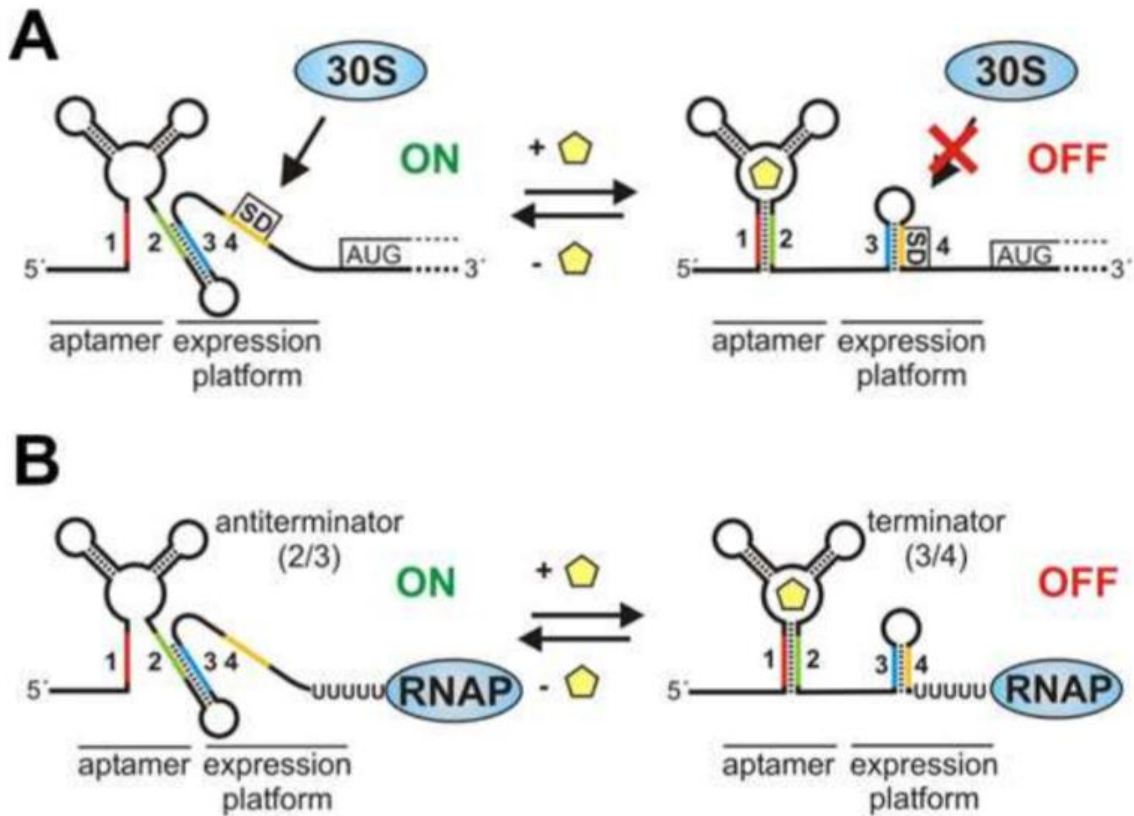


Figure 1.2 Mécanismes de régulation de riboswitchs

En (A) est la représentation du mécanisme de régulation de la traduction par un riboswitch. En absence du ligand la structure permet la liaison de la sous-unité 30S du ribosome à la séquence SD et la traduction peut être initiée. En présence du ligand il y a un changement de conformation du riboswitch et la séquence SD s'apparie pour ne plus être accessible à la liaison de la sous-unité 30S (Wittmann & Suess, 2012). En (B) est la représentation du mécanisme de régulation de la transcription par un riboswitch. En absence du ligand il y a formation d'une structure anti-terminateur et en présence du ligand une structure terminateur se forme ce qui cause la terminaison de transcription par le RNAP (Wittmann & Suess, 2012). Également, certains riboswitchs fonctionnent à l'inverse et c'est la liaison d'un ligand qui cause une structure favorable à l'expression d'un gène. (Image créée par Wittmann A. et Suess B., © 2012 Federation of European Biochemical Societies, <https://febs.onlinelibrary.wiley.com/doi/pdf/10.1016/j.febslet.2012.02.038>, CC BY-NC-ND 4.0 license <https://creativecommons.org/licenses/by-nc-nd/4.0/> (Wittmann & Suess, 2012))

1.2 Les riboswitchs SAM

Pour la plupart des ligands, on connaît une seule famille de riboswitchs mais pour certains, tel que la prequeuosine-1, la guanidine, la S-adénylméthionine (SAM) et la 3'-5'-di-guanosine monophosphate cyclique, plusieurs classes de riboswitchs ont été découvertes (Lotz & Sues, 2018). De tous les ligands, le ligand SAM et son produit déméthylé S-adényl-L-homocystéine (SAH), ont le plus de familles de riboswitchs connues (Price *et al.*, 2014), soit huit familles différentes (Lotz & Sues, 2018). Parmi ces huit familles, il en existe une qui reconnaît le produit déméthylé de SAM, SAH, ainsi qu'une famille qui reconnaît les deux métabolites, SAM et SAH, avec des affinités similaires (Price *et al.*, 2014). Les huit familles de riboswitchs SAM ont toutes des architectures suffisamment différentes pour nécessiter des modèles de covariance différents pour effectuer des recherches d'homologie avec Infernal—l'outil utilisé pour trouver tous les exemples de structures d'ARN connues dans les bases de données de séquences publiques (décrit plus en détail dans la section 1.3.1). De plus, certaines familles peuvent être regroupées en trois superfamilles (ou clans), selon la pochette de liaison de SAM (Trausch *et al.*, 2014). Cela suggère que trois voies évolutives indépendantes ont menés à l'apparition des pochettes de liaisons différentes pour SAM (Price *et al.*, 2014). La superfamille de SAM-I comprend les familles SAM-I (S-box), SAM-IV, et SAM-I/IV et la superfamille de SAM-II comprend les familles SAM-II et SAM-V (Price *et al.*, 2014). La dernière famille est celle de SAM-III (Price *et al.*, 2014). De plus, une nouvelle famille a été récemment découverte, soit la famille SAM-VI et celle-ci serait apparentée à la famille SAM-III (Mirihana Arachchilage *et al.*, 2018). Certaines des familles de riboswitchs, contiennent dans leurs structures des pseudo-nœuds qui sont des topologies de repliement causées par les interactions d'appariement de bases entre l'hélice d'une boucle, faisant partie d'une structure tige-boucle, et une hélice d'ARN adjacente pour former des courtes structures de double hélices (Moore, 1999).

Les riboswitchs contrôlent souvent les gènes d'une manière rétro-inhibitrice, pour réguler soit le transport, la biosynthèse ou la dégradation de leurs ligands (Price *et al.*, 2014). La molécule SAM est la deuxième molécule la plus utilisée dans les réactions enzymatiques après l'adénosine triphosphate (ATP) et sa fonction principale est celle d'un donneur de groupement méthyle (Batey, 2011). SAM est impliqué dans les réactions de méthylation de l'ADN, de l'ARN, de protéines et d'autres métabolites (Leyn *et al.*, 2014).

La molécule SAM cède son groupement méthyle à travers une réaction d'attaque nucléophile qui génère le produit SAH (Batey, 2011) (Figure 1.3). La synthèse de SAM ainsi que son recyclage nécessite la disponibilité de la méthionine (Batey, 2011), un acide aminé comprenant

le soufre. Les autres groupes fonctionnels présents dans la molécule SAM sont un groupement aminoalkyl, un ribosyl, un méthyle, et un groupement amine (Batey, 2011). Les voies métaboliques impliquées dans la synthèse de SAM sont interconnectées avec le métabolisme du soufre qui est un élément limitant dans l'environnement.

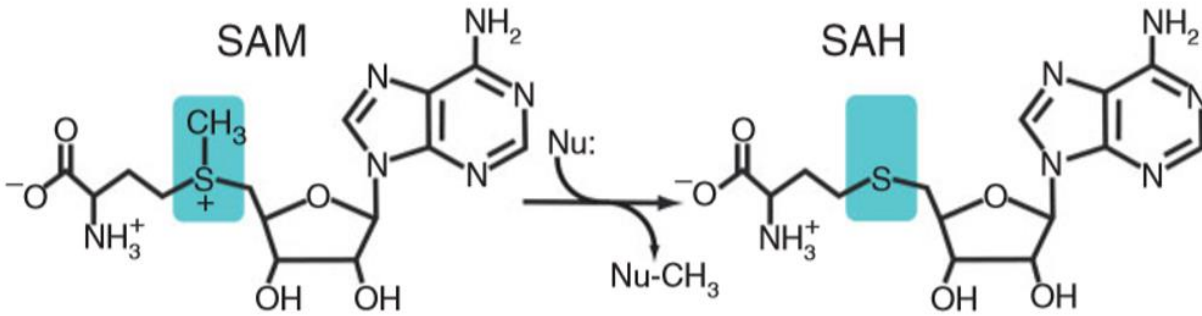


Figure 1.3 Structures de SAM et SAH

SAM est un donneur de méthyle. Lorsque le groupement méthyle est perdu via une attaque nucléophile, la molécule SAH est alors produite. (Image tirée avec permission de (Batey, 2011) Copyright © 2011 John Wiley & Sons, Ltd.)

Chez les bactéries, la biosynthèse et la dégradation de SAM et de méthionine sont régulées précisément (Batey, 2011) par des riboswitchs ainsi que par des protéines. Les voies métaboliques impliquées dans la synthèse de méthionine chez les protéobactéries sont illustrées dans la Figure 1.4. Chez *B. subtilis* onze différents riboswitchs de type SAM-I régulent des gènes associés au métabolisme de méthionine. Chez plusieurs groupes de γ -protéobactéries incluant *Enterobacteriales*, *Pasteurellales*, *Vibrionales*, *Aeromonadales* et *Alteromonadales* (Leyn *et al.*, 2014), la protéine MetJ, qui est retro-inhibée par SAM, est un régulateur transcriptionnel de plusieurs gènes des voies de la biosynthèse de la méthionine (Batey, 2011).

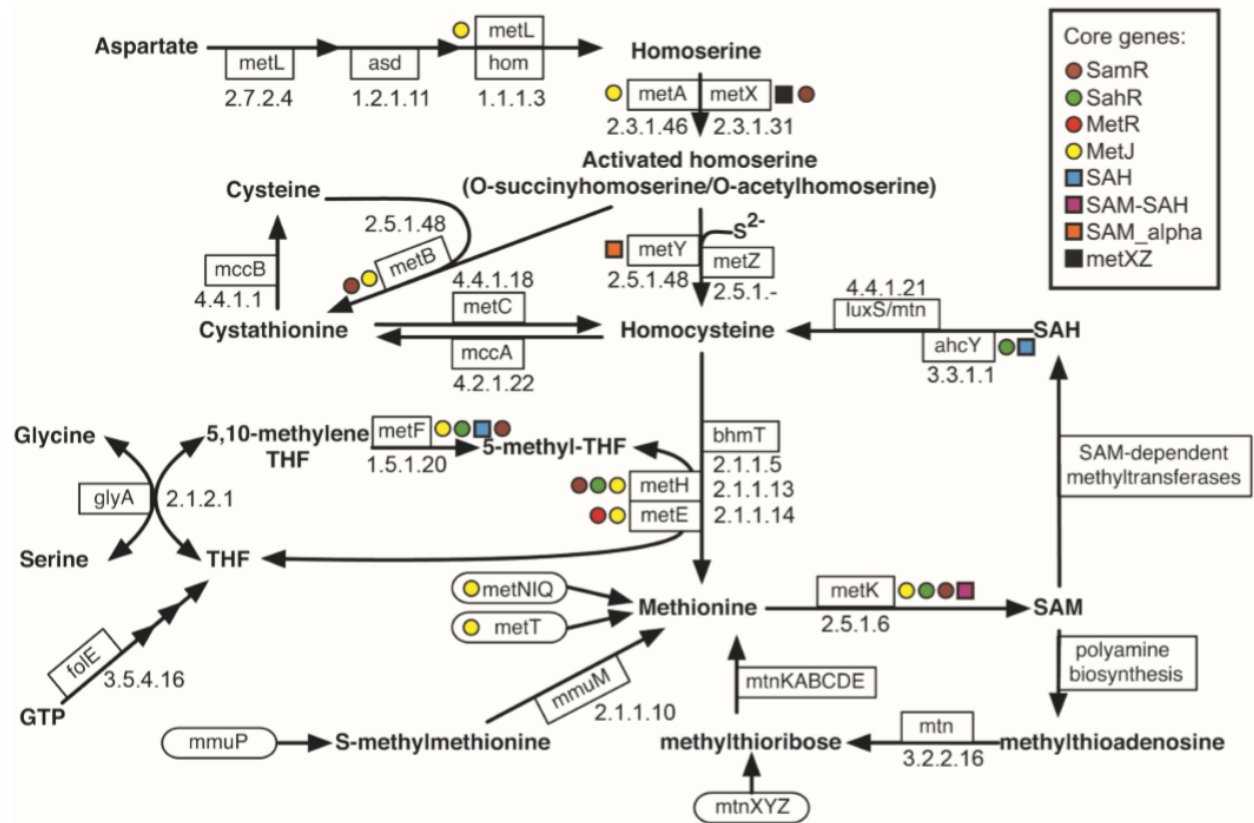


Figure 1.4 Régulation des voies métaboliques impliquées dans la synthèse de méthionine chez les Protéobactéries

Les enzymes sont encadrées et sont accompagnées de leur nomenclature de la Commission des enzymes. Encerclés sont les transporteurs de la méthionine, la méthylméthionine, ou la méthylthioribose. Les protéines régulatrices (de type facteurs de transcription) sont représentées par des cercles colorés et les riboswitchs sont représentés par des carrés colorés. (Image créée par Leyn SA, Suvorova IA, Kholina TD, Sherstneva SS, Novichkov PS, Gelfand MS, et al. © 2014 Leyn et al. <https://doi.org/10.1371/journal.pone.0113714> CC BY 4.0 license <http://creativecommons.org/licenses/by/4.0/> (Leyn et al., 2014))

1.2.1 Superfamille SAM-I

La superfamille de riboswitchs SAM-I comprend les familles SAM-I (S-box), SAM-IV, et SAM-I/IV (Price et al., 2014). Les trois familles comprises dans cette superfamille ont des pochettes de liaisons de SAM conservées qui se retrouvent au cœur de la jonction des nombreuses tiges de chacune des structures, comme illustré dans la Figure 1.5. Chacune des familles contient un ou deux sous-domaines périphériques qui sont caractérisés par une interaction de type pseudo-nœud (Trausch et al., 2014). Les riboswitchs SAM-I régulent l'expression génique principalement au niveau de la transcription grâce à la formation d'un terminateur rho-indépendant lorsque SAM est lié (Batey, 2011). Une analyse de l'arbre phylogénique de ces trois familles suggère que la famille SAM-IV a évolué à partir de la famille SAM I/IV et que la perte de la tige P4 trouvée dans

la structure de la famille SAM I/IV s'est produite à de multiples reprises dans cette famille, menant notamment à la famille SAM-IV qui n'a pas cette tige, et au moins une fois dans la famille de SAM-I, qui a normalement la tige P4. Ces cas, phylogénétiquement isolés, représenteraient donc une évolution convergente, c'est-à-dire dû à une évolution indépendante à travers le temps avec le même résultat (Tausch *et al.*, 2014).

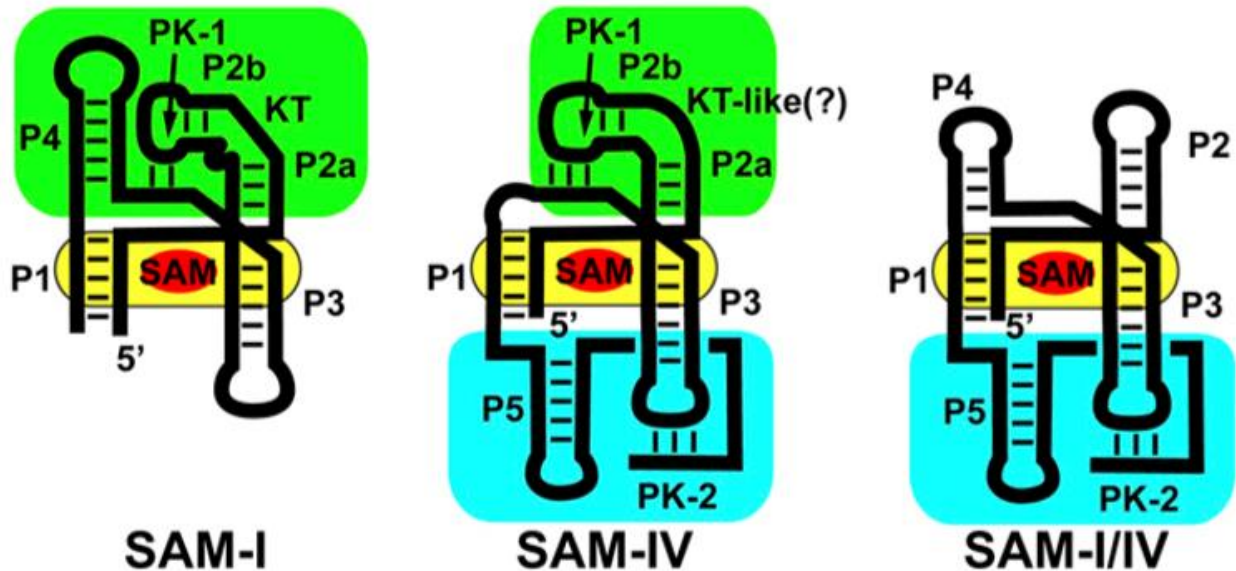


Figure 1.5 La superfamille SAM-I

Les trois familles de la superfamille de riboswitchs SAM-I sont représentées dans leurs états liés à SAM (rouge). Les pochettes de liaison conservées à travers les trois familles sont en jaune. Les tiges (P1-P5), les renflements de type kink-turn (KT), ainsi que les pseudo-nœuds (PK-1 et PK-2) sont annotés. Les deux sous-domaines comprenant chacun un pseudo-nœud sont colorés en vert et en bleu. (Image tirée de (Tausch *et al.*, 2014).)

1.2.2 Superfamille SAM-II

La superfamille de riboswitchs SAM-II comprend les familles de SAM-II et de SAM-V (Price *et al.*, 2014). Ces deux familles se ressemblent structurellement par la présence d'un pseudo-nœud de type H (Batey 2011), ou de type *hairpin loop*. Un pseudo-nœud de type H est une structure tertiaire d'ARN caractérisée par l'interaction de type Watson-Crick (WC) entre les nucléotides d'une boucle, faisant partie d'un élément structural tige-boucle, et des nucléotides situés dans une région simple brin à l'extérieure de cette boucle (Peselis & Serganov, 2014)(Figure 1.6). Les nucléotides conservés de la pochette de liaison se trouvent dans la partie L1 et P2 (Figure 1.7) (Batey, 2011). Les riboswitchs de SAM-II sont surtout retrouvés chez les α -Proteobacteria et ne se retrouvent jamais dans les génomes qui contiennent déjà le riboswitch SAM-I (Batey, 2011).

Pour un cas de riboswitch SAM-II qui a été résolu par la cristallographie aux rayons X, le SD se trouve dans la pochette de liaison de SAM ce qui suggère qu'il est possible que ce riboswitch régule l'expression génique par un mécanisme de séquestration du SD (Batey, 2011). Les riboswitchs SAM-II ont aussi été repérés en tandem avec le riboswitch SAM-V et dans cette configuration, un terminateur de transcription est toujours présent en aval du riboswitch SAM-II, suggérant un mécanisme de régulation transcriptionnel (Poiata *et al.*, 2009).

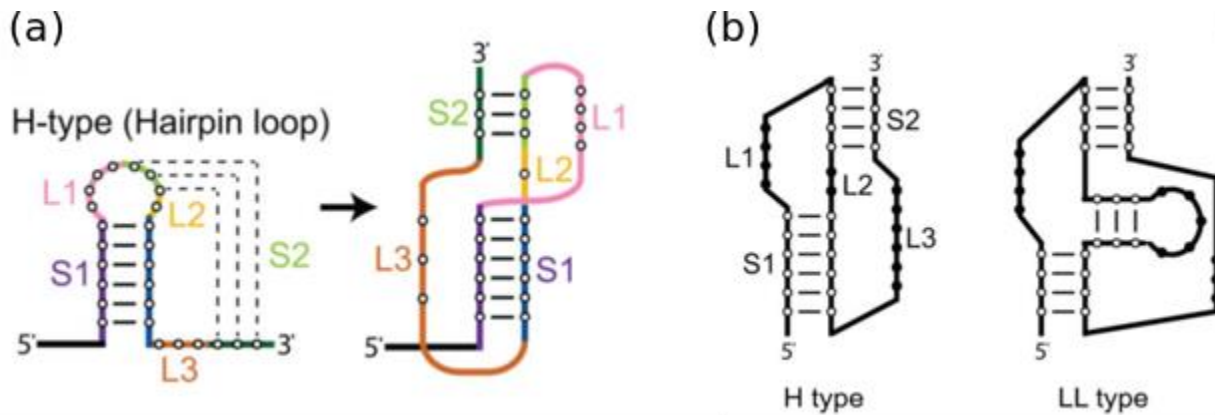


Figure 1.6 Les pseudo-nœuds faisant partie des riboswitchs SAM

En (a) l'interaction de type Watson-Crick des nucléotides de la boucle apicale d'une structure tige-boucle avec les nucléotides en simple brin à l'extérieur de cette boucle est représenté. Cette disposition mène à la formation d'un pseudo-nœud de type H, présent chez les riboswitchs de la superfamille SAM-II. En (b) un pseudo-nœud de type H est comparé à un pseudo-nœud de type LL. (Image tirée avec permission de (Peselis & Serganov, 2014) © 2014 John Wiley & Sons, Ltd.)

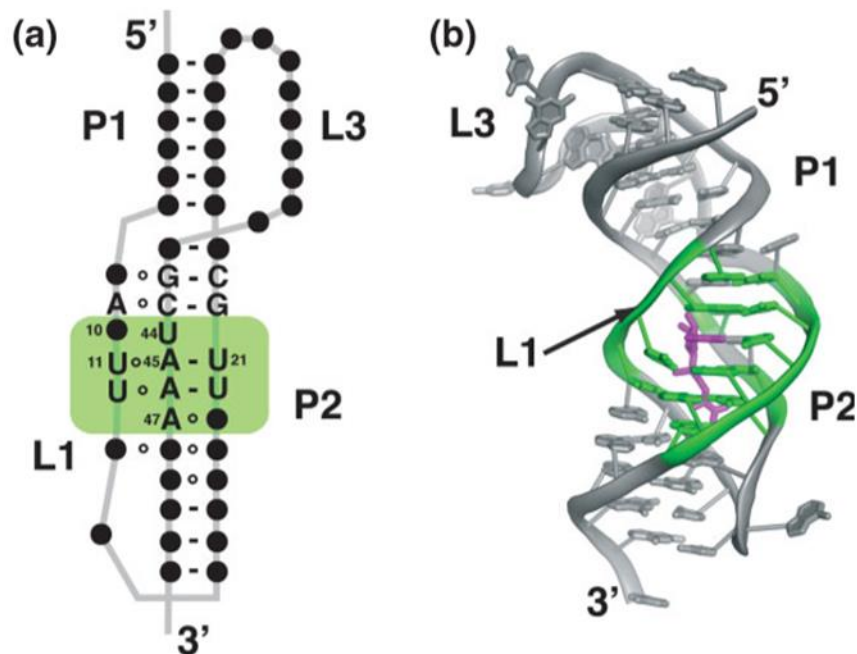


Figure 1.7 Riboswitch SAM-II de la superfamille SAM-II

En (a) la structure secondaire du riboswitch SAM II est représentée dans son état lié à SAM. Les bases dont l'identité est conservée à >97% sont montrées et le lieu de la pochette de liaison de SAM est en vert. En (b) la structure obtenue par la cristallographie aux rayons X de SAM-II est représentée. Le ligand SAM est en magenta. (Image tirée avec permission de (Batey, 2011) © 2011 John Wiley & Sons, Ltd.)

1.2.3 Superfamille SAM-III

Le riboswitch SAM-III, aussi nommé le motif SMK-box se trouve principalement chez les espèces *Enterococcus*, *Streptococcus*, et *Lactococcus* (Batey, 2011) et sa pochette de liaison pour SAM se trouve dans une jonction à trois tiges et dans une boucle interne voisine (Batey, 2011) (Figure 1.8). Ce riboswitch est un régulateur de traduction car la séquence SD se retrouve dans la partie aptamère et est séquestrée lorsque SAM est lié (Batey, 2011). De plus, récemment dans le groupe de Breaker (Mirihana Arachchilage *et al.*, 2018) une nouvelle famille, ayant une pochette de liaison similaire à celle du riboswitch SAM-III a été découverte et nommée SAM-VI (Mirihana Arachchilage *et al.*, 2018).

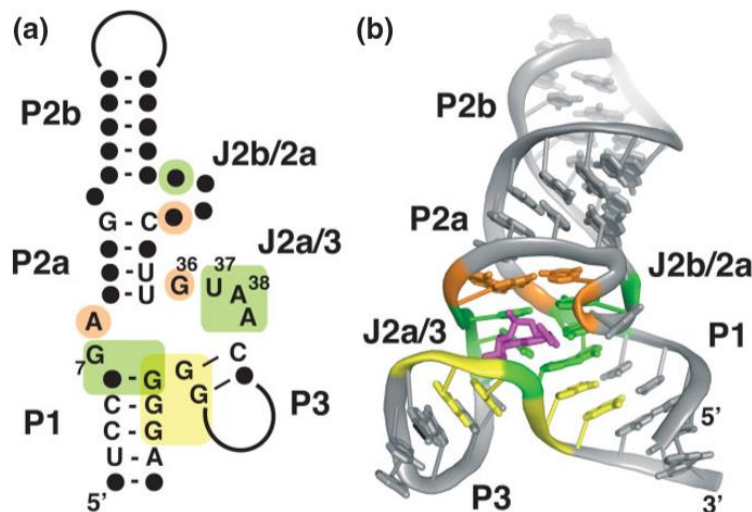


Figure 1.8 Riboswitch SAM-III

En (a) la structure secondaire du riboswitch SAM III est représentée dans son état lié à SAM. Les bases dont l'identité est conservée à >97% sont montrées et le lieu de la pochette de liaison de SAM est en vert. En (b) est une représentation de la structure obtenue par la cristallographie aux rayons X de SAM III. Le ligand SAM est en magenta. (Image tirée avec permission de (Batey, 2011) © 2011 John Wiley & Sons, Ltd.)

1.2.4 Riboswitchs SAM/SAH et SAH

Lorsque le groupement méthyle de la molécule SAM est utilisé par la cellule, le produit déméthylé, SAH, peut s'accumuler. Puisque les structures des deux molécules sont presque identiques (Figure 1.3), il y a une possibilité que SAH pourrait compétitionner avec SAM et inhiber certaines enzymes qui utilisent SAM (Batey, 2011). Le groupe de Breaker a découvert que le recyclage de SAH par les gènes *ahcY* et *metH* est régulé par un riboswitch chez certaines protéobactéries qui est spécifique pour SAH (Batey, 2011). Ce riboswitch se replie en une structure ayant un pseudo-nœud de type LL, qui est similaire au pseudo-nœud de type H mais qui contient aussi dans la boucle apicale de la structure tige-boucle—en aval des nucléotides qui forment les paires de

bases WC avec une région à l'extérieure de la boucle apicale—une structure tige-boucle supplémentaire (Peselis & Serganov, 2014) (Figure 1.6). L'affinité pour SAH de ce riboswitch est d'environ 20 nM (Batey, 2011). De plus, chez les *Rhodobacterales*, un motif en amont du gène *metK* qui se replie en un pseudo-nœud de type H, typiquement présent chez les riboswitchs SAM-II, a été découvert et son affinité a été confirmée pour SAM et SAH (Batey, 2011). Cette séquence d'ARN lie également SAH et SAM avec une affinité d'environ 10 μ M (Batey, 2011).

1.3 Les riboswitchs fluor

Des motifs structurels conservés correspondant à des riboswitchs ont été découverts chez les bactéries et les archées en amont des gènes *crcB* et *eriC*. Ces gènes codent pour des protéines antiports régulant le transport du fluor et lorsqu'ils sont régulés par un riboswitch sont activés en présence de cet ion (Baker *et al.*, 2012; Stockbridge *et al.*, 2012; Weinberg *et al.*, 2010). Il est suggéré que chez certaines bactéries comme chez *Bacillus cereus* le riboswitch fluor en amont du gène *crcB* régulerait au niveau transcriptionnel tandis que chez d'autres bactéries tel que *P. syringae* la régulation se passe au niveau de la traduction (Baker *et al.*, 2012).

1.4 Méthodes de recherche et de validation des riboswitchs

La recherche et la validation de nouveaux riboswitchs se fait grâce à de nombreuses stratégies impliquant un premier tri des séquences d'intérêt, habituellement par génomique comparative *in silico*, suivi par une validation de l'affinité et de la spécificité de l'aptamère pour son ligand par des réactions *in vitro*, et terminé par la confirmation de la fonction du riboswitch à travers des essais *in vivo* dans un micro-organisme modèle.

1.4.1 Méthodes *in silico*

La recherche de nouveaux riboswitchs commence par l'analyse des séquences des régions intergéniques (IGR) conservées provenant de diverses espèces bactériennes (Sun & Rodionov, 2014). La nature de l'opéron en amont duquel un riboswitch potentiel se retrouve peut donner un indice sur sa fonction biologique. La conservation des ribonucléotides de la pochette de liaison d'un ligand a permis de repérer une multitude de riboswitchs dans les génomes bactériens (Sun & Rodionov, 2014). Une caractéristique encore plus importante pour l'identification de nouveaux riboswitchs, et aussi une caractéristique cible lors des recherches *in silico*, est la conservation des éléments structuraux. Cette conservation issue de l'appariement ou non-appariement de paires de bases WC ou Wobble qui constituent les tiges et les boucles, peut être un indice de la

présence d'un ARN régulateur. Une analyse comparative de génomes recherchant la conservation des séquences et en même temps la conservation des éléments structuraux à l'intérieur des IGR permet d'établir un modèle de covariance pour un riboswitch donné (Sun & Rodionov, 2014). CMfinder (Yao et al., 2006) est un outil bio-informatique qui permet le criblage *in silico*, à grande échelle, des variations phylogénétiques d'un riboswitch. Le résultat de ce criblage permet de créer de nouveaux modèles de covariance, qui ensuite, peuvent servir à des nouvelles recherches utilisant d'autres outils bio-informatiques comme Infernal, dans le but de trouver d'autres exemplaires de cet ARN (Sun & Rodionov, 2014).

Infernal (Nawrocki & Eddy, 2013) est un outil bio-informatique qui construit un profil de séquences, le modèle de covariance, en tenant compte d'un consensus de séquences primaires ainsi que d'un consensus des structures secondaires, qui est issu des interactions entre les bases des séquences primaires (Sun & Rodionov, 2014). Bien qu'il soit l'outil de choix pour annoter les ARNnc, Infernal ne tient pas compte des pseudo-nœuds. Une limite du logiciel Infernal est la nécessité d'une pré-connaissance de la structure de l'ARN qu'on cherche à aligner lorsque la séquence contient des éléments structuraux qui ont une faible conservation de séquence (Sun & Rodionov, 2014).

L'analyse *in silico* des séquences d'ARN à l'aide des logiciels bio-informatiques qui prédisent les structures les plus stables thermodynamiquement, comme Sfold (Ding *et al.*, 2004), Mfold (Zuker, 2003) ou MC-Fold (Parisien & Major, 2008) peut aussi aider à construire des hypothèses sur les mécanismes de régulation pour un ARN structuré. Sfold utilise un algorithme pour échantillonner les structures secondaires d'ARN de l'ensemble de Boltzmann—une distribution largement utilisée pour décrire une variété de systèmes physiques en équilibre thermodynamique—et génère un échantillon statistique qui représente l'ensemble des populations structurales et leurs probabilités statistiques, pour une séquence d'ARN donné, à une température de 37°C (Ding *et al.*, 2004). Parmi d'autres informations, Sfold génère la structure centroïde de l'ensemble échantillonné—cette structure représente la tendance centrale de l'échantillon en repérant la structure qui a le plus de similarité avec les autres structures de l'ensemble (Ding *et al.*, 2005). Le centroïde est une alternative à la structure d'énergie minimal (MFE) qui est prédit par l'algorithme de Mfold, un autre outil bio-informatique pour prédire les structures des séquences d'ARN (Zuker, 2003). Une limite de Sfold et de Mfold est que ces deux outils ne tiennent pas en compte les possibilités de repliement des pseudo-nœuds. Pour les prédictions des structures contenant des pseudo-nœuds un outil comme MC-Fold peut être utilisé (Parisien & Major, 2008).

L'arrangement spatial des motifs structuraux des plateformes d'expression des riboswitchs vis-à-vis la séquence SD et des potentiels motifs de terminateurs de transcription sont des indices potentiels pour la prédiction des mécanismes de régulation (Sun & Rodionov, 2014).

1.4.2 Méthodes *in vitro*

De nombreuses stratégies *in vitro* existent pour valider les domaines aptamères des riboswitchs. La caractéristique recherchée des séquences hypothétiques des riboswitchs est le changement de conformation causé par l'ajout d'un ligand. Des techniques ayant recours à des réactifs chimiques qui réagissent selon les structures locales peuvent être utilisés pour valider les changements structuraux qui se produisent lors de la liaison du ligand. Par exemple, des réactifs de *Selective 2'-hydroxyl acylation analyzed by primer extension* (SHAPE), tel que le N-méthylisatoïc (NMIA), réagissent avec le groupement 2'hydroxyl de l'ARN, préférentiellement aux positions non-appariées (Wilkinson *et al.*, 2006) sans spécificité pour la séquence (Mlýnský & Bussi, 2017). Un autre exemple est le réactif sulfate de diméthyle (DMS) qui réagit avec les adénosines et les cytidines non-appariées (Tijerina *et al.*, 2007). La détection de ces réactions préférentielles se fait par transcription inverse suivie par l'extension des amorces et la migration par électrophorèse sur gel polyacrylamide (Wilkinson *et al.*, 2006).

Comme SHAPE, la technique de *in-line probing* permet de visualiser les sites de clivage spontanés du squelette phosphodiester indépendamment de la séquence en nucléotides de l'ARN. Cette technique peut être appliquée à travers une plage de concentrations du ligand après incubation de l'ARN dans un tampon approprié pour ensuite permettre de repérer les bases non-appariées de la séquence en présence et en absence du ligand. La réaction de transestérification (Mlýnský & Bussi, 2017), aussi nommée la réaction de *in-line*, se manifeste surtout aux positions non-appariées puisque le squelette de l'ARN est plus flexible et il y a une plus grande probabilité que l'oxygène en 2' s'aligne avec le phosphore et l'oxygène 5' du ribose adjacent, pour que l'attaque nucléophile par l'oxygène en 2' ait lieu sur le phosphore (Regulski & Breaker, 2008) (Figure 1.6). Les différents motifs de clivage pour un même ARN incubé avec des concentrations différentes de ligand peuvent être visualisés grâce à une migration de l'ARN, préalablement marqué, sur un gel dénaturant de polyacrylamide. Cette technique permet non seulement de quantifier la modulation de clivage pour déterminer l'affinité de l'aptamère pour son ligand mais aussi de cartographier les positions de changement de conformation de simple brin en double brin (Regulski & Breaker, 2008).

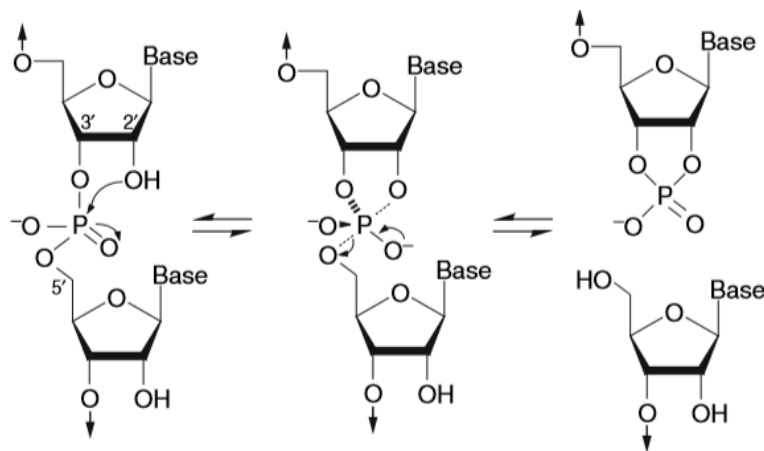


Figure 1.9 Clivage spontané de l'ARN

Lorsque l'oxygène 2' s'aligne avec le phosphore et l'oxygène 5', une attaque nucléophile prend place causant une rupture de l'ARN entre les deux nucléotides. (Image tirée avec permission de (Regulski & Breaker, 2008) https://doi.org/10.1007/978-1-59745-033-1_4 Copyright © 2008, Humana Press, a part of Springer Science Business Media, LLC.)

1.4.3 Méthodes *in vivo*

La dernière étape de validation d'un nouveau riboswitch est la vérification de sa capacité de fonctionner comme un régulateur génique à l'intérieur d'un organisme vivant (Regulski & Breaker, 2008). Pour ceci, des essais *in vivo* sont conçus pour confirmer l'induction ou la répression de l'expression d'un gène rapporteur lorsque le ligand est introduit à l'intérieure de la cellule. Le riboswitch est fusionné de façon à conserver le cadre de lecture de la protéine du gène rapporteur dans une cassette appropriée pour ensuite être transformé dans un organisme hôte. Le ligand, ou un précurseur du ligand, est introduit dans le milieu de culture de l'organisme hôte contenant la construction et l'émission du signal détectable, soit la fluorescence, la luminescence ou la production de la β -galactosidase, est quantifié pour des milieux de culture contenant différentes concentrations du ligand. Il est important de conserver l'environnement génique du riboswitch lors de la construction du vecteur rapporteur pour assurer le bon repliement du riboswitch.

1.5 Validation d'un nouveau variant du riboswitch SAM

1.5.1 Problématique

Des travaux récents dans le laboratoire (inclus dans Sellamuthu et al. dans ce mémoire) ont révélé l'existence de nouvelles structures d'ARN conservées en amont des gènes faisant partie du métabolisme du soufre, de la méthionine, et de SAM dans les génomes d'une multitude de β -

protéobactéries, suggérant la présence d'un potentiel riboswitch SAM. L'élément structural clé constitue un pseudo-nœud de type 'H' tel que retrouvé dans les structures consensus du riboswitch SAM-II et du riboswitch SAM-V. En revanche, la tige de base qui participe à la formation du pseudo-nœud est d'une longueur plus importante comparativement à celle des riboswitchs SAM-II et SAM-V déjà connus. De plus, dans plusieurs espèces une tige accessoire d'une longueur de 4 à 6 nt terminée par une boucle de longueur variable est présente. Cette tige accessoire n'a pas été rapportée précédemment. Dans le but de caractériser et de comprendre la régulation de ces nouvelles structures, des analyses *in vitro* et *in vivo* sont nécessaires. À cet effet, à notre connaissance, la littérature ne décrit pas de vecteurs conçus pour des fusions traductionnelles compatibles avec les bactéries du genre *Burkholderia*.

1.5.2 Hypothèse et objectif

L'objectif général de cette étude est de caractériser et de valider les nouvelles variantes du riboswitch SAM-II/V par l'évaluation de leurs affinités pour SAM et par la vérification de leur fonction comme régulateurs génique dans un organisme modèle, soit *Burkholderia thailandensis* E264. L'hypothèse est que les différentes versions du riboswitch sont des adaptations différentes, vraisemblablement pour une régulation correspondant à chacun des gènes trouvés en aval. Pour évaluer cette hypothèse, l'objectif principal a été subdivisé en trois sous-objectifs, décrits ci-bas avec leur hypothèse sous-jacente.

1- Mettre au point un plasmide rapporteur d'expression pour des fusions traductionnelles compatible avec *Burkholderia*.

Le rapporteur choisi est celui de la luciférase, qui permettra d'estimer l'expression en temps réel. Nous émettons l'hypothèse que cet avantage permettra de voir des différences autrement difficiles à percevoir. Ce rapporteur sera d'abord validé pour l'utilisation avec un riboswitch connu, soit celui du fluor. Une des particularités de ce riboswitch est que son ligand, l'ion F^- , ne fait pas partie du métabolisme, sa concentration peut donc plus facilement être contrôlée.

2- Valider expérimentalement les riboswitchs découverts par bio-informatique.

L'hypothèse pour la caractérisation *in vitro* est que les différentes constructions du riboswitch SAM provenant de plus de dix organismes différents démontreront des affinités apparentes pour SAM, lorsque soumis à des réactions de clivage spontané en présence de SAM.

3- Déterminer de possibles particularités de la modulation médiée par différents riboswitchs.

Comme chez *B. thailandensis* E264 il existe trois variantes de nouveaux riboswitchs SAM, soit en amont des gènes *metK*, *metX*, et *metZ*—ces riboswitchs ont été étudiés pour mieux comprendre leurs rôles respectifs. L'hypothèse de la régulation *in vivo* est que les trois versions du riboswitch SAM fusionnés à leurs promoteurs respectifs, chez *B. thailandensis* E264 auront différentes dynamiques de régulation en temps réel du gène rapporteur de la luciférase d'une culture en milieu liquide. Ces différences pourraient être au niveau de l'allure des courbes d'expression à travers le temps de l'essai, ce qui peut impliquer: le degré d'induction (*fold induction*), la chronologie de l'induction, le niveau de base, le maximum, le nombre de pics et/ou l'aire totale sous la courbe.

4- Tester l'importance de composantes structurales propres à l'aptamère du riboswitch SAM en amont de *metK* chez *B. thailandensis* E264 sur l'affinité pour SAM et sur la régulation génique.

Des mutations ayant pour but de soit déstabiliser la tige de base, la tige accessoire ou de supprimer complètement la tige accessoire devraient exhiber un changement d'affinité pour SAM lorsque comparé avec le construit de type sauvage (*WT*).

2 BIBLIOGRAPHIE DU CHAPITRE 1

- Baker JL, Sudarsan N, Weinberg Z, Roth A, Stockbridge RB & Breaker RR (2012) Widespread genetic switches and toxicity resistance proteins for fluoride. *Science* 335(6065):233-235.
- Batey RT (2011) Recognition of S-adenosylmethionine by riboswitches. *Wiley Interdiscip Rev RNA* 2(2):299-311.
- Cech TR & Steitz JA (2014) The noncoding RNA revolution-trashing old rules to forge new ones. *Cell* 157(1):77-94.
- Ding Y, Chan CY & Lawrence CE (2004) Sfold web server for statistical folding and rational design of nucleic acids. *Nucleic Acids Res* 32(Web Server issue):W135-141.
- Ding Y, Chan CY & Lawrence CE (2005) RNA secondary structure prediction by centroids in a Boltzmann weighted ensemble. *RNA* 11(8):1157-1166.
- Hor J, Gorski SA & Vogel J (2018) Bacterial RNA Biology on a Genome Scale. *Mol Cell* 70(5):785-799.
- Leistra AN, Curtis NC & Contreras LM (2019) Regulatory non-coding sRNAs in bacterial metabolic pathway engineering. *Metabolic Engineering* 52:190-214.
- Leyn SA, Suvorova IA, Kholina TD, Sherstneva SS, Novichkov PS, Gelfand MS & Rodionov DA (2014) Comparative genomics of transcriptional regulation of methionine metabolism in Proteobacteria. *PLoS One* 9(11):e113714.
- Lotz TS & Suess B (2018) Small-Molecule-Binding Riboswitches. *Microbiol Spectr* 6(4).
- Mirihana Arachchilage G, Sherlock ME, Weinberg Z & Breaker RR (2018) SAM-VI RNAs selectively bind S-adenosylmethionine and exhibit similarities to SAM-III riboswitches. *RNA Biol* 15(3):371-378.
- Mlýnský V & Bussi G (2017) Understanding In-line Probing Experiments by Modeling Cleavage of Non-reactive RNA Nucleotides. *RNA* 23(5):712-720.
- Moore PB (1999) Structural Motifs in RNA. *Annu. Rev. Biochem.* 68(1):287-300.
- Parisien M & Major F (2008) The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* 452(7183):51-55.
- Peselis A & Serganov A (2014) Structure and function of pseudoknots involved in gene expression control. *Wiley Interdiscip Rev RNA* 5(6):803-822.
- Poiata E, Meyer MM, Ames TD & Breaker RR (2009) A variant riboswitch aptamer class for S-adenosylmethionine common in marine bacteria. *RNA* 15(11):2046-2056.
- Price IR, Grigg JC & Ke A (2014) Common themes and differences in SAM recognition among SAM riboswitches. *Biochim Biophys Acta* 1839(10):931-938.
- Ramírez-Clavijo S & Montoya-Ortíz G (2013) *Gene expression and regulation*. El Rosario University Press. <https://www.ncbi.nlm.nih.gov/books/NBK459456/files/724/NBK459456.html> (Consulté le 2020/06/28/17:44:23)
- Regulski EE & Breaker RR (2008) In-Line Probing Analysis of Riboswitches. *Post-Transcriptional Gene Regulation*, Wilusz J (Édit.) Humana Press, Totowa, NJ Vol 419. p 53-67.

- Santiago-Frangos A, Kavita K, Schu DJ, Gottesman S & Woodson SA (2016) C-terminal domain of the RNA chaperone Hfq drives sRNA competition and release of target RNA. *Proc Natl Acad Sci USA* 113(41):E6089-E6096.
- Stockbridge RB, Lim HH, Otten R, Williams C, Shane T, Weinberg Z & Miller C (2012) Fluoride resistance and transport by riboswitch-controlled CLC antiporters. *Proc Natl Acad Sci U S A* 109(38):15289-15294.
- Sun EI & Rodionov DA (2014) Computational analysis of riboswitch-based regulation. *Biochim Biophys Acta* 1839(10):900-907.
- Tijerina P, Mohr S & Russell R (2007) DMS footprinting of structured RNAs and RNA-protein complexes. *Nat Protoc* 2(10):2608-2623.
- Trausch JJ, Xu Z, Edwards AL, Reyes FE, Ross PE, Knight R & Batey RT (2014) Structural basis for diversity in the SAM clan of riboswitches. *Proc Natl Acad Sci U S A* 111(18):6624-6629.
- Vogel J & Luisi BF (2011) Hfq and its constellation of RNA. *Nat Rev Microbiol* 9(8):578-589.
- Weinberg Z, Wang JX, Bogue J, Yang J, Corbino K, Moy RH & Breaker RR (2010) Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol* 11(3):R31.
- Wilkinson KA, Merino EJ & Weeks KM (2006) Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat Protoc* 1(3):1610-1616.
- Wittmann A & Suess B (2012) Engineered riboswitches: Expanding researchers' toolbox with synthetic RNA regulators. *FEBS Lett* 586(15):2076-2083.
- Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31(13):3406-3415.

3 RIBOSWITCH-MEDIATED EXPRESSION DAMPENING

Fluoride riboswitch-based dampening of gene expression for cloning potent promoters

Diminution de l'expression génique via le riboswitch fluor pour le clonage des promoteurs forts

Auteurs :

Vesta Korniakova¹, Marie-Christine Groleau¹, Eric Déziel¹ et Jonathan Perreault¹

¹INRS - Centre Armand-Frappier Santé Biotechnologie, 531 boulevard des Prairies, Laval, QC, H7V 1B7, Canada

Titre de la revue ou de l'ouvrage :

Frontiers in Genetics

Date de soumission : 4 août 2020

Contribution des auteurs :

Vesta Korniakova a conçu et construit les plasmides, exécuté tous les essais *in vivo*, analysé les résultats et a contribué à la rédaction de l'article.

Marie-Christine Groleau et Éric Déziel ont partagé des conseils techniques lors des expériences *in vivo* ainsi que des souches bactériennes, et ont contribué à la révision de l'article.

Jonathan Perreault a guidé la conception des expériences et a contribué à la rédaction de l'article.

3.1 Abstract

Bioreporter systems based on detectable enzyme activity, such as that of beta-galactosidase or luciferase, are key in novel bacterial promoter discovery and study. While these systems permit quantification of gene expression, their use is limited by the toxicity of the expressed reporter enzymes in a given host. Indeed, the most potent promoters may be overlooked if their activity causes a lethal overproduction of the reporter genes when screening for transcriptional activity of potential promoter sequences with the *luxCDABE* cassette. To overcome this limitation, a variation of the mini-CTX-*lux* plasmid has been designed which allows reduction of promoter activity via the addition of an adjacent fluoride riboswitch. The riboswitch adds a layer of regulation between the promoter and the reporter gene, allowing cloning of stronger promoters by weakening expression, while still providing a good signal for weaker promoters by inducing with fluoride, thus circumventing limitations associated with reporter toxicity. We noticed some unexpected behavior of the riboswitch in terms of portability between species, suggesting caution when using riboswitches non-native to the species where it is being used. This study introduces a new molecular biology tool which will allow for the identification of previously unverifiable or uncharacterized potent promoters and also provides a cloning vector for translational fusion with luciferase in a plasmid compatible with many species such as from the genera *Burkholderia* and *Pseudomonas*.

Key words: plasmid, reporter, promoter, regulatory region, fluoride riboswitch, *Burkholderia*, *Pseudomonas*, *E. coli*, luciferase, toxicity, cloning, induction, repression, translation fusion

3.2 Introduction

Reporter genes encoding for proteins which are easily detectable through sensitive and simple means (colorimetry, fluorescence, luminescence) are key elements to numerous gene expression studies and critical to decipher regulatory elements, including the discovery of new promoters and their characterization in terms of strength and dynamics. Common reporter proteins include β -galactosidase, Green fluorescent protein (GFP) and luciferase; detected either by spectrophotometry, fluorimetry or luminometry, respectively. As a rule of thumb, when gene regulatory elements are cloned upstream of a reporter gene, a high reporter protein signal indicates a strong promoter, while a low signal is attributed to a weak promoter. Strategies have been developed to allow for weak promoter detection and characterization via reporter gene assays (Guo *et al.*, 2019), however to our knowledge, no strategy for detection and study of circumstantially lethal potent promoters, has been suggested.

While current strategies allow for the discovery of weak promoters (Guo *et al.*, 2019), several screening approaches using classical gene reporter assays may have biases against the most potent promoters. The toxicity caused by overexpression of reporter proteins could inhibit the growth of potential clones causing an important gap in new promoter discovery.

Many DNA cloning experiments are not successful and are deemed to be technical mysteries. This failure may appear initially as a cloning gap or a disassembly in sequencing results or as an absence of transformed colonies for a given construct in a species of interest other than the shuttle species. A possible reason for these failures may be that an overexpression of the detectable protein in the designed construct has caused a lethal metabolic burden for the cell and thus an absence of viable construct-validated clones. Previously, it has been shown through genome sequencing of clone-based assemblies that many occurring cloning gaps were not technical failures but rather a consequence of the sequences coding for toxic products (Kimelman *et al.*, 2012). As cloning and transformation experiments often involve propagating the construct across different species, constructs must be viable across all species being manipulated in order for an experiment to be successful. While toxicity level thresholds of different reporter genes, their substrates, or by-products vary depending on the host organism, overexpression of any protein can potentially be toxic (Bolognesi & Lehner, 2018). In fact, toxicity has been previously reported for luciferase substrate N-decyl aldehyde in *Saccharomyces cerevisiae* and *Caenorhabditis elegans* (Hollis *et al.*, 2001); for constitutive expression of *Gaussia princeps* luciferase (Gluc) in *Escherichia coli* (Liu *et al.*, 2014); for β -galactosidase expression in *E. coli* under osmotic stress (Malakar *et al.*, 2014); and for GFP in *S. cerevisiae* (Kintaka *et al.*, 2016). One study concluded that for a number of proteins, the overexpression burden limit in *S. cerevisiae* is achieved for normally non-harmful proteins once it constitutes up to 15% of the total cellular proteins (Eguchi *et al.*, 2018).

The goal of the cloning strategy described in this paper is to decouple the cloning and transformation process from the evaluation of promoter activity in a host organism. To the best of our knowledge, no such strategy has previously been described. To this end we believe riboswitches, ligand-specific RNA *cis*-acting gene regulatory elements, may be important tools for dampening the strength of potent constitutive or potent uncharacterized inducible promoters. By

sandwiching an appropriate riboswitch between the potent promoter and the reporter gene, expression levels may be controlled and adjusted down to viable levels during promoter screening or characterization assays (Figure 2.1A). Many expression vectors are designed for protein induction via inducible promoters to avoid toxicity of the protein to be purified during cloning and growth, however to our knowledge no vector exists for controlled promoter read-out via inducible 5' UnTranslated Regions (UTRs).

When choosing the right promoter-dampening riboswitch for an experiment it is important to consider its compatibility with the shuttle and host species, its regulation range dynamics including repression capacity, fold induction (FI) and regulation mechanism, and whether or not the trigger ligand is endogenous to the species and what systems exist, if any, to control ligand concentrations inside the cell. For this study, we identified the fluoride riboswitch (F RS) as a potential candidate for mitigating the potency of two promoters to be used in a reporter gene system: the constitutive S7 ribosomal protein gene promoter (P_{S7}) from *Paraburkholderia xenovorans* strain LB400 and the P1 integron promoter originally from R388, a trimethoprim-resistance broad-host-range plasmid (DeShazer & Woods, 1996; Zolg & Hanggi, 1981). These promoters were meant to be cloned upstream of the *lux* cassette in a strategically redesigned version of mini-CTX-*lux* (Becher & Schweizer, 2000), a high copy plasmid in *E. coli* or a single copy chromosomal integration plasmid for *Burkholderia thailandensis* and *Pseudomonas aeruginosa*. The fluoride riboswitch has the advantage of controlling gene expression according to concentrations of fluoride, a non-cellular metabolite. This conserved RNA structure is widespread across bacteria and archaea and is known to upregulate, in the presence of fluoride ions, the expression of proteins which manage its exportation, such as the CrcB proteins and the CLC proteins, fluoride-specific channels which act as fluoride/proton antiporters (Baker *et al.*, 2012; Stockbridge *et al.*, 2012; Weinberg *et al.*, 2010). Additionally, the fluoride riboswitch from a thermophilic archaeon has been previously used as an alternative strategy to inducible promoters for regulating gene expression in hyperthermophiles (Speed *et al.*, 2018).

Our redesigned plasmid, which we named pVK-f-*lux*, features optimal cloning features for allowing to easily swap promoters and 5'UTRs as needed in order to find the right combination for a particular experiment. Additionally, it is optimized to allow for scar-free translational fusion cloning, a feature not often present in reporter vectors, but essential for studying many *cis*-regulatory RNA

elements. Our plasmid is designed for Gibson assembly (GA) cloning but may also be used with a restriction enzyme (RE) digestion and ligation approach.

3.3 Methods

3.3.1 DNA amplification and reporter plasmid construct assembly.

Oligonucleotides from Integrated DNA Technologies (25 nmoles DNA oligonucleotides and 500 ng of gBlocks® Gene Fragments) were used. DNA parts for GA cloning were amplified using the Q5® High-Fidelity DNA Polymerase (New England Biolabs) using a touchdown-gradient PCR protocol as previously described (Korbie & Mattick, 2008) and appropriate primers and template as specified in Table S1. Touchdown annealing cycles (-1.2°C/cycle for 10 cycles) were performed from 71°C down to 60.2°C and were followed by 20 cycles of constant annealing temperatures (with five tubes in a gradient from 55°C to 72°C). Reactions with the most specificity were chosen for further cloning steps. Backbone vectors, as per specific cloning attempts (Annexe I: Table S1 and Table S2) were linearized using restriction enzymes (Thermo Fisher Scientific) as described. GA cloning was carried out using diluted PCR products, unpurified restriction digestion products and the NEBuilder® HiFi DNA Assembly Master Mix kit (New England Biolabs) according to the manufacturer protocol.

3.3.2 Bacterial strains and construction of reporter strains.

All strains and clones used in this study are enumerated in Annexe I: Table S3. All strains were grown at 37°C on a rotary agitator in liquid Luria Broth (LB) (Alpha Biosciences) or on petri dishes of LB agar (Alpha Biosciences). GA-cloned plasmid constructs were transformed into either *E. coli* strain DH5 α or strain SM10 λ pir as follows. Either 2 μ L of the GA reaction or 25 ng of the plasmid of interest was added to 100 μ L of thawed chemically competent cells on ice and incubated for 20 min. A thermal shock was performed for 40 seconds at 42°C followed by a 3-minute incubation on ice. 300 μ L of LB was added to the mix and cells were incubated at 37°C for 1 hour with agitation at 250 rpm. 150 μ L of cells were spread on a prewarmed selection plate and incubated overnight at 37°C.

The constructs were integrated into the chromosome of *B. thailandensis* E264 by bi-parental conjugation with *E. coli* SM10 λ pir as follows. Pellets from 1.5 mL of overnight cultures diluted to 0.5 OD₆₀₀ for both *E. coli* SM10 λ pir donor strains and for *B. thailandensis* E264 WT strain were

obtained by centrifugation at 7000 x g. Each pellet was resuspended in 25 μ L of LB and pooled into a single drop on an antibiotic-free LB agar dish for overnight incubation at 37°C. The resulting growth was resuspended in 1 mL of liquid LB using a sterile Q-tip and 100 μ L was plated on LB agar selection plates for *B. thailandensis* E264 (see below).

Liquid and solid selection media were supplemented with 15 μ g/mL tetracycline for *E. coli* strains; and with 25 μ g/mL tetracycline, 50 μ g/mL gentamycin and 15 μ g/mL polymyxin for *B. thailandensis* E264. When required, FH₄KO₂ was added to Luria agar selection plates or to liquid media in concentrations ranging from 0 to 31 mM. Transformed reporter strains were verified for luminescence signal using a microplate reader (Cytation 3; BioTek Instruments, Inc.). Plasmids were extracted from transformed *E. coli* strains using the Presto™ Mini Plasmid Kit (Geneaid) and sequences were confirmed by Sanger sequencing carried out at Genome Quebec (Montreal, Canada).

3.3.3 Testing Gibson assembly-based cloning in the plasmid

The designed overhangs of the 5'UTR DNA sequence which excludes the promoter region, were tested for compatibility for GA with a *Scal* and *AarI* digested backbone, by carrying out GAs with inserts containing the *NcoI-SdaI-Scal* left overlap sequence of 17 nts and the *AarI/luxC* right overlap sequence of 21-24 nts.

3.3.4 lux reporter assay

To assess time-course riboswitch regulation dynamics in bacteria, *E. coli* DH5 α , *E. coli* SM10 and *B. thailandensis*, strains transformed with constructs of interest containing the P1 promoter and a 5'UTR translational fusion with the bacterial operon *luxCDABE* were first cultured overnight in LB supplemented with the same antibiotic composition as during transformation. Next, cells were pelleted by centrifugation at 15000 x g for 3 minutes and washed twice with M9 Minimal Media (M9-MM). Cells were then suspended in fresh M9-MM. M9-MM was prepared by combining 200 mL of sterile M9 salts (64 g/L Na₂HPO₄-7H₂O, 15 g/L KH₂PO₄, 2.5 g/L NaCl, 5.0 g/L NH₄Cl with 2 mL of sterile 1M MgSO₄, 20 ml of 20% glucose and 100 μ L of sterile 1M CaCl₂ in a total volume of 1000 mL of sterile reverse osmosis water. Assays were carried out in 96-well microplates from Greiner Bio-One (Microplate, 96 well, PS F-bottom [chimney well], white, med. binding Ref: 655095). Each well contained a total culture volume of 200 μ L of antibiotic supplemented 0.5X

M9-MM. Cultures were adjusted to an initial optical density at 600 nm (OD₆₀₀) of approximately 0.06. Luminescence and OD₆₀₀ readings were recorded at 20-minute intervals with a multi-mode microplate reader (Cytation 3; BioTek Instruments, Inc.) for assay total run times between 30 and 60 hours.

3.3.5 Luciferase reporter data analysis

Average blank values for each time-point were subtracted from each corresponding well sample reading. All readings were cropped to start as soon as the OD₆₀₀ reading hit 0.1 for an individual well. For each assay, the number of time points used was adjusted to be the same for all samples unless otherwise stated. For determining the plateau OD₆₀₀ value, the average value of the last 58 OD₆₀₀ readings of a time-course luciferase assay was calculated.

Fold Induction (FI) of total luciferase activity:

$$FI \text{ of lum total} = \frac{\sum_n^N f^{-ligand}(n)}{\sum_n^N f^{+ligand}(n)}$$

Equation 1

where:

$$f^{-ligand}(n) = \frac{\sum_n^N lum^{-ligand}(n)}{\sum_n^N OD^{-ligand}(n)}$$

Equation 2

$$f^{+ligand}(n) = \frac{\sum_n^N lum^{+ligand}(n)}{\sum_n^N OD^{+ligand}(n)}$$

Equation 3

Where *ligand* is fluoride ions (F⁻);

lum^{-ligand}(n) is the luminescence reading for a culture containing a P1+ F RS-*lux* construct of interest in absence of supplementary fluoride for the time point *n*.

lum^{+ligand}(n) is the luminescence reading for a culture containing a P1+ F RS-*lux* construct of interest in presence of supplementary fluoride for the time point *n*.

$OD^{\text{ligand}}(n)$ is the OD_{600} reading for a culture containing a P1+ F RS-*lux* construct of interest in absence of supplemented fluoride for the time point n .

$OD^{+\text{ligand}}(n)$ is the OD_{600} reading for a culture containing a P1+ F RS-*lux* construct of interest in presence of supplemented fluoride for the time point n .

Peak lum/ OD_{600} :

The maximum peak luminescence value, in Relative Luminescence Units (RLU), was determined for a time-course reporter assay. This value was divided by the corresponding time-point OD_{600} .

Average lum/ OD_{600} :

The average RLU value for all time points of each triplicate was divided by the average OD_{600} value for all time points of each triplicate, respectively.

3.4 Results and Discussion

3.4.1 The pVK-f-*lux* Plasmid Allows for Rapid Mixing and Matching Promoters and 5'UTRs

To study *cis*-regulatory RNA elements, we wanted to devise a luciferase reporter with a potent promoter to provide a strong signal. Cloning of the RBS sequence 5'-AGGAGC-3' was attempted downstream of P_{S7} . Restriction enzyme digestion of clones resulting from this vector assembly suggested that 7 of the 9 obtained clones contained misassembled constructs (Annexe I: Figure S1A and C). After multiple cloning attempts, sequencing of the only two clones with inserts of the correct size revealed mutations which would explain the lack of luminescence in these clones (Annexe I: Figure S1B). In that context, the most likely reason for these failed cloning experiments appeared to be a selective pressure against strong expression of this reporter; we thus used a fluoride riboswitch as a way to dampen expression independently of the promoter cloned upstream to design a new reporter tool with unique features.

For our design we chose to include three different single cutter restriction enzyme sites as GA overlap flanking sequences for the promoter region, both for the 5' end, designated as the Multiple Cloning Site 1 (MCS1), and for the 3' end, designated as the Multiple Cloning Site 2 (MCS2) (Figure 2.1B and C). Two versions of the plasmid were constructed, containing either the *P. syringae* fluoride riboswitch (*P. syr* F RS), termed pVK-f-*lux*, or the *B. thailandensis* E264 fluoride

riboswitch (*B. thai* F RS), termed pVK-f2-*lux* (Figure 3.1C). For the 3' end of the 5' UTR part, AarI, a type IIS restriction site was incorporated. If the inner restriction sites are used to cut the backbone to insert a promoter sequence, then the remainder of the MCS1 and MCS2 sequences on the linearized backbone are sufficient as GA overhangs (GA overlap 1 and GA overlap 2 in Figure 3.1B) and can be added to the insert of interest. Similarly, by digesting the backbone for a 5' UTR part insert using the innermost (in relation to the insert position) MCS2 site, and the AarI site, the backbone remainder of the MCS2 site may be used for the 5' GA overhang (GA overlap 3 in Figure 3.1B) and the beginning 15-20 nucleotides of *luxC* may be used as the 3' end GA overhang (GA overlap 4 in Figure 3.1B). Short inserts (18 and 21 bp) could not be cloned, because a small insert size is already known to be detrimental for Gibson assembly (Roth *et al.*, 2014).

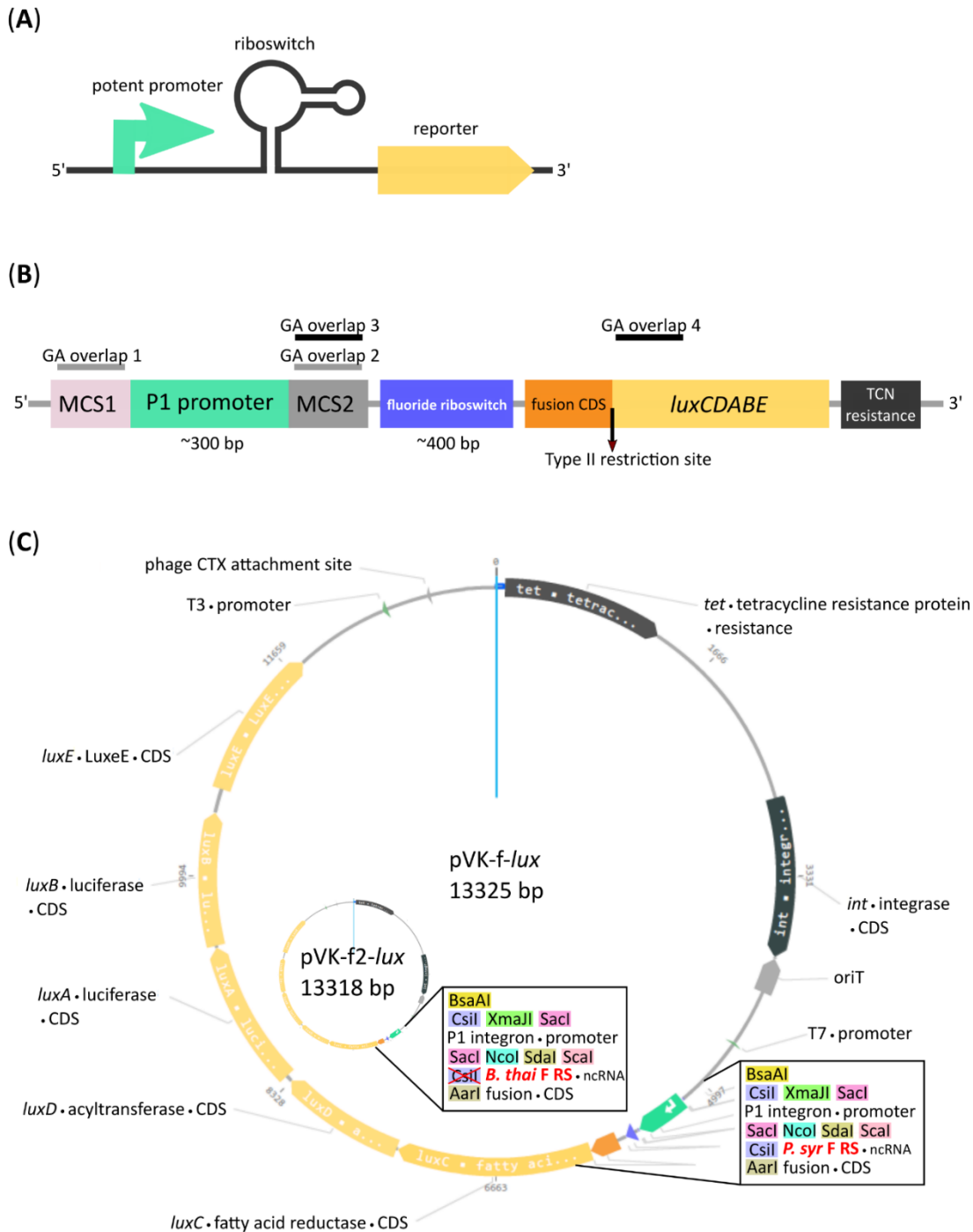


Figure 3.1 Riboswitch-mediated reporter expression under the control of a potent promoter.

(A) Schematic diagram of a potent promoter dampening strategy. (B) Key features of pVK-f-lux, a parts-swappable mini-CTX-lux derived backbone for convenient cloning of promoters and RNA regulatory parts. MCS1 contains CsiI-XmaJI-SacI restriction sites and MCS2 contains NcoI-SdaI-Scal restriction sites. A type IIS restriction site is located at the beginning of luxC allowing digestion into the second codon for scar-free translational fusion. GA overlaps for double digestions with SacI and NcoI or Scal and AarI have been designed for allowing interchangeability of parts as described (Annexe I: Quick User Manual for pVK-f-lux). These double digestions are buffer compatible and yield fragments visible on an agarose gel (approximately 300 and 400 bp

in length, see Figure S2 for more details). (C) Plasmid map of pVK-f-*lux* and pVK-f2-*lux*. Note that, while not annotated in the original mini-CTX-*lux*, Bprom predicts a rpoD17 promoter between *luxD* and *luxE*.

The AarI RE site is positioned to cut the backbone directly after the 2nd nucleotide of the 2nd codon of *luxC*, which allows achieving a scar-free translational fusion with a choice of the desired start codon. This could be useful given that some known examples of non-AUG start codons are important for translational regulation (Hecht *et al.*, 2017). If scar-free translational fusion is not a priority it is recommended to re-incorporate the AarI site and all depleted restriction sites into the constructs to allow for the flexibility of further cloning by using new plasmid constructs as backbones for new experiments rather than needing to start over from the original backbone, as required for many plasmid assembly methods. If at any point the innermost restriction enzyme recognition site of a Multiple Cloning Site (MCS), in relation to the insert, is not ideal for a particular cloning strategy, when using a newly assembled backbone, there are still two other restriction sites which may be used at each MCS, thus avoiding, in most cases, the necessity to domesticate any inserts (Annexe I: Figure S2). Additional information for using this plasmid is available (Annexe I: Quick User Manual for pVK-f-*lux*).

3.4.2 The Fluoride Riboswitch Dampens Reporter Gene Expression and Facilitates Potent Promoter Cloning

Having chosen the fluoride riboswitch as the candidate regulatory RNA for independent promoter dampening triggered by a non-cellular metabolite, we first needed to evaluate its usefulness to repress elevated expression levels with the option of re-activating expression. For this, we measured luminescence for a fluoride riboswitch reporter construct (pVK-f2-*lux*) in its original host organism (*B. thailandensis* E264) (Figure 3.2). Depending on the conditions, fluoride supplementation caused up to a ~20X induction of luminescence (RLU/OD₆₀₀) (Figure 3.2A). The addition of fluoride, or general ion content of media, had no apparent impact on osmotic pressure, since equivalent amounts of chloride made no difference (data not shown). In *B. thailandensis* E264, for OFF conformations (absence of fluoride), the repression capacity at the maximum peak expression of a 54h time-course luciferase expression curve of the *B. thai* F RS was 28X greater than that of the *P. syr* F RS, and 117X greater than that of the *B. thailandensis* E264 *metX* SAM riboswitch (*B. thai metX* SAM RS; Figure 3.2C).

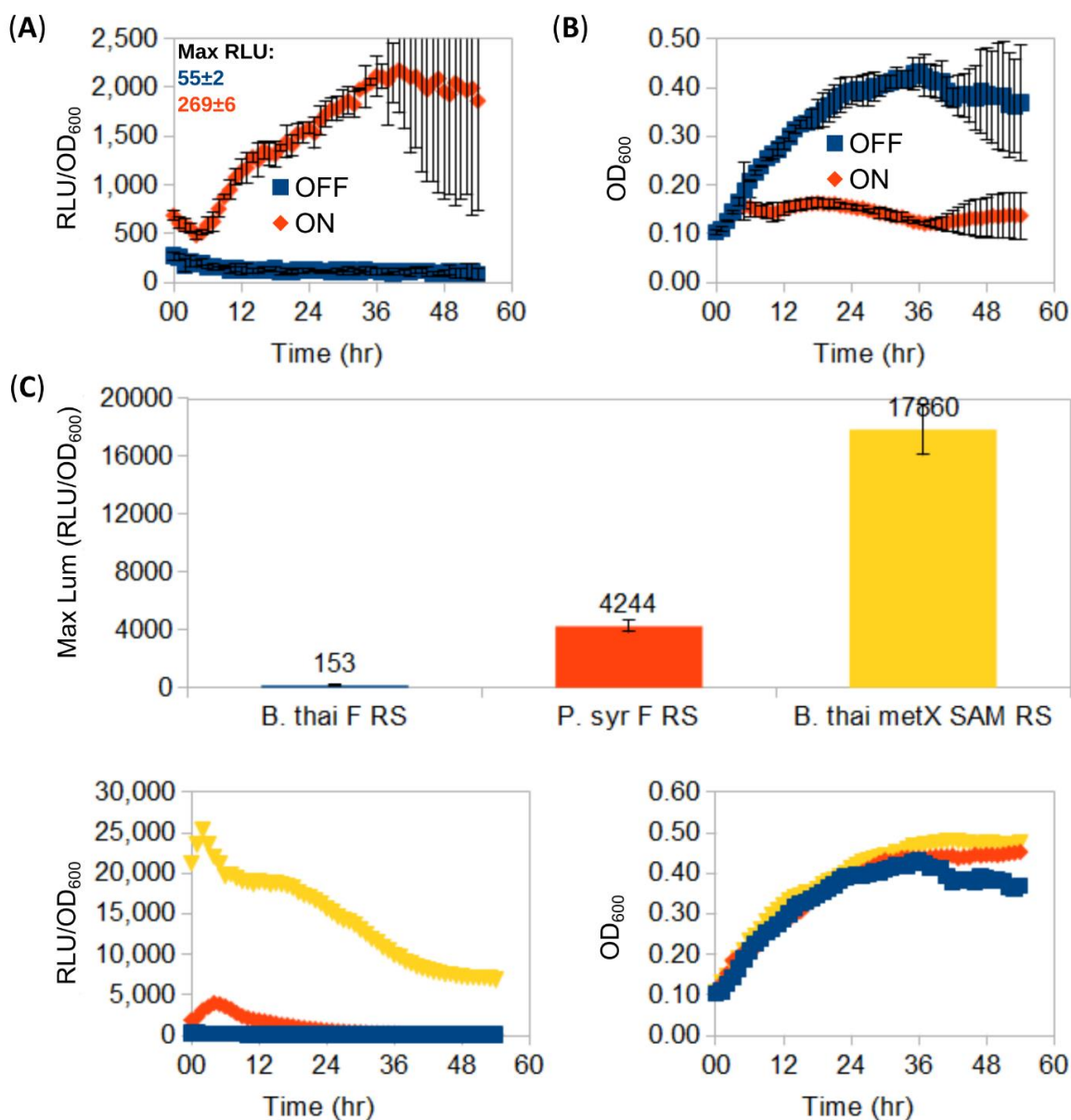


Figure 3.2 Time-course luminescence induction curves of *B. thailandensis* E264 clones containing the *B. thailandensis* fluoride riboswitch and repression effect.

A 54h time-course *lux* expression assay comparing *lux*/OD₆₀₀ (A) and growth curves (B) of *B. thailandensis* E264 containing the P1-*B. thai* F RS-*lux* (pVK-f2-*lux*) construct (See Table S3: *B. thailandensis* E264/ P1 + *B. thai* F) for either an ON-induced state of the fluoride riboswitch in the presence of 31 mM F⁻, or an OFF-induced state fluoride riboswitch in the absence of F⁻. Cultures were grown and measured on the same 96-well microplate assay run and maximum peak levels of luciferase expression are indicated above curves in relative light units (RLU) for un-normalised-to-OD signal strength comparison. The data points represent the means and standard deviations of triplicate values. (C) *In vivo lux* expression levels during repression by three OFF-induced riboswitch-containing constructs in *B. thailandensis* E264: *B. thailandensis* fluoride riboswitch construct (pVK-f2-*lux*—*B. thai* F RS), the *P. syringae* fluoride riboswitch (pVK-f-*lux*—*P. Syr* F RS), and the *B. thailandensis metX* SAM riboswitch (*B. thai metX* SAM RS). OFF-induction was achieved with 0 mM F⁻ for fluoride riboswitch constructs and with 0.05 mM methionine for the *metX* SAM riboswitch construct. RLU per OD₆₀₀ and growth curves are shown as well (bottom left and right respectively). In absence of plasmid, RLU

values were similar to RLU values of blank medium, while the OD₆₀₀ values were similar to (B). The data points represent the means of triplicate values.

Cloning attempts of P_{S7} were unsuccessful in spite of riboswitch addition, but cloning attempts of the P1 integron promoter were only successful in presence of a riboswitch—either the fluoride riboswitch sequence from *B. thailandensis* E264, the fluoride riboswitch sequence from *P. syringae*, three different S-adenosylmethionine (SAM) riboswitches from *B. thailandensis* E264 (recently discovered by our group, unpublished data) or the *yypP-ykoY* riboswitch from *P. aeruginosa* PA14. This suggests that a 5'UTR dampening tool is imperative to repress promoter potency and reduce its toxicity. Cloning the 5'-AGGAGC-3' RBS by itself downstream of the P1 promoter was unsuccessful using many GA design strategies (Annexe I: Table S1). However, cloning this RBS was successful when integrated within one of the above-mentioned riboswitches or when in tandem with a weak promoter such as that of the *metK* promoter from *B. thailandensis* E264.

Not all riboswitches enabled successful cloning of the strong P1 promoter, and none enabled cloning of very strong P_{S7} promoter, suggesting that the tested riboswitches did not sufficiently repress *lux* expression in their OFF conformations. As a reference, one study reported that mRNA coding for the S7 protein was among the top 3% in terms of total RNA quantity, highlighting how strong this promoter is, while *metK* mRNA was in the top 10% (Gorochofski *et al.*, 2019). A full list of successful and unsuccessful cloning experiments enumerated in Table 3.1 highlights the relation between expression levels and successful cloning of the promoter.

Tableau 3.1 Cloning success of the promoters upstream of the lux operon depends on the choice of 5'UTR.

Cloning attempt		Cloning outcome
Promoter	5' UTR	
P _{S7}	5'-AGGAGC-3' RBS	failed
P _{S7}	<i>B. thailandensis</i> fluoride riboswitch	failed
P _{S7}	<i>P. syringae</i> fluoride riboswitch	failed
P1 integron	5'-AGGAGC-3' RBS	failed
P1 integron	5'-AGGAGU-3' RBS	failed
P1 integron	<i>E. coli thiM</i> TPP riboswitch	failed
P1 integron	<i>B. thailandensis thiC</i> riboswitch	failed
P1 integron	<i>B. thailandensis</i> mini-ykkC riboswitch	failed
P1 integron	<i>B. thailandensis metK</i> SAM riboswitch ^{1,2}	failed
P1 integron	<i>B. thailandensis</i> fluoride riboswitch	successful
P1 integron	<i>P. syringae</i> fluoride riboswitch	successful
<i>B. thailandensis metK</i> promoter	5'-AGGAGC-3' RBS	successful
P1 integron	<i>B. thailandensis metX</i> SAM riboswitch ²	successful
P1 integron	<i>B. thailandensis metZ</i> SAM riboswitch ²	successful
P1 integron	<i>P. syringae</i> fluoride riboswitch + <i>B. thailandensis</i> fluoride riboswitch	successful
P1 integron	<i>P. aeruginosa</i> PA14 yybP-ykoY riboswitch	successful

¹ Three different construct designs of varying lengths were attempted (not shown).

²The *metK*, *metX* and *metZ* riboswitches from *B. thailandensis* are newly discovered riboswitches not yet published.

The fluoride riboswitch may be a useful tool for screening for a wide range of promoters and not only those which are potent. For example, a library of unknown sequences containing possible promoters may be cloned into pVK-f-*lux* with an up-regulating fluoride riboswitch in the target host such that the same library of clones can be screened for reporter gene activity in the presence or absence of fluoride. In this way, a screen in the absence of fluoride (with maximum repression), would yield those clones containing the most potent promoters. In parallel, another screen with added fluoride to re-activate riboswitch-mediated repression would allow to detect weaker promoters. Ideally, fluoride threshold tolerance of the target species during transformation and reporter assays as well as the timing of the expression pattern should be determined prior to screening. In this study it has been observed that *B. thailandensis* E264 is more tolerant to fluoride than *E. coli* SM10λpir, but less tolerant to high luciferase expression than *E. coli* SM10λpir (results not shown).

To test the possibility of using the fluoride riboswitch as a screening tool to mediate promoter potency during transformation of strong promoters (such as P_{S7} or P1 integron promoters), the viability of *E. coli* DH5α and *E. coli* SM10λpir transformant cells was assessed in the presence of 10 mM and 15 mM fluoride on selection plates. *E. coli* DH5α showed reduced viability at both concentrations as illustrated by the reduced number of visible colonies (Annexe I: Figure S3). Colonies which grew on fluoride supplemented plates were also visibly smaller for all tested constructs. Sequenced plasmid extractions of overnight inoculations of the P_{S7} promoter + *B.thai* F RS-*lux* in liquid media with corresponding fluoride concentrations revealed non-functional mutants for all cases even those for which *lux* expression should have been repressed in fluoride supplemented media. We suspect that the cloning failure of P_{S7} in this experiment was due to insufficient repression of *lux* expression rather than fluoride concentration levels in the transformation media as sequencing results were similar to previous GA attempts in absence of fluoride and using different GA designs for the same construct (Annexe I) and because cloning the unmodified backbone mini-CTX-*lux* in presence of 10 mM fluoride was also successful (Figure S3 D). Overall 33% of sequenced clones of the P_{S7} promoter + *B.thai* F RS-*lux* attempts from Figure S3 contained a 56 nt addition and a point mutation of C285T, in reference to the P_{S7} sequence, 37% had a gap, and 30% had an unrelated sequence included as the insert (Annexe I: Table S4). The effect of fluoride on growth in liquid expression media was also tested and *B. thailandensis* E264 clones with chromosomally integrated P1+riboswitch-*lux* constructs for the *B. thailandensis* or the *P. syringae* fluoride riboswitches. All were shown to grow equally well in 31.25 mM and 62.5 mM fluoride supplemented liquid media, demonstrating their high tolerance to fluoride presence (Figure S4 (C)), in contrast with *E. coli*.

3.4.3 Portability of riboswitches from one species to another can be unexpected

Luciferase expression was tested in *E. coli* DH5α, *E. coli* SM10λpir and *B. thailandensis* E264 each transformed with a construct containing the fluoride riboswitch originating either from *B. thailandensis* or from *P. syringae*. Unexpectedly, riboswitch modulation was inversed from one species to another for the fluoride riboswitch construct originating from *B. thailandensis* (Figure 3.3). While *B. thailandensis* E264 demonstrated an up-regulation, with a 5.1-fold change with the addition of fluoride, *E. coli* DH5α and *E. coli* SM10λpir yielded a down-regulation with a fold change of 9.2 and 1.7, respectively. For the *P. syringae* riboswitch construct, up-regulation was recorded for *E. coli* DH5α and *B. thailandensis* E264 as expected, with a fold change of 4.5 and 11 respectively (Figure 3.3). Other riboswitches were tested as well (Table 3.1), but either

dampened expression enough to allow cloning or did not provide as good of a modulation (less than two fold)—the latter was the case for the *B. thailandensis* SAM riboswitch upstream of *metK*. Regulation sensitivity thresholds were also tested for the P1 + *B. thai* F RS (pVK-f2-*lux*, Annexe I: Table S2) constructs transformed into *B. thailandensis* E264 and into *E. coli* SM10 λ pir and visible regulation effect was achieved at 3.9 mM for *B. thailandensis* E264/ P1 + *B. thai* F, and 1.95 mM for *E. coli* SM10 λ pir/P1 + *B. thai* F RS (Annexe I: Figure S4) further confirming sensitivity variability of a same riboswitch across different species. Characterization of presented constructs across different strains is important for understanding the limitations of a fluoride riboswitch-mediated *lux* reporter system. To this end, plasmid replication levels and luciferase expression levels were determined in *E. coli* SM10 λ pir and *E. coli* DH5 α . *E. coli* SM10 λ pir expresses far more luciferase than *E. coli* DH5 α at similar fitness levels based on the plateau OD₆₀₀ value, however counterintuitively *E. coli* DH5 α produces 1.4X more of pVK-f2-*lux* than the former (Annexe I: Table S5).

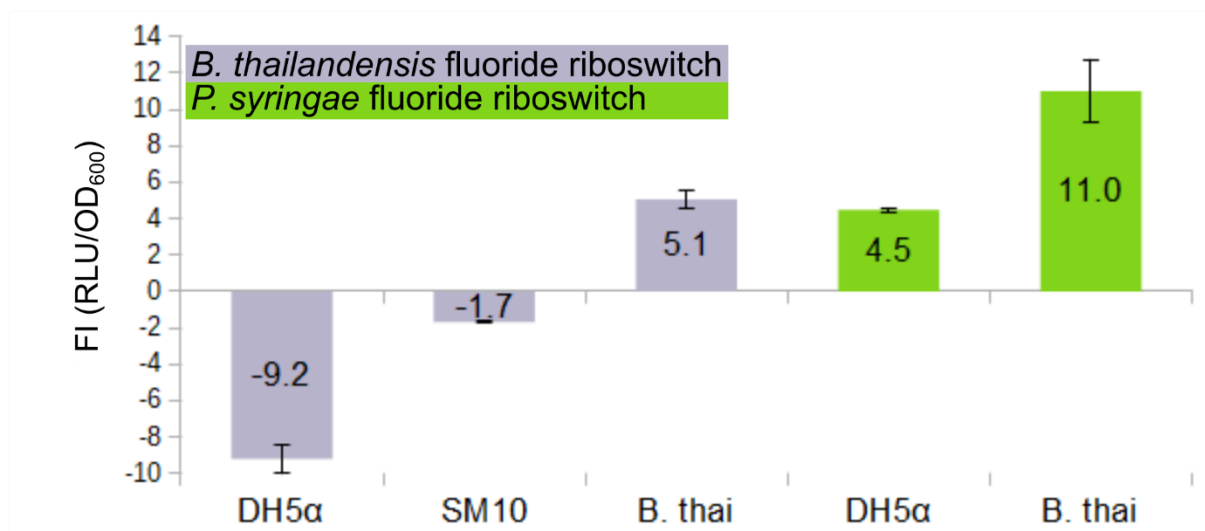


Figure 3.3 The *B. thailandensis* E264 fluoride riboswitch upregulates in original host and downregulates in *E. coli* DH5 α and *E. coli* SM10 λ pir

Fold induction (FI) of total luciferase activity in cultures was calculated over an incubation period of ~40 h (*E*

. coli DH5 α —DH5 α , *E. coli* SM10 λ pir—SM10, and *B. thailandensis*—B. thai transformed with the *B. thailandensis* E264 fluoride riboswitch construct—pVK-f2-*lux*) and ~55h (*E. coli* DH5 α and *B. thailandensis* E264 transformed with *P. syringae* fluoride riboswitch construct—pVK-f-*lux*). Both plasmids are replicative in *E. coli* strains and integrative in *B. thailandensis* E264. Both pre-culture media (LB) and expression media (0.5X M9-MM) were supplemented with appropriate antibiotics depending on the species (see Methods). Fluoride induction concentrations were 16 mM for *E. coli* DH5 α and *E. coli* SM10 λ pir transformants and 31 mM for *B. thailandensis* E264. Concentrations were chosen according to the maximum induction effect. The FI values for each sample represent the means and standard deviations of triplicates on the same microplate.

When running a time-course expression assay, the duration should be optimized to ensure that an expression peak is attained for the given strain and media conditions. Even if addition of 8 mM or 16 mM fluoride induced luciferase expression in all relevant assays, some discrepancies were observed between some clones, both with regards to exact expression quantitation and growth curves. Additionally, peaks were reached at different times in different fluoride concentrations highlighting the importance of a sufficiently long assay run (Annexe I: Figure S5). Oscillating expression (with ups and downs) may also be observed when using the fluoride riboswitch in reporter assays as fluoride concentrations inside bacteria will vary according to the activity of the fluoride export pumps (Annexe I: Figure S5).

Characterization is not only important for understanding how different conditions affect riboswitch dynamics but also how they may affect the *lux* cassette enzymes, as the system is composed of five different enzymes (*luxCDABE*). We did test the system's sensitivity to unrelated inducers/repressors. Effect of chloride (up to 15.6 mM) was tested on *E. coli* clones carrying pVK-f2-*lux*, however no regulatory effect was observed (results not shown). Methionine addition (up to 125 mM) to *B. thailandensis* E264 clones carrying pVK-f2-*lux* also did not have a regulatory effect.

3.5 Conclusion

In this study we designed a plasmid which allows for straightforward swapping of promoters and 5'UTR translationally fused sequences directly from PCR amplified inserts using Gibson assembly-based homologous cloning. We also included the fluoride riboswitch as a tool for modulating reporter gene expression under the control of strong constitutive promoters, such as the P1 integron promoter, in order to circumvent possible reporter overexpression toxicity in both shuttle and final host species, even if it still has limitations, as exemplified by the unsuccessful cloning of the particularly strong P_{S7} promoter. We also illustrate that riboswitches used as cloning tools need to be characterized across shuttle species as well as the target species to ensure optimal use. Indeed, we discovered that in its native species the *B. thailandensis* fluoride riboswitch upregulates expression when supplemented with fluoride, yet it surprisingly downregulates expression in *E. coli*. To our knowledge this is the first time such a phenomenon has been observed, even if there are several accounts of riboswitches that function normally in their native species, but do not modulate gene expression in a different species (usually more phylogenetically remote). While we did not further study this occurrence, we reason that this may be due to the difference in riboswitch expression platform folding kinetics due to difference in RNA polymerase activity across species or the wide gap in the GC% of their respective genomes (67%

vs 50% for *B. thailandensis* and *E. coli*, respectively). Nevertheless, this work may serve as an example of riboswitch use to improve current cloning tools. Other riboswitches whose ligands are independent of the host organism's metabolism, such as the theophylline synthetic riboswitch, may provide alternatives to apply the same approach, potentially circumventing some of the project-specific limitations that can be encountered the same way different resistance markers can be more or less appropriate for a given cloning experiment.

4 BIBLIOGRAPHIE DU CHAPITRE 3

- Baker JL, Sudarsan N, Weinberg Z, Roth A, Stockbridge RB & Breaker RR (2012) Widespread genetic switches and toxicity resistance proteins for fluoride. *Science* 335(6065):233-235.
- Becher A & Schweizer HP (2000) Integration-proficient *Pseudomonas aeruginosa* vectors for isolation of single-copy chromosomal lacZ and lux gene fusions. *Biotechniques* 29(5):948-950, 952.
- Bolognesi B & Lehner B (2018) Reaching the limit. *Elife* 7.
- DeShazer D & Woods DE (1996) Broad-host-range cloning and cassette vectors based on the R388 trimethoprim resistance gene. *Biotechniques* 20(5):762-764.
- Eguchi Y, Makanae K, Hasunuma T, Ishibashi Y, Kito K & Moriya H (2018) Estimating the protein burden limit of yeast cells by measuring the expression limits of glycolytic proteins. *Elife* 7.
- Gorochowski TE, Chelysheva I, Eriksen M, Nair P, Pedersen S & Ignatova Z (2019) Absolute quantification of translational regulation and burden using combined sequencing approaches. *Mol Syst Biol* 15(5):e8719.
- Guo Y, Hui CY, Liu L, Zheng HQ & Wu HM (2019) Improved Monitoring of Low-Level Transcription in *Escherichia coli* by a beta-Galactosidase alpha-Complementation System. *Front Microbiol* 10:1454.
- Hecht A, Glasgow J, Jaschke PR, Bawazer LA, Munson MS, Cochran JR, Endy D & Salit M (2017) Measurements of translation initiation from all 64 codons in *E. coli*. *Nucleic Acids Res* 45(7):3615-3626.
- Hollis RP, Lagido C, Pettitt J, Porter AJ, Killham K, Paton GI & Glover LA (2001) Toxicity of the bacterial luciferase substrate, n-decyl aldehyde, to *Saccharomyces cerevisiae* and *Caenorhabditis elegans*. *FEBS Lett* 506(2):140-142.
- Kintaka R, Makanae K & Moriya H (2016) Cellular growth defects triggered by an overload of protein localization processes. *Sci Rep* 6:31774.
- Korbie DJ & Mattick JS (2008) Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat Protoc* 3(9):1452-1456.
- Liu M, Blinn C, McLeod SM, Wiseman JW, Newman JV, Fisher SL & Walkup GK (2014) Secreted *Gaussia princeps* Luciferase as a Reporter of *Escherichia coli* Replication in a Mouse Tissue Cage Model of Infection. *PLOS ONE* 9(3):e90382.
- Malakar P, Singh VK, Karmakar R & Venkatesh KV (2014) Effect on beta-galactosidase synthesis and burden on growth of osmotic stress in *Escherichia coli*. *Springerplus* 3:748.
- Roth TL, Milenkovic L & Scott MP (2014) A Rapid and Simple Method for DNA Engineering Using Cycled Ligation Assembly. *PLoS ONE* 9(9).
- Speed MC, Burkhart BW, Picking JW & Santangelo TJ (2018) An Archaeal Fluoride-Responsive Riboswitch Provides an Inducible Expression System for Hyperthermophiles. *Appl Environ Microbiol* 84(7).
- Stockbridge RB, Lim HH, Otten R, Williams C, Shane T, Weinberg Z & Miller C (2012) Fluoride resistance and transport by riboswitch-controlled CLC antiporters. *Proc Natl Acad Sci U S A* 109(38):15289-15294.

- Weinberg Z, Wang JX, Bogue J, Yang J, Corbino K, Moy RH & Breaker RR (2010) Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol* 11(3):R31.
- Zolg JW & Hanggi UJ (1981) Characterization of a R plasmid-associated, trimethoprim-resistant dihydrofolate reductase and determination of the nucleotide sequence of the reductase gene. *Nucleic Acids Res* 9(3):697-710.

5 STRUCTURAL DIVERSITY OF NEW SAM-II/V RIBOSWITCH VARIANTS

Structural diversity of new SAM riboswitch variants extends the scope of the SAM-II/V clan

La diversité structurelle d'une nouvelle variante de riboswitch SAM élargit le clan des riboswitchs SAM-II/V

Auteurs :

Balasubramanian Sellamuthu^{1,2*}, Vesta Korniakova^{1*}, Xiaoling Yang¹, Mohammad Reza Naghdi¹, Radia Benmiloud¹, Michelle Mageau-Dubuc¹ and Jonathan Perreault¹

¹INRS - Centre Armand-Frappier Santé Biotechnologie, 531 boulevard des Prairies, Laval, QC, H7V 1B7, Canada.

²Centre de Recherche du CHUM, Université de Montréal, 900 rue st-Denis, Montréal, QC, H2X 0A9, Canada

Titre de la revue ou de l'ouvrage :

Il est prévu de soumettre cet article prochainement à Nature Microbiology

Contribution des auteurs :

Les deux premiers auteurs sont considérés comme ayant une contribution équivalente.

Balasubramanian Sellamuthu a conçu les séquences d'aptamères pour les expériences *in vitro*, a exécuté et analysé les résultats des expériences *in vitro* et a contribué à la rédaction de l'article, incluant la prédiction du mécanisme du riboswitch *metK*.

Vesta Korniakova a exécuté et analysé les résultats des expériences *in vitro* pour des aptamères SAM provenant de différents organismes. Elle a conçu, exécuté et analysé toutes les expériences *in vivo* et a contribué à la rédaction de l'article.

Xiaoling Yang a fait certains des essais préliminaires *in vitro* qui ont confirmé l'existence d'un nouveau riboswitch SAM.

Mohammad Reza Naghdi et Radia Benmiloud ont fait des analyses bio-informatiques, notamment pour des analyses phylogénétiques et pour évaluer la covariation des sous-familles de riboswitchs (incluant leurs plateformes d'expression).

Michelle Mageau-Dubuc a exécuté et analysé les résultats des expériences *in vitro* pour des aptamères SAM provenant de différents organismes.

Jonathan Perreault a fait la découverte initiale des riboswitch « SAM-beta » par bio-informatique, a guidé les différents aspects du projet et a rédigé l'article.

Lien entre l'article du chapitre 3 et l'article du chapitre 5 :

Dans le chapitre 3, on a réussi à surmonter l'occurrence de la toxicité due à la surexpression par un promoteur fort du gène rapporteur dans le cas d'utilisation de la cassette *luxCDABE* grâce à l'addition adjacente d'un riboswitch fluor qui permet de réduire l'expression malgré la force du promoteur. Ce rajout du niveau de régulation entre le promoteur et le gène rapporteur a permis le clonage du promoteur fort de l'intégron P1, en évitant la toxicité létale.

On a aussi découvert une régulation inverse lors de l'utilisation d'un riboswitch fluor provenant de *B. thailandensis* E264 dans *E. coli*, indiquant que les riboswitchs doivent être utilisés avec prudence dans les hôtes autres que les hôtes natifs du riboswitch.

L'étude dans le chapitre 3 a introduit un nouvel outil pour la découverte des promoteurs précédemment non-caractérisables et aussi un vecteur de clonage pour la fusion traductionnelle avec la luciférase, pour un plasmide intégratif dans plusieurs espèces du genre *Burkholderia* et *Pseudomonas*.

Cette dernière caractéristique du plasmide, soit la possibilité de concevoir des fusions traductionnelles des riboswitchs avec la cassette *luxCDABE*, permettra son utilisation pour les essais *in vivo* effectués dans ce chapitre, le chapitre 5, suivant en conjonction avec les trois riboswitchs SAM chez *B. thailandensis* en amont de *metK*, *metX*, ou *metZ*. De plus, la caractérisation du promoteur P1 chez *B. thailandensis* E264 de cette étude permettra de tester des construits alternatifs pour la comparaison des riboswitchs SAM découplés de leurs promoteurs natifs

5.1 Abstract

The S-adenosylmethionine-V (SAM-V) riboswitch was discovered with a bioinformatics pipeline aiming to discover completely new structured RNA. Yet, the SAM-V riboswitch was found to have the same overall architecture as the previously described SAM-II. Here we report several additional families that are members of this SAM-II/V “clan” of SAM-riboswitches which we initially discovered in Betaproteobacteria, but also found in numerous species with a wide phylogenetic distribution. Within the proposed new “universal consensus” for SAM-II/V, we reveal a diversity of different secondary structures with accessory stems and competing alternative stems, of which several were tested and confirmed to bind SAM. Our results suggest that hundreds to thousands of riboswitch instances remain unannotated, largely due to a structural diversity difficult to capture with simple homology searches. Likely, this assortment of secondary structures contributes to fine-tuning the affinity and switching mechanisms of riboswitches for appropriate regulation of the downstream genes.

5.2 Introduction

Riboswitches are structured RNA elements within mRNAs that control gene expression through selective binding of their ligand. They are generally located in 5'-UTRs (Untranslated regions) of mRNAs (Edwards *et al.*, 2010). Their aptamer domain is responsible for binding their cognate metabolite ligand, which exerts a conformational rearrangement of an expression platform for gene regulation (Mandal & Breaker, 2004). The region of the riboswitch common to both the aptamer and expression platform is called the “switching sequence.” Riboswitches control gene expression by various mechanisms like transcription (by an intrinsic terminator stem) and translation control (by ribosome binding site occlusion) (Breaker, 2012). To date, at least 40 known families of riboswitches have been validated for at least 30 different ligands (Breaker, 2012; McCown *et al.*, 2017). Riboswitch families have been categorized based on their ligand specificity and their secondary structure (Montange & Batey, 2008), the Rfam riboswitch families are essentially based on the ability of each covariance model to find all instances of the corresponding family.

SAM (S-adenosyl methionine) binding riboswitches are unique for the diversity of structures produced by convergent evolution. In the RNA family database, there are at least seven families of SAM riboswitches reported to regulate SAM and/or methionine and sulfur metabolism in bacteria (including survival and virulence genes in pathogens) (Corbino *et al.*, 2005; Epshtein *et al.*, 2003; Fuchs *et al.*, 2006; McDaniel *et al.*, 2003; Naghdi *et al.*, 2020; Poiata *et al.*, 2009;

Trausch *et al.*, 2014; Weinberg *et al.*, 2008; Weinberg *et al.*, 2010; Winkler & Breaker, 2003). However, a few of these families, which can be grouped according to the similarity of their binding pocket (McCown *et al.*, 2017), appear less likely to have distinct phylogenetic origin (Trausch *et al.*, 2014). One such “clan” includes the SAM-I, IV and I/IV riboswitches (found, respectively, in gram-positive organisms, Actinomycetales and mostly in proteobacteria) (Trausch *et al.*, 2014; Weinberg *et al.*, 2008; Weinberg *et al.*, 2010). Another clan is that of SAM-II and SAM-V riboswitches, both found mostly in alphaproteobacteria and Bacteroides (Corbino *et al.*, 2005; Poiata *et al.*, 2009). More recently, SAM-III (found in Lactobacillales) (Fuchs *et al.*, 2006) and SAM-VI (found in Bifidobacterium) (Mirihana Arachchilage *et al.*, 2018; Sun *et al.*, 2019) were also found to be likely members of the same clan. However, SAM/SAH riboswitches (reported in α -proteobacteria) (Weinberg *et al.*, 2010) is different in terms of structure and ligand binding pocket and cannot be grouped into a clan. In spite of similarities made evident by atomic resolution structures (Gilbert *et al.*, 2008; Huang & Lilley, 2018; Stoddard *et al.*, 2010; Zhang *et al.*, 2019), powerful structure homology search tools such as Infernal are not able to uncover the corresponding homologs from the same clan if they are not part of the same family. Using a *de novo* ncRNA search pipeline (Naghdi *et al.*, 2017), we have found a “new” SAM riboswitch which corresponds to the SAM-II/V clan. Using this search strategy, adapted for the SAM-II/V clan, we have unravelled numerous new instances of riboswitches and an unforeseen structural versatility of this clan. Although previously two examples of these riboswitches, one upstream of *metX* and another upstream of *metZ* in beta-proteobacteria, have been proposed, they have not been confirmed experimentally and their structures were incorrectly predicted (Leyn *et al.*, 2014). Results of characterization through bioinformatics, *in vitro* structural probing and reporter gene constructs indicate that evolution exploits this structural plasticity to fine-tune each riboswitch for expression requirements of the downstream gene by adapting each aptamer to its expression platform.

5.3 Results and Discussion

5.3.1 Discovery of new, structurally diverse, riboswitch variants

An initial search aiming to discover new structural motifs using a pipeline which compares sequences upstream of related genes (Naghdi *et al.*, 2017) led us to find an RNA motif upstream of *metK* in Betaproteobacteria which seemed to have nothing in common with any of the thousands of secondary structures catalogued in Rfam (Annexe II: Figure S1). Given that several riboswitches were known to regulate *metK* according to SAM concentration, we compared our

sequences with these riboswitches. Despite the incapacity of Infernal, the RNA structure annotation tool of reference, to find homology between our aligned sequences and covariance models from Rfam, careful evaluation of the sequences corresponding to the new model indicated that they could fold in a secondary structure forming a SAM binding pocket identical to that of SAM-II (Doshi *et al.*, 2012; Gilbert *et al.*, 2008) or SAM-V (Huang & Lilley, 2018; Poiata *et al.*, 2009). In the case of the SAM II/V clan, the main caveat for RNA annotation may be that close to half of the riboswitch structure is found within a pseudoknot (Figure 5.1ABC), a characteristic which is not taken into consideration by Infernal thus limiting annotations.

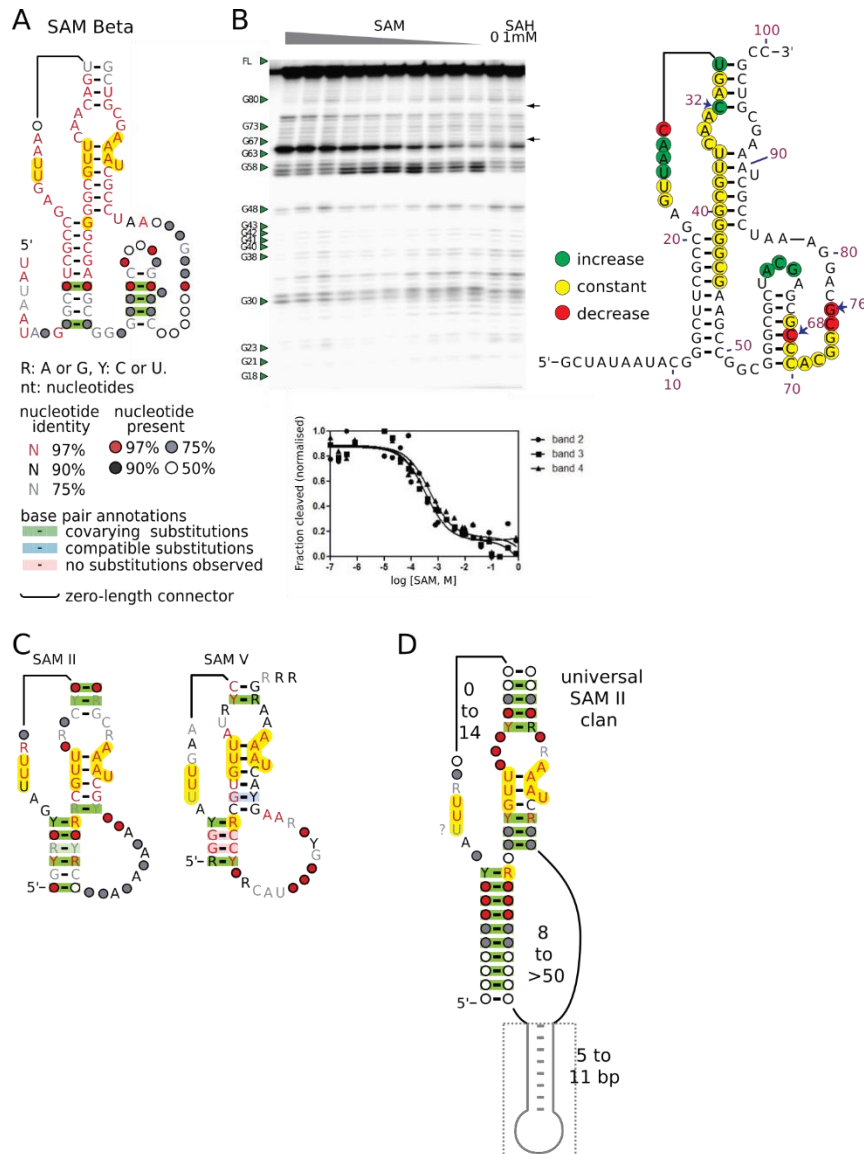


Figure 5.1 A novel riboswitch family within the SAM-II clan.

(A) Consensus secondary structure of the novel riboswitch family found by comparative genomics drawn according to the SAM-II clan architecture (initial structure predicted pictured in Annexe II: Fig. S1). Color legend

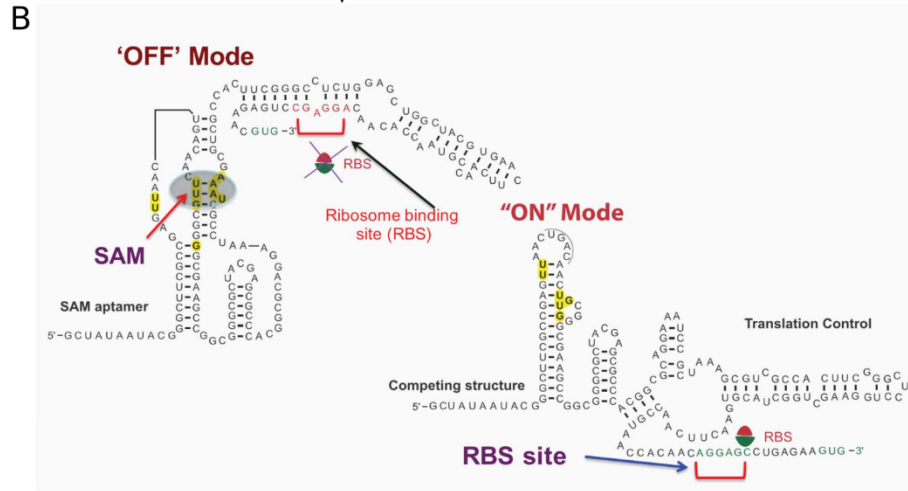
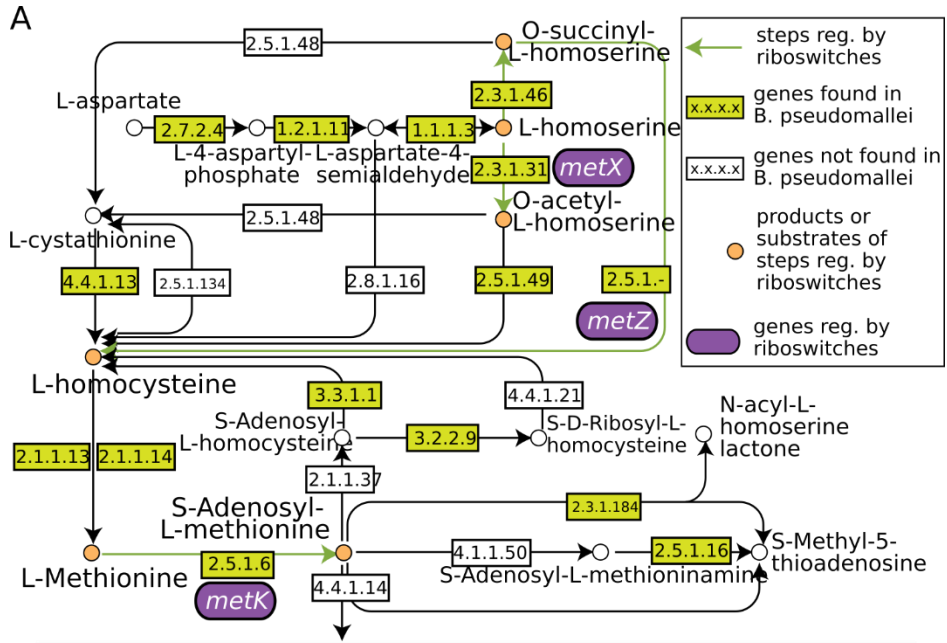
is shown within figure. (B) In-line probing of the RNA riboswitch instance from *B. thailandensis* pictured on the right. Green, yellow and red correspond to greater, same or lesser band intensity, respectively. Curves used to calculate a K_d of 0.3 μM is shown below. (C) Consensus secondary structures of the SAM-II and V riboswitch families, both belonging to the same clan. (D) “Universal” consensus structure of the entire SAM-II-V clan determined from additional instances found with RNArobo (Annexe II: Fig. S2).

Realizing this could mean that beyond SAM-II, SAM-V and the new SAM variant riboswitch upstream of Betaproteobacteria *metK*, several other so-called “new” riboswitches might still be unannotated, we designed a motif descriptor based on the combined consensus of these three RNA families (Annexe II: Supplementary Material) to search for other instances using the pseudoknot-compatible search tool RNArobo (Webb *et al.*, 2009). Bioinformatics results exhibited many more riboswitch candidates across different bacterial phyla. These revealed an unexpected diversity of structure including length variability of loop 3 ranging from 8 to 60 bases, and the possibility of additional stems within that loop (example structures in Annexe II: Figure S2). This led us to redefine the consensus of the SAM-II/V clan (Figure 5.1D). Interestingly, the *metK* RNA motif that we uncovered in Betaproteobacteria, which we initially thought had nothing in common with known SAM riboswitches, corresponds to what appears to be a stable stem-loop that would presumably compete with the aptamer structure (Annexe II: Figure S3), and possibly contribute to this riboswitch’s mechanism. This riboswitch also harbors a stem within loop 3, which we termed “accessory stem”, that shows covariation (Figure 5.1A), suggesting a functional role. Moreover, several other examples of variant riboswitches also present such accessory stems (Annexe II: Figure S2).

5.3.2 Three SAM riboswitches regulating SAM and methionine synthesis in *Burkholderia thailandensis*

Beyond the finding of the initial new SAM riboswitch upstream of *metK* in *B. thailandensis*, two other SAM riboswitches of the same superclass were found upstream of *metX* and *metZ* in this species and many other *Burkholderiales*. These genes are situated upstream of methionine synthesis and, while *metX* is essential, *metZ* is one out of two alternative pathways to produce homoserine, direct precursor of methionine (Fig. 5.2A). Thus, even if the bacteria use SAM to determine whether all three genes should be expressed or not, the requirements differ for each. Accordingly, the sensing is also expected to differ, either in terms of amplitude of response or in the threshold of sensing and the modulation range. This is reflected by the completely different expression platforms (Fig. 5.2BCD) as well as by the affinity of the corresponding aptamers to SAM (Fig. 5.1B and Annexe II: Fig. S2) and by the different expression profiles of a reporter gene in M9 medium supplemented, or not, with methionine (Fig. 5.2E). Since methionine can be

considered the limiting precursor of SAM synthesis, adding it to M9 medium is expected to increase levels of methionine. Surprisingly, none of the concentrations of methionine assayed had any impact on reporter gene activity when under the control of the *metZ* riboswitch (Fig. 5.2F). O-succinyl-L-homoserine is the substrate of *metZ* to incorporate sulfur in the methionine pathway directly from sulfide, as an alternative pathway to *metY* that uses O-acetyl-L-homoserine instead (Fogliano *et al.*, 1995), both to produce L-homocysteine (Fig. 5.2). Thus, as opposed to *metK* and *metX*, which are both essential steps in the methionine and SAM biosynthesis, the requirement to de-repress *metZ* might be more stringent and extremely low concentrations of SAM might be required to activate expression of this gene, concentrations which we apparently could not attain in our assays. Even when using cycloleucine (an inhibitor of *metK*) in absence of methionine in the medium for the reporter assay, only a relatively small difference could be observed, and not in all conditions tested (Fig. 5.2F).



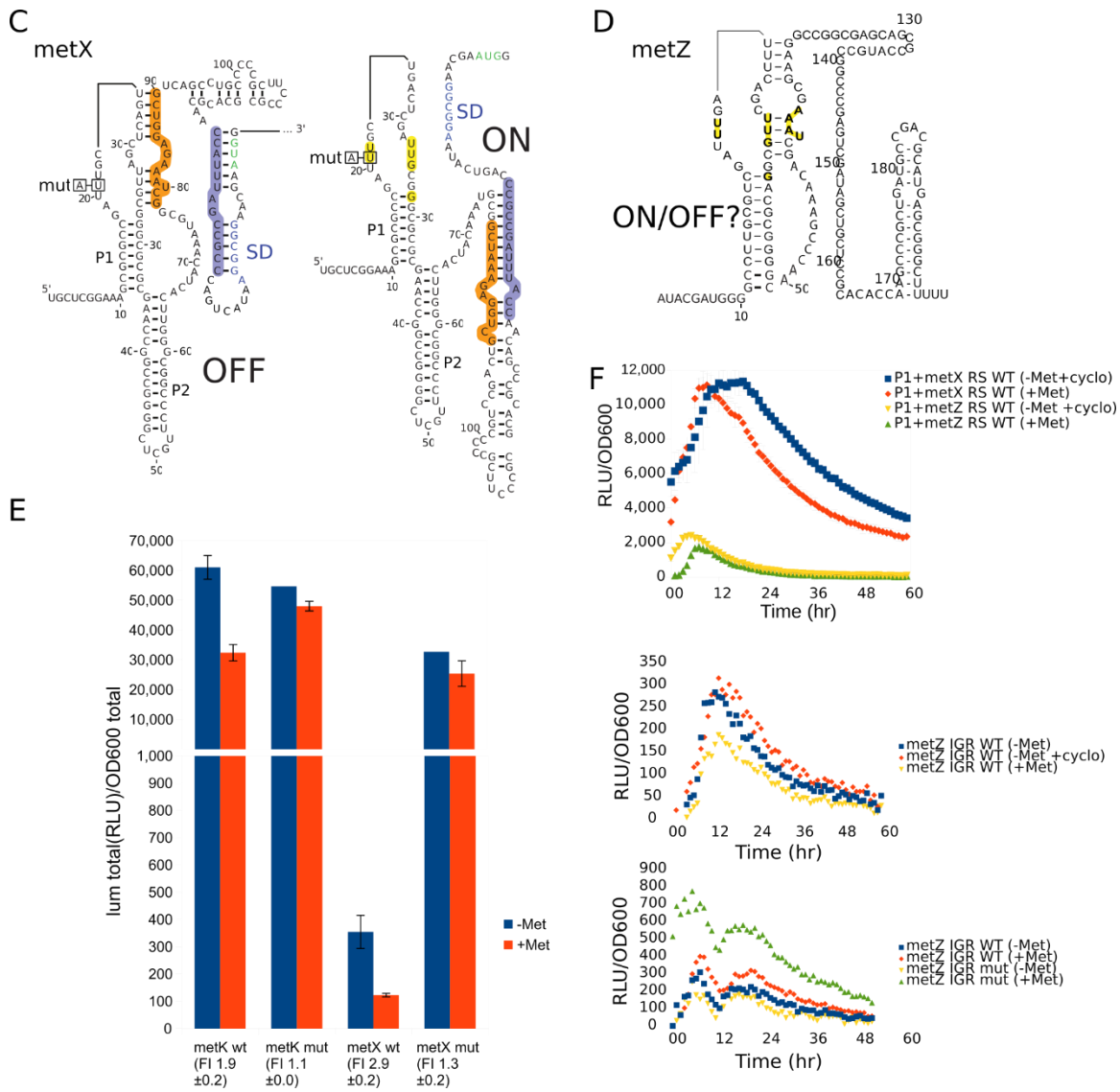


Figure 5.2 Differential regulation of the methionine/SAM biosynthesis pathways by different SAM riboswitches in *Burkholderia*.

(A) Biosynthesis pathways of SAM and methionine. Steps regulated by the novel riboswitches uncovered in this study are highlighted as per in-figure legend. (B, C and D) SAM riboswitches with predicted expression platforms for *metK*, *metX* and *metZ*, respectively. (E) Modulation of gene expression of *lux* reporter by the *metK* and *metX* riboswitches according to supplementation of 3.75 mM methionine, and absence of modulation for mutant versions of the riboswitch. (F) Luminescence curves for a 60 (top) and 50 (middle and bottom) hour time course for the *metX* and *metZ* riboswitches under the control of the P1 integron promoter (top), or the native IGR promoters of *metX* or *metZ* (middle and bottom). MetX RS and metZ RS correspond to riboswitches that putatively regulate *metX* and *metZ* respectively; either with WT sequence or mutant (mut) shown in (C). The cultures were supplemented with methionine (+Met) or not (-Met) or supplemented with cycloleucine (+cyclo), a SAM synthase inhibitor to further reduce intracellular SAM concentrations.

Modulation of gene expression by the riboswitches upstream of *metX* and *metK* also showed marked differences (Annexe II: Fig. S4). Both aptamer versions of *metX* and *metK* have an

accessory stem within their loop 2 (Fig. 5.2BC), but it differs in its size and its position relative to P, as well as its surrounding sequence. Moreover, the *metK* version of the aptamer has a potential competing structure within the structure (Annexe II: Fig S3). Probing experiments with radiolabeled RNA suggest a k_d of 0.3 μM for the *metK* riboswitch and of $\sim 20 \mu\text{M}$ for the *metX* riboswitch (Fig. 5.1 and Annexe II: Fig. S2) and gene expression is also modulated, ~ 2 -fold for the *metK* SAM riboswitch and ~ 3 -fold for the *metX* SAM riboswitch, when 3.75 mM methionine is added to 0.5X MM-S medium (Fig. 5.2E). However, the *metK* riboswitch appears to modulate like a “threshold-based” switch at concentrations as low as 0.039 mM methionine, with apparently no increase in modulation at higher concentrations, while the *metX* riboswitch appears to act more like a “dimmer”, with repression increasing progressively from 0.1 mM (perhaps less) up to 3.75 mM (Annexe II: Fig. S5). Previously, an S-box–*lacZ* fusion of the *B. subtilis metK* SAM riboswitch in its original host did not exhibit a fold induction upon addition of methionine to starvation conditions, however for other S-box gene-*lacZ* transcriptional fusions such as of S-box *yitJ*, *ykrT*, *metE*, *ykrW*, *yusC* and *yoaD*, a 15 to 2300 fold induction was observed, where generally more tightly repressed genes had the biggest fold induction (Tomsic *et al.*, 2008). In this present study, induction levels for the *metK* and *metX* SAM riboswitches follow this correlation as well, while the *metZ* riboswitch, the most tightly repressed, did not exhibit apparent induction under the tested conditions. Also, mutant versions of each riboswitch have different levels of expressions relative to their WT functional counterparts. The *metK* inactive mutant riboswitch has an expression level (maximum) similar to the WT version in absence of added methionine in the medium, conversely the *metX* inactive mutant riboswitch has an expression level about 100 fold higher than the WT version in absence of methionine (and ~ 300 fold higher than in presence of methionine) (Fig. 5.2EF). The much higher expression observed for this mutant could suggest that, like hypothesized for the *metZ* riboswitch, our current experimental setting does not allow us to decrease the SAM concentration enough to attain maximal expression. Interestingly, probing assays of several new betaproteobacteria variant riboswitches harboring an accessory stem show an unusual modulation pattern. In these riboswitches, the SAM concentration range for which conformational changes occur spans 4 to 5 logs, instead of the usual 2 logs observed for most riboswitches (Annexe II and Annexe II: Fig. S2), further suggesting a “dimmer” type of switch as compared to typical riboswitches or the cooperative riboswitches deemed to be more “digital” and for which changes occur within only 1 log of concentrations (Dann *et al.*, 2007; Furukawa *et al.*, 2015; Kwon & Strobel, 2008).

It is possible that the three distinct SAM riboswitches in *Burkholderia* originated from a single riboswitch, for example by reshuffling in the genome, ending with a copy of the aptamer (and

maybe also a part of the expression platform) sequence upstream of other genes within the methionine metabolic pathway, to eventually re-adapt its expression platform to the new RNA environments. Alternatively, the riboswitches could have originated from different species, untraceable to a common ancestor, to incorporate into the methionine metabolic pathway of one bacterium by incorporation of full riboswitch-gene blocks by horizontal gene transfer. A previous study concluded that aptamers for the same ligand can evolve into variations having up to a 21-fold difference in their ligand binding affinity (Mulhbacher & Lafontaine, 2007), further highlighting evolutionary fine-tuning and differences in regulatory requirements for different genes.

5.4 Conclusion

Infernal is the current gold standard for RNA annotation, it is a powerful tool to find sequences that fit an RNA structure-annotated alignment, taking in consideration both structure and sequence conservation. This allows the annotation of instances distantly related from sequences in the alignment, such as for TPP riboswitches, for which sequence homology alone would be insufficient to find all instances; e.g. tools like BLAST find no homology between TPP riboswitches from *Acidobacteria* and from *Penicillium*, while both fit the Infernal covariance model of the TPP riboswitch in Rfam. Nevertheless, Infernal has some limitations, the main one being that it cannot take into account base pairs in pseudoknots. This has already been highlighted for structures similar to Hepatitis Delta Virus ribozyme but not for other families. In the case of riboswitches belonging to the same clan as SAM-II, it limits searches of families to homologs where the pseudoknot base pair identities are conserved, because covariation of the pseudoknot cannot be detected. This also reduces the weight of sequence conservation of the binding pocket, which will appear as having a widely diverse positioning relative to P1, because it is actually embedded in the undetected P2/P3 pseudoknot. Consequently, the “SAM-beta-*metK*” instances, were discovered with a pipeline destined to find novel structures, because they had never been annotated before. This prompted us to use a different approach to find more instances and led us to unravel three novel families, part of the SAM-II clan, for which the covariance models have very little overlap with the other known Rfam families.

This diversity of novel riboswitch families also revealed a very flexible riboswitch architecture, which apparently evolved to include multiple adaptations to fine-tune each of these riboswitches for the expression of the downstream genes. Indeed, the three new SAM-riboswitches we have studied in *Burkholderia thailandensis* have distinct aptamer structures, different expression

platforms and appear to modulate expression in diverse manners as well. These differences in modulation include more pronounced repression at the peak of expression, or during the entire time-course (spanning >2 days) as well as different levels of modulation from undetectable (in the tested conditions), to ~2 fold, but with a potential for ~100 fold according to binding pocket mutants. Finally, even if the idea of “ribo-dimmers” rather than riboswitches was evocated previously (Baird *et al.*, 2010), to our knowledge, as opposed to the “digital” cooperative riboswitches which have been experimentally verified, there has been no account of “dimmer-like” riboswitches. More work will be required regarding the structure and conformation dynamics of these new families to better understand this phenomenon. Overall, these novel riboswitches are a hint towards evolutionary tricks for riboswitch adaptation and tuning to new contexts.

5.5 Acknowledgements

This work was supported by the Natural Sciences and Engineering Council of Canada (NSERC) (418240 to J.P.). J.P. is a junior 2 FRQS research scholar. We wish to acknowledge support from Compute Canada for some of the bioinformatics searches of novel SAM-II clan riboswitches.

6 BIBLIOGRAPHIE DU CHAPITRE 5

- Baird NJ, Kulshina N & Ferre-D'Amare AR (2010) Riboswitch function: flipping the switch or tuning the dimmer? *RNA Biol* 7(3):328-332.
- Breaker RR (2012) Riboswitches and the RNA world. *Cold Spring Harb Perspect Biol* 4(2).
- Corbino KA, Barrick JE, Lim J, Welz R, Tucker BJ, Puskarz I, Mandal M, Rudnick ND & Breaker RR (2005) Evidence for a second class of S-adenosylmethionine riboswitches and other regulatory RNA motifs in alpha-proteobacteria. *Genome Biol* 6(8):R70.
- Dann CE, 3rd, Wakeman CA, Sieling CL, Baker SC, Irnov I & Winkler WC (2007) Structure and mechanism of a metal-sensing regulatory RNA. *Cell* 130(5):878-892.
- Doshi U, Kelley JM & Hamelberg D (2012) Atomic-level insights into metabolite recognition and specificity of the SAM-II riboswitch. *RNA* 18(2):300-307.
- Edwards AL, Reyes FE, Heroux A & Batey RT (2010) Structural basis for recognition of S-adenosylhomocysteine by riboswitches. *RNA* 16(11):2144-2155.
- Epshtein V, Mironov AS & Nudler E (2003) The riboswitch-mediated control of sulfur metabolism in bacteria. *Proc Natl Acad Sci U S A* 100(9):5052-5056.
- Fogolino M, Borne F, Bally M, Ball G & Patte JC (1995) A direct sulfhydrylation pathway is used for methionine biosynthesis in *Pseudomonas aeruginosa*. *Microbiology* 141 (Pt 2):431-439.
- Fuchs RT, Grundy FJ & Henkin TM (2006) The S(MK) box is a new SAM-binding RNA for translational regulation of SAM synthetase. *Nat Struct Mol Biol* 13(3):226-233.
- Furukawa K, Ramesh A, Zhou Z, Weinberg Z, Vallery T, Winkler WC & Breaker RR (2015) Bacterial riboswitches cooperatively bind Ni(2+) or Co(2+) ions and control expression of heavy metal transporters. *Mol Cell* 57(6):1088-1098.
- Gilbert SD, Rambo RP, Van Tyne D & Batey RT (2008) Structure of the SAM-II riboswitch bound to S-adenosylmethionine. *Nat Struct Mol Biol* 15(2):177-182.
- Huang L & Lilley DMJ (2018) Structure and ligand binding of the SAM-V riboswitch. *Nucleic Acids Res* 46(13):6869-6879.
- Kwon M & Strobel SA (2008) Chemical basis of glycine riboswitch cooperativity. *RNA* 14(1):25-34.
- Leyn SA, Suvorova IA, Kholina TD, Sherstneva SS, Novichkov PS, Gelfand MS & Rodionov DA (2014) Comparative genomics of transcriptional regulation of methionine metabolism in Proteobacteria. *PLoS One* 9(11):e113714.
- Mandal M & Breaker RR (2004) Gene regulation by riboswitches. *Nat Rev Mol Cell Biol* 5(6):451-463.
- McCown PJ, Corbino KA, Stav S, Sherlock ME & Breaker RR (2017) Riboswitch diversity and distribution. *RNA* 23(7):995-1011.
- McDaniel BA, Grundy FJ, Artsimovitch I & Henkin TM (2003) Transcription termination control of the S box system: direct measurement of S-adenosylmethionine by the leader RNA. *Proc Natl Acad Sci U S A* 100(6):3083-3088.

- Mirihana Arachchilage G, Sherlock ME, Weinberg Z & Breaker RR (2018) SAM-VI RNAs selectively bind S-adenosylmethionine and exhibit similarities to SAM-III riboswitches. *RNA Biol* 15(3):371-378.
- Montange RK & Batey RT (2008) Riboswitches: emerging themes in RNA structure and function. *Annu Rev Biophys* 37:117-133.
- Mulhbacher J & Lafontaine DA (2007) Ligand recognition determinants of guanine riboswitches. *Nucleic Acids Res* 35(16):5568-5580.
- Naghdi MR, Boutet E, Mucha C, Ouellet J & Perreault J (2020) Single Mutation in Hammerhead Ribozyme Favors Cleavage Activity with Manganese over Magnesium. *Noncoding RNA* 6(1).
- Naghdi MR, Smail K, Wang JX, Wade F, Breaker RR & Perreault J (2017) Search for 5'-leader regulatory RNA structures based on gene annotation aided by the RiboGap database. *Methods* 117:3-13.
- Poiata E, Meyer MM, Ames TD & Breaker RR (2009) A variant riboswitch aptamer class for S-adenosylmethionine common in marine bacteria. *RNA* 15(11):2046-2056.
- Stoddard CD, Montange RK, Hennelly SP, Rambo RP, Sanbonmatsu KY & Batey RT (2010) Free state conformational sampling of the SAM-I riboswitch aptamer domain. *Structure* 18(7):787-797.
- Sun A, Gasser C, Li F, Chen H, Mair S, Krasheninina O, Micura R & Ren A (2019) SAM-VI riboswitch structure and signature for ligand discrimination. *Nat Commun* 10(1):5728.
- Tomsic J, McDaniel BA, Grundy FJ & Henkin TM (2008) Natural variability in S-adenosylmethionine (SAM)-dependent riboswitches: S-box elements in bacillus subtilis exhibit differential sensitivity to SAM In vivo and in vitro. *J Bacteriol* 190(3):823-833.
- Trausch JJ, Xu Z, Edwards AL, Reyes FE, Ross PE, Knight R & Batey RT (2014) Structural basis for diversity in the SAM clan of riboswitches. *Proc Natl Acad Sci U S A* 111(18):6624-6629.
- Webb CH, Riccitelli NJ, Ruminski DJ & Luptak A (2009) Widespread occurrence of self-cleaving ribozymes. *Science* 326(5955):953.
- Weinberg Z, Regulski EE, Hammond MC, Barrick JE, Yao Z, Ruzzo WL & Breaker RR (2008) The aptamer core of SAM-IV riboswitches mimics the ligand-binding site of SAM-I riboswitches. *RNA* 14(5):822-828.
- Weinberg Z, Wang JX, Bogue J, Yang J, Corbino K, Moy RH & Breaker RR (2010) Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol* 11(3):R31.
- Winkler WC & Breaker RR (2003) Genetic control by metabolite-binding riboswitches. *ChemBiochem* 4(10):1024-1032.
- Zhang K, Li S, Kappel K, Pintilie G, Su Z, Mou TC, Schmid MF, Das R & Chiu W (2019) Cryo-EM structure of a 40 kDa SAM-IV riboswitch RNA at 3.7 Å resolution. *Nat Commun* 10(1):5511.

7 CARACTÉRISATION DES ÉLÉMENTS STRUCTURELS DU RIBOSWITCH SAM EN AMONT DU GÈNE *METK*

Lien entre le chapitre 5 et le chapitre 7 :

Dans le chapitre 5, on a présenté des nouveaux membres de riboswitch SAM-II/V qui nous amènent à proposer un nouveau consensus « universel » qui inclut des éléments structuraux jamais remarqués auparavant pour les familles de ce clan. Bien que ce chapitre confirme l'existence de plusieurs de ces nouveaux membres par des études *in vitro* et des études *in vivo*, grâce au plasmide conçu et décrit dans le chapitre 3, le rôle des nouveaux éléments structurels présent chez ces membres n'a pas été élucidé. Il est probable que cette diversité de structures secondaires contribue au raffinement de l'affinité ou des mécanismes de *switch* des riboswitchs pour la régulation des gènes en aval.

Le chapitre 7, qui est la dernière étude incluse dans ce mémoire, explore ces éléments structurels par la caractérisation plus approfondie du riboswitch SAM en amont du gène *metK*, un riboswitch modèle parmi les nouveaux membres découverts dans la deuxième étude (Chapitre 5). Ce riboswitch contient les deux éléments structurels caractéristiques de ce groupe de nouveaux membres du riboswitch SAM-II/V, soit une apparente tige compétitrice et une petite tige accessoire. Pour comprendre l'impact de ces deux éléments structurels sur le fonctionnement de l'aptamère de ce riboswitch, une série de mutants visant à déstabiliser ou à renforcer ces éléments structurels sont testés pour leur affinité pour SAM *in vitro* ainsi que pour leur changement de capacité à réguler *in vivo*.

7.1 Introduction

Comme le riboswitch SAM en amont du gène *metK* semble être une nouvelle variante des riboswitchs SAM-II, on a voulu l'étudier plus en profondeur pour valider sa structure. Pour décortiquer la nature de la relation entre différents éléments structurels présents dans la structure liée et non-liée, on visait à confirmer l'existence d'une de ces structures non-liées, que l'on présumait être en compétition avec l'aptamère.

Basé sur la théorie que le domaine aptamère d'un riboswitch est dynamique et transitionnel dans une plage de structures intermédiaires (Cruz & Westhof, 2009; Savinov *et al.*, 2014), une série de mutants, ayant pour but d'affaiblir ou de renforcer une des structures hypothétiques non-liées, ont été testés pour leur affinité avec SAM par la technique d'*in-line probing*. En visant à moduler les structures hypothétiques des mutants on obtiendrait des indices par rapport à l'importance de l'élément structural ciblé par chaque mutant, c'est-à-dire sa contribution à la liaison de SAM et au bon repliement de l'aptamère. L'aptamère de cette nouvelle version du riboswitch SAM présente deux éléments caractéristiques, soit une tige principale d'une plus grande taille que d'habitude (environ 9 pb en

conformation liée et présumément de 14 pb en conformation non-liée ou alternative), la grande tige P1; ainsi qu'une tige accessoire, la petite tige P2, absente dans les riboswitchs SAM rapportés à date. Nos études mutationnelles ciblaient ces deux éléments. De plus, en conformation liée notre riboswitch semble avoir deux pseudo-nœuds qui sont présents : PK-1 et PK-2.

L'hypothèse qu'on teste est que deux structures compétitionnent entre elles de manière à déterminer la capacité du riboswitch de réguler l'expression génique—soit la structure non-liée qui inclurait la tige qu'on a appelé « alternative », et la structure liée, l'aptamère, qui est stabilisé par les deux pseudo-nœuds. À notre surprise, nos résultats indiquent que la structure alternative ne compétitionne pas (ou peu) avec la structure liée, mais contribue plutôt à la liaison de son ligand SAM. Cette caractéristique d'une structure de l'aptamère non-liée révèle des indices sur les déterminants structuraux essentiels au mécanisme d'action des riboswitchs.

7.2 Matériels et méthodes

7.2.1 Amplification de l'ADN et construction de plasmides

Les construits pour les essais de *in-line probing* ont été produit tel que décrit dans le chapitre 5 en utilisant les oligonucléotides rapportés dans le tableau 7.1. La construction des plasmides a été effectuée par la méthode de l'assemblage par Gibson tel que décrit dans le chapitre 3. Les inserts, le vecteur de clonage, ainsi que les plasmides construits sont énumérés dans le tableau 7.1 et 7.2 avec les détails de clonage.

Tableau 7.1 Liste des oligonucléotides utilisés pour ce chapitre

Construit	Nom de l'oligonucléotide	Séquence 5'-3' ^{1,2}	Gabarit pour PCR et combinaison d'oligonucléotides
Parties pour essais de <i>in-line probing</i>			
<i>WT</i>	BS018-F	taatacgcactactataggctataatacgggcttc	PCR: BS018-F/BS019_R/(gDNA <i>B.thailandensis</i> E264)
	BS019_F	ggcgacgcttagcggat	
<i>inactive</i>	JP216_M4-F(c-alt)	taatacgcactactataggctataatacgggcttcgccgagAtaactgacaTc ttgcggggcgaagccggcg	PCR: JP216_M4-F(c-alt)/BS019_R/(gDNA <i>B.thailandensis</i> E264).
<i>P2 dis.</i>	JP213-M1-R	gctataatacgggcttcgccgagttaactgacaactgccccggaagccgg cggCGgctacgagcGccacggcgcaggaatccgctaaagcgtcgcc	BS018-F/JP213-M1-R/(PCR <i>WT</i>)
<i>P2 res.</i>	JP214-M2-R	gctataatacgggcttcgccgagttaactgacaactgccccggaagccgg cggCGgctacgagcGccacggcgcaggaatccgctaaagcgtcgcc	BS018-F/JP214-M2-R/(PCR <i>WT</i>)

$\Delta P2(55...69)$	VK026_M28-F	gctataatacgggcttcgcccagttaactgacaactgCGgggCGgaagccggcgacgCGcaggaatccgctaaagcgtcgcc	PCR par assemblage: VK026_M28-F/VK027-M28-R
	VK027-M28-R	ggcgacgcttagcggattcctgcgccgtgcgggcttcgccccgaagttgtcagtaa	
<i>P1 dis.1.</i>	JP219_M4(d)-F	gctataatacgggcttcgcccagttaactgacaTcttgcgggCGgaagccggcgggcgctacgagcGCCacggCGcaggaatccgctaaagcgtcgcc	PCR: JP219_M4(d)-F/BS019_R/(PCR WT)
<i>P1 dis.2.</i>	VK022_M26-F	gctataatacgggcttcgcccagttaactgacaTAttgcgggCGgaagccggcgggcgctacgagcGCCacggCGcaggaatccgctaaaAAgtcgcc	PCR par assemblage: VK022_M26-F/VK023.2_M26-R
	VK023.2_M26-R	ggcgacTTtttagcggattcctgcgccgtggcgctcgtagcggcgccgcttcgccccgcaATAtg	
<i>P1 bulge</i>	JP224_M8(d)-F	gctataatacgggcttcgcccagttaactgacaactgGCgggCGgaagccggcgggcgctacgagcGCCacggCGcaggaatcGCctaaagcgtcgcc	PCR: JP224_M8(d)-F/JP225_M8(d)-R/(PCR WT)
	JP225_M8(d)-R	ggcgacgcttagGCgattcctgcgccgtggcg	
Parties pour les clonages			
<i>Insert P2 dis.</i>	VK146_F	acgctctgtgacaattcgagaaagccgctataatacgggctt	PCR: VK146_F/VK147_R/(PCR P2 dis.)
	VK147_R	cgtagccagctccagagggcccgaagtggcgacgcttag	
<i>Insert P2 res.</i>	--	--	PCR: VK146_F/VK147_R/(PCR P2 res.)
<i>Insert $\Delta P2(55...69)$</i>	--	--	PCR: VK146_F/VK147_R/(PCR $\Delta P2(55...69)$)
<i>Insert P1 dis.2.</i>	VK148_R	cgtagccagctccagagggcccgaagtggcgacttttagc	PCR: VK146_F/VK148_R/(PCR P1 dis.2.)
<i>Insert P1 bulge</i>	--	--	PCR: VK146_F/VK147_R/(PCR P1 bulge)

¹Les bases mutés sont en majuscule.

Tableau 7.2 Liste des plasmides utilisées pour ce chapitre

Plasmide	Description	Source	Procédure de clonage
pP1.0	Vecteur de clonage pour les mutants	Chapitre 3	n/a
p180222-3-3 (pP1.2)	Construit <i>WT</i>	Chapitre 3	n/a
P1.6_1	Construit <i>inactive</i>	Chapitre 5	n/a
P1.4_3	Construit <i>P2 dis.</i>	Ce chapitre	Assemblage par Gibson du vecteur pP1.0 digéré avec Eco53KI et avec le produit PCR <i>Insert P2 dis.</i>
P1.5_1	Construit <i>P2 res.</i>	Ce chapitre	Assemblage par Gibson du vecteur pP1.0 digéré avec Eco53KI et avec le produit PCR <i>Insert P2 res.</i>
P1.3_1	Construit $\Delta P2(55...69)$	Ce chapitre	Assemblage par Gibson du vecteur pP1.0 digéré avec Eco53KI et avec le produit PCR <i>Insert $\Delta P2(55...69)$</i>
P1.7_3	Construit <i>P1 dis.2.</i>	Ce chapitre	Assemblage par Gibson du vecteur pP1.0 digéré avec Eco53KI et avec le produit PCR <i>Insert P1 dis.2.</i>
P1.8_1	Construit <i>P1 bulge</i>	Ce chapitre	Assemblage par Gibson du vecteur pP1.0 digéré avec Eco53KI et avec le produit PCR <i>Insert P1 bulge.</i>

7.2.2 Souches bactériennes et transformations

Toutes les transformations ont été faites comme décrit dans le chapitre 3. La liste des souches utilisés pour ce chapitre sont énumérées dans le tableau 7.3.

Tableau 7.3 Liste des souches utilisées pour ce chapitre

Souche	Catalogue du laboratoire	Description	Source
<i>E. coli</i> DH5 α (Chen <i>et al.</i> , 2018), 2018)	B1A4	Utilisé pour la transformation de plasmides	NCBI:txid668369
<i>E. coli</i> SM10 λ pir (Simon <i>et al.</i> , 1983)	B1A5	Utilisé pour la conjugaison des plasmides avec la souche <i>Burkholderia</i> sp. strain E264.	BCCM: LMBP 3889
<i>B. thailandensis</i> E264 (Kim <i>et al.</i> , 2005)	B1A7	Utilisé pour la conjugaison des plasmides intégratifs.	NCBI:txid271848
<i>B. thailandensis</i> E264/ <i>metK</i> IGR	B1E8; p180222-3-3/pP1.2	Souche pour les essais de luminescence	Chapitre 3
<i>B. thailandensis</i> E264/ <i>metK</i> mut.	B3G9; pP1.6-1-1-2	Souche pour les essais de luminescence	Chapitre 5
<i>E. coli</i> DH5 α /P2 <i>dis.</i>	B3B5; P1.4_1	Souche transformée pour l'extraction des plasmides séquencées.	Ce chapitre
<i>E. coli</i> DH5 α /P2 <i>res.</i>	B3A6 ; P1.5_1	Souche transformée pour l'extraction des plasmides séquencées.	Ce chapitre
<i>E. coli</i> DH5 α / Δ P2(55...69).	B1G8; P1.3_1	Souche transformée pour l'extraction des plasmides séquencées.	Ce chapitre
<i>E. coli</i> DH5 α /P1 <i>dis.2.</i>	B3B7; P1.4_3	Souche transformée pour l'extraction des plasmides séquencées.	Ce chapitre
<i>E. coli</i> DH5 α /P1 <i>bulge</i>	B1H5; P1.8_1	Souche transformée pour l'extraction des plasmides séquencées.	Ce chapitre

<i>E. coli</i> SM10λpir/P2 <i>dis.</i>	B4A2; P1.4-3-2	Souche pour la conjugaison d'un plasmide séquencée avec la souche <i>Burkholderia</i> sp. strain E264.	Ce chapitre
<i>E. coli</i> SM10λpir/P2 <i>res.</i>	B3C4; P1.5-1-1	Souche pour la conjugaison d'un plasmide séquencée avec la souche <i>Burkholderia</i> sp. strain E264.	Ce chapitre
<i>E. coli</i> SM10λpir/ΔP2(55...69).	B3C2; P1.3-1-1	Souche pour la conjugaison d'un plasmide séquencée avec la souche <i>Burkholderia</i> sp. strain E264.	Ce chapitre
<i>E. coli</i> SM10λpir/P1 <i>dis.2.</i>	B4A2; P1.4-3-2	Souche pour la conjugaison d'un plasmide séquencée avec la souche <i>Burkholderia</i> sp. strain E264.	Ce chapitre
<i>E. coli</i> SM10λpir/P1 <i>bulge</i>	B3D1; P1.8-1-1	Souche pour la conjugaison d'un plasmide séquencée avec la souche <i>Burkholderia</i> sp. strain E264.	Ce chapitre
<i>B. thailandensis</i> E264/P2 <i>dis.</i>	B4A9; P1.4-3-2-3	Souche pour les essais de luminescence	Ce chapitre
<i>B. thailandensis</i> E264/P2 <i>res.</i>	B3G7; P1.5-1-1-3	Souche pour les essais de luminescence	Ce chapitre
<i>B. thailandensis</i> E264/ΔP2(55...69).	B3G2; P1.3-1-1-1	Souche pour les essais de luminescence	Ce chapitre
<i>B. thailandensis</i> E264/P1 <i>dis.2.</i>	B4A9; P1.4-3-2-3	Souche pour les essais de luminescence	Ce chapitre
<i>B. thailandensis</i> E264/P1 <i>bulge</i>	B3H7 ; P1.8-1-1-3	Souche pour les essais de luminescence	Ce chapitre

7.2.3 Essais *in-line* et essais *in vivo* de luminescence

Les essais *in line*, les essais *in vivo*, ainsi que l'analyse des données ont été effectués tel que décrit dans le chapitre 5.

7.3 Résultats et discussion

Les mutants du riboswitch ont été conçus pour favoriser l'équilibre vers une structure à la fois (soit la tige alternative prolongeant P1, soit les pseudo-nœuds de la structure de l'aptamère), pour améliorer ou réduire l'affinité pour SAM tout en faisant attention à ne pas affecter la pochette de liaison. Les structures des mutants correspondants sont présentées dans la Figure 7.1. Nos résultats sur la contribution des différents éléments structuraux à la stabilité thermodynamique estimée, à l'affinité pour le ligand *in vitro* ainsi qu'au fonctionnement de l'aptamère *in vivo* sont résumés dans le Tableau 7.4 avec les structures des mutants correspondants dans la Figure 7.1. Ces résultats démontrent que certains mutants qui visaient à défaire la structure alternative (*P1 dis.1*) n'amélioreraient pas l'affinité de l'aptamère pour SAM comme attendu, mais augmentait le K_D par plus de 100x (abolissant pratiquement l'interaction avec SAM), c'est à dire que notre hypothèse initiale était infirmée. Un autre mutant, *P1 dis.2*, qui avait aussi pour but de défaire la tige alternative comme *P1 dis.1*, semble démontrer une amélioration d'affinité pour SAM. Ce mutant a été conçu pour affecter la structure alternative tout en assurant la stabilité de l'aptamère, en répliquant la boucle d'un aptamère SAM-V rapporté par Poiata *et al.* dans la bactérie marine, *Candidatus Pelagibacter* (Poiata *et al.*, 2009). Ce mutant concorde donc mieux avec l'hypothèse de départ, suggérant un modèle de repliement impliquant tout de même un certain degré de compétition entre la conformation « aptamère » et la structure alternative.

Tableau 7.4 Sommaire de la caractérisation par étude mutationnelle du riboswitch SAM en amont de *metK* chez *B. thailandensis* E264.

Construit	Rationale	ΔG de la structure alternative la plus probable ¹	ΔG de la structure alternative la plus stable (MFE) ¹	ΔG de la structure aptamère ²	K_D <i>in vitro</i> (μM) et 95% CI ³	Ratio d'induction <i>in vivo</i> ⁴
Contrôles						
<i>WT</i>	Contrôle positif	-37.2	-37.2	-81.94	0.336 (0.057 à 0.869)	1.87 ± 0.05
<i>inactive</i>	Contrôle négatif	--	--	--	>100 (très large), interaction non-significative	0.94 ± 0.03
Mutants de la petite tige P2						
<i>P2 dis.</i>	Déstabilisation de P2 par dissociation de 2 pb	-41.3	-41.3	-71.62	2.329 (1.829 à 2.911)	1.20 ± 0.04
<i>P2 res.</i>	Rétablissement de P2 complémentaire à <i>P2 dis.</i>	-33.6	-33.6	-81.94	0.3676 (0.2075 à 0.6158)	1.26 ± 0.03
$\Delta P2(55...69)$	Délétion de P2	-25.6	-20.4	-57.38	3.951 (2.845 à 5.302)	1.69 ± 0.21
Mutants de la grande tige P1						
<i>P1 dis.1.</i>	Déstabilisation de la structure alternative par dissociation de 1 pb	-41.0	-41.0	-82.54	85.15 (59.97 à 155.1), interaction non-significative	Clonage échoué
<i>P1 dis.2.</i>	Déstabilisation de la structure alternative par dissociation de 1 pb	-35.0	-35.0	-83.75	<0.001 (inconnu à 0.1847)	1.26 ± 0.08
<i>P1 bulge</i>	Renforcement de la structure alternative par le déplacement de la boucle vers la base de P1	-44.9	-44.9	-80.76	>100 (très large), interaction non-significative	0.99 ± 0.04

¹La structure la plus probable est considérée comme la structure *centroïde* du cluster le plus statistiquement probable. Le ΔG de cette structure et de la structure MFE sont prédits par Sfold/Srna pour des conditions à 37°C et 1M de NaCl.

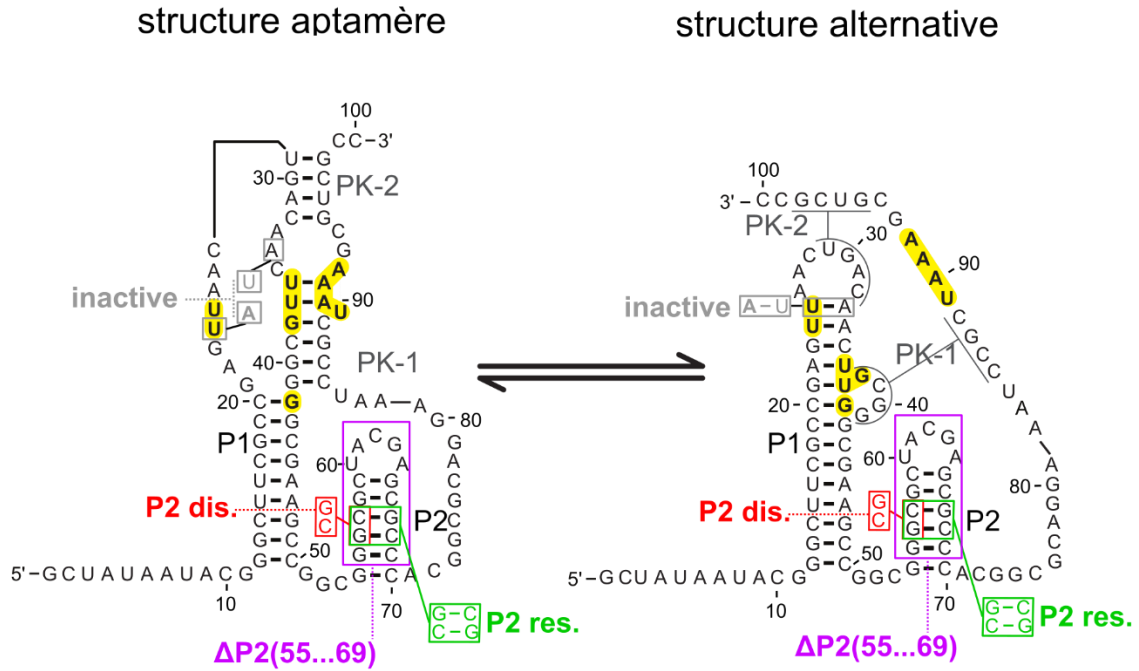
²Valeur générée par MC-Fold

³Valeurs calculées avec le logiciel GraphPad Prism 8.4.3 à partir des résultats de *in-line probing* en utilisant l'équation pour Non-linear regression/specific binding with Hill-slope. CI est l'intervalle de confiance de 95%.

⁴Le ratio d'induction (FI) est calculé tel que dans l'article du chapitre 3

(A)

Mutants de la petite tige P2



(B)

Mutants de la grande tige P1

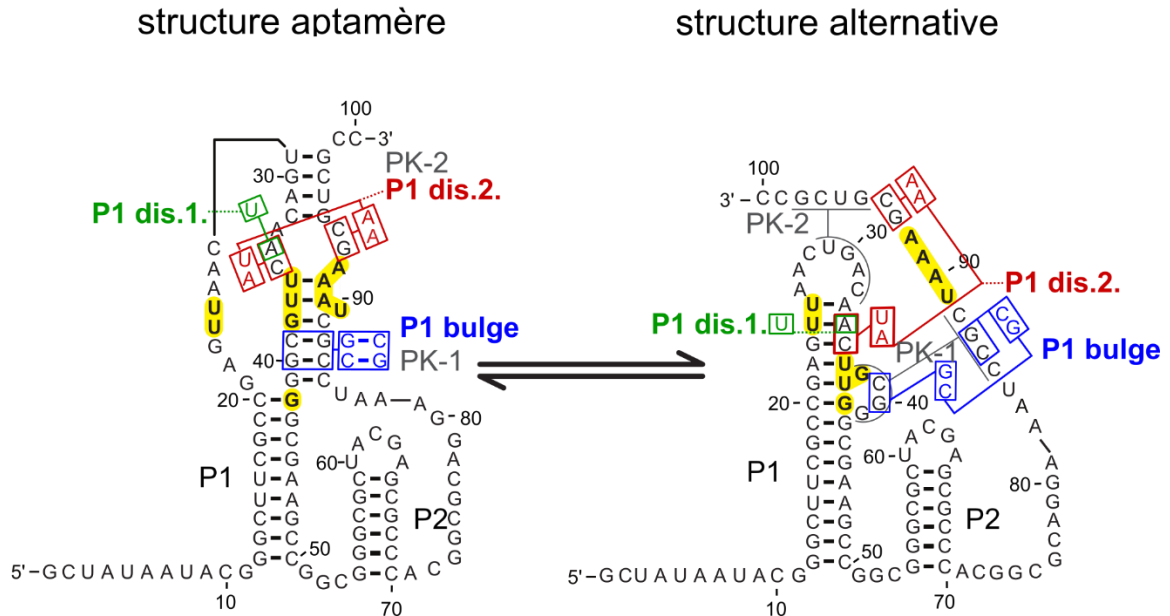


Figure 7.1 Divers mutants de l'aptamère du riboswitch SAM en amont de *metK* chez *B. thailandensis* E264.

En (A) les mutants visant à étudier la petite tige de l'aptamère et le mutant du contrôle négatif sont représentés et en (B), les mutant visant à étudier la grande tige P1 sont représentés. Les pseudo-nœuds sont indiqués par PK-1 et PK-2.

Pour le mutant *P1 bulge*, on a essayé de renforcer la structure alternative par le déplacement de la boucle de la tige principale vers la base de cet tige (Figure 7.2). La courbe de liaison de ce mutant ressemble plus à deux courbes distinctes, ayant des K_D différents, soit de 0.100 μM pour la population 1 et de $>100 \mu\text{M}$ pour la population 2 (Figure 7.3). Puisque ce résultat et d'autres résultats *in vitro* d'autres mutants mentionnés dans cette section ne semblent pas tous concorder avec l'hypothèse que la structure alternative compétitionnerait avec l'aptamère, on croit que la structure alternative serait un précurseur important pour la voie de repliement vers la structure de l'aptamère. Ceci supporte la spéculation que les deux courbes apparentes seraient dues à l'existence de deux populations de conformations dont une est d'une structure alternative transitoire qui puisse se replier correctement en structure aptamère pour lier SAM (population 1) et une autre qui est une structure alternative renforcée ou métastable, incapable de transitionner en aptamère (population 2), causant une baisse importante d'affinité (c'est-à-dire un K_D très élevé, voire une absence d'interaction). L'alignement des boucles de la structure alternative, qui ressemblerait à la structure de la population 1, permet une voie directe pour la transition au bon repliement pour la liaison de SAM. Pour la population 2, qui aurait une boucle déplacée, il est possible que les interactions du pseudo-nœud PK-2 ne puissent plus avoir lieu et il y aurait alors une perte de bon repliement.

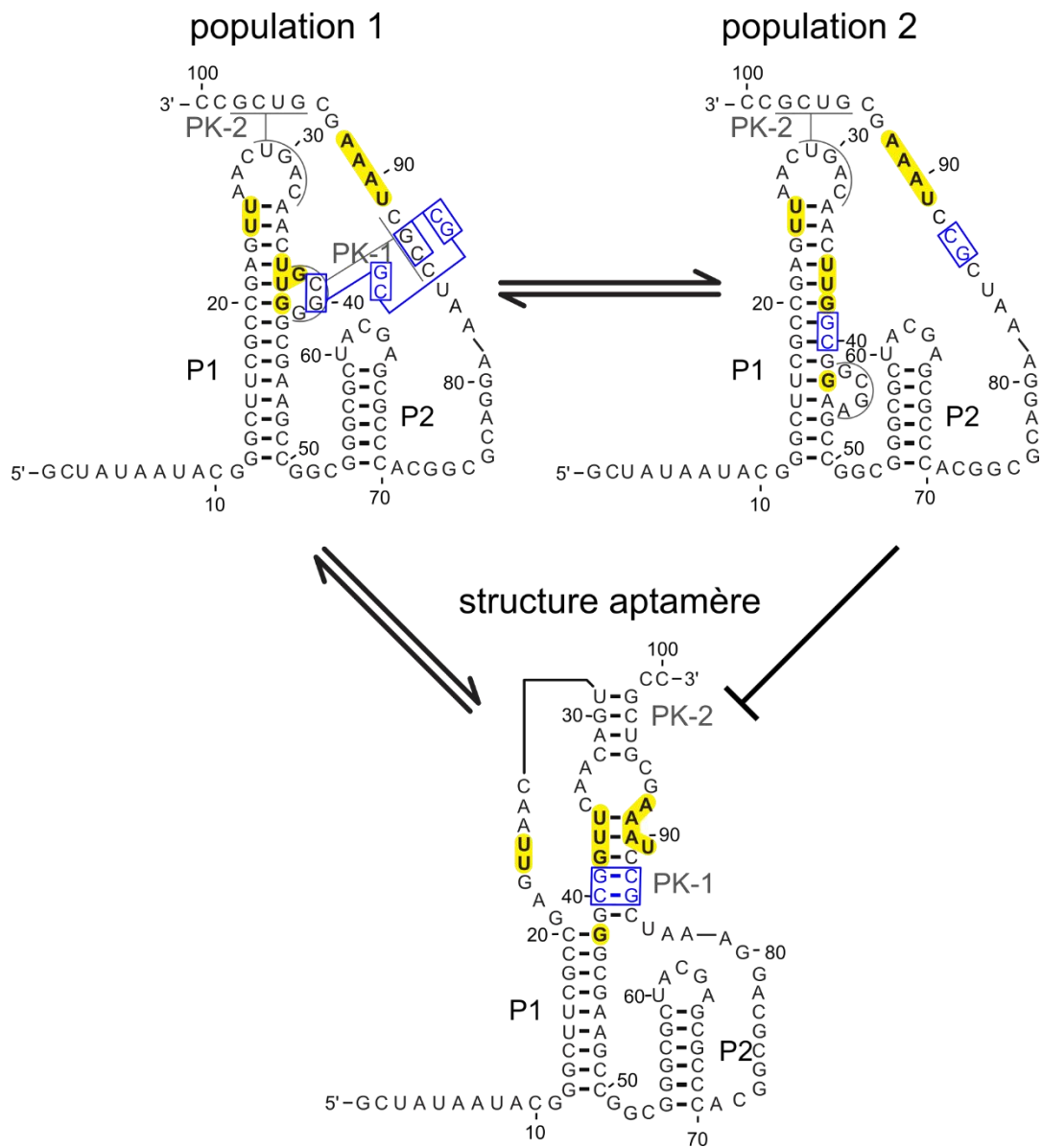


Figure 7.2 Restructuration du mutant P1 bulge

Les bases noires encadrées de la population 1 sont les bases originales de la structure *WT*, tandis que les bases bleues encadrées sont les mutations conçues pour le déplacement du bulge de la tige P1.

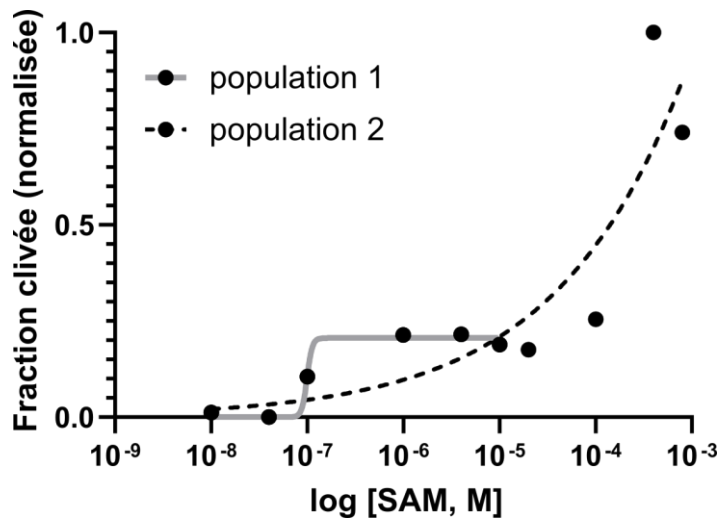


Figure 7.3 Courbe de K_D générée à partir d'une expérience de *in line probing* pour le mutant *P1 bulge*.

L'allure des points suggère une possibilité de l'existence de 2 populations structurales, chacune avec une affinité pour SAM différente.

L'étude *in silico* par rapport à la thermostabilité des différents mutants en conformation liée et non-liée a été faite en utilisant Sfold pour les structures non-liées transitoires et/ou métastables et MC-Fold pour les structures liées, soit les aptamères. Ces résultats combinés avec les résultats *in vitro* sont combinés dans la Figure 7.4 qui démontre la distribution des construits selon leurs rangs relatifs en thermostabilité et en affinité pour SAM, ainsi que leur population statistique.

Distribution des structures des mutants

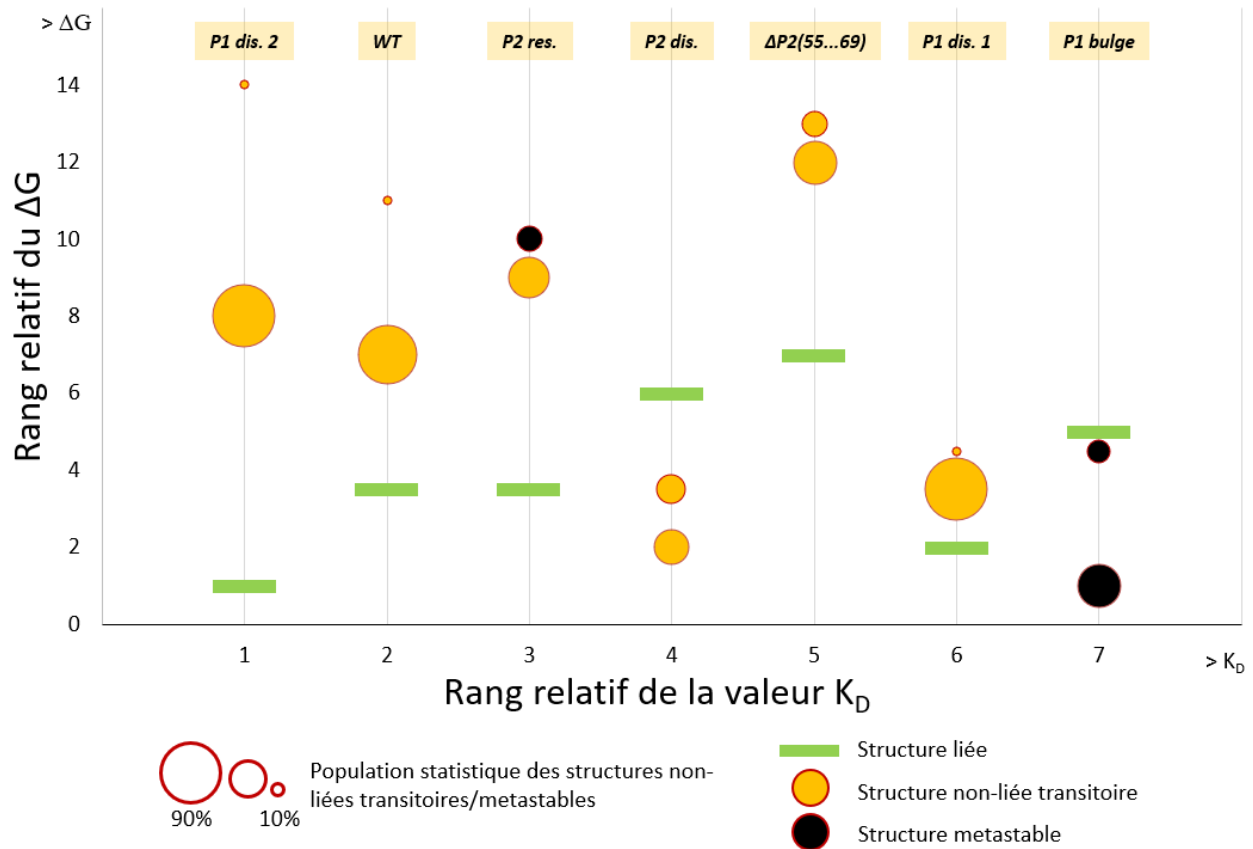


Figure 7.4 Distribution des structures des mutants et du WT selon leur population statistique, leur stabilité thermodynamique, et de leur affinité pour SAM.

Pour conclure sur la partie *in silico* et *in vitro*, mes résultats suggèrent que notre riboswitch et peut-être d'autres structures d'ARN régulateurs ont évolué sous l'effet d'un réseau d'interactions (dans un même ARN) de type protagoniste et antagoniste ce qui permettrait d'ajuster leur capacité de réguler et/ou d'avoir un processus de repliement (*folding pathway*) plus robuste. Les interactions protagonistes sont celles qui promouvoient le bon repliement de l'aptamère ou le bon repliement des structures non-liées transitoires et les interactions antagonistes sont celles qui promouvoient le repliement des structures métastables qui ne permettent pas la formation de l'aptamère. Le bilan de toutes ces structures pour un riboswitch donné est ce qui lui attribue son empreinte unique de régulation. Une représentation de ce réseau est présentée dans la Figure 7.5. Dans une étude ultérieure, il a été démontré *in vitro* qu'un riboswitch SAM II transitionne de l'ARN non replié (quelques secondes), à de l'ARN qui change rapidement entre une structure tige-boucle et une structure pseudo-nœud (Haller *et al.*, 2011). Il est suggéré que la liaison du ligand au riboswitch SAM-II se produit par un processus de sélection conformationnel, où un ensemble de structures est échantillonné pour permettre la reconnaissance du ligand (Chen *et al.*, 2012; Haller *et al.*, 2011).

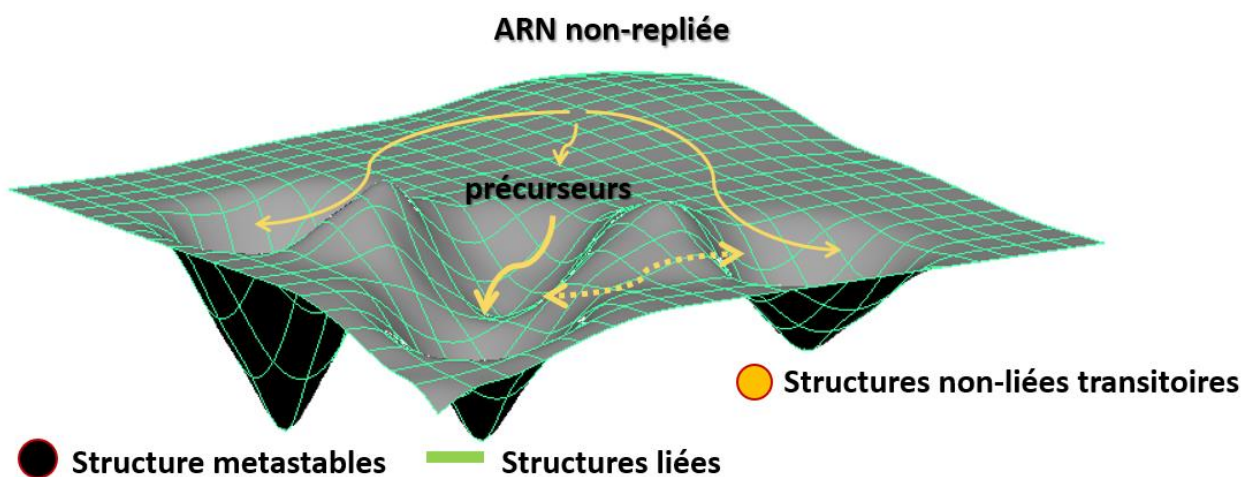


Figure 7.5 Représentation d'un réseau d'intra/interaction de la distribution de la stabilité des structures d'ARN régulateurs pour une séquence donnée.

Chaque creux représente une structure thermodynamiquement stable et la surface plane représente l'état thermodynamique instable des structures d'ARN en transition. Les flèches représentent le sens des transitions structurales hypothétiques. Le « creux » de la structure liée serait théoriquement rendu encore plus profond suite à la liaison de SAM qui stabiliserait cette conformation. Cette image créée avec l'aide du logiciel Maya de Autodesk vise à donner un exemple et n'est pas une quantitative de la distribution des stabilités des différentes structures d'ARN.

De plus, l'impact des mutations a été évalué *in vivo* en adaptant le plasmide pVK-f-*lux* pour l'étude des séquences IGR mutants de *metK* (Tableau 7.1). Dans les conditions testées—milieu de préculture de LB, milieu d'expression de 0.5X M9-MM, les antibiotiques comme précisé dans l'article du chapitre 5, et une concentration de méthionine de soit 0 ou de 0.1 mM—tous les mutants semblaient perdre complètement ou en partie leur fonction de régulateur, sauf celui du mutant $\Delta P2(55...69)$, qui avait une FI proche de 2 comme le *WT*. Les mutants *inactive* et *P1 bulge* ont perdu leur fonction régulatrice complètement tandis que *P2 dis.*, *P2 res.* et *P1 dis.2* ont gardé partiellement leur capacité de réguler (FI~1.2). La luminescence maximale pour tous les construits a été atteinte au milieu de la phase exponentielle. Les résultats *in vivo* démontrent que la tige accessoire n'est pas nécessaire pour l'obtention d'une fonction régulatrice du riboswitch SAM et que possiblement c'est une structure qui permettrait de raffiner le repliement du riboswitch. Possiblement, cette tige aurait un effet sur la dynamique soit en raffinant la vitesse de repliement de l'aptamère, des structures non-liées transitoires, ou des structures métastables, ou encore tout simplement la probabilité de se replier correctement, tout ça pour assurer une réponse optimale régulatrice de ce riboswitch dans le contexte dans lequel se retrouve la bactérie. Comme les résultats *in vitro* ne prenaient en compte que la partie aptamère du riboswitch, il est possible que les mutations aient affecté le repliement de la plateforme d'expression *in vivo* et dans ce cas il serait difficile de tirer des conclusions sur l'impact de ces mutants sur le niveau de régulation. D'ailleurs, au-delà d'un FI semblable, nous avons noté une expression diminuée de moitié pour le mutant avec la tige P2 déléetée, suggérant que la plateforme d'expression a été affectée, même si l'aptamère ne l'avait pas, ou peu, été. Aussi, il se peut que les conditions de l'essai *in vivo* ne

permettent pas d'atteindre des niveaux suffisamment bas de méthionine et de SAM pour observer une modulation de l'expression. Il a été démontré dans l'article du chapitre 5 que les plus grands FI ont été achevés lorsque tous les milieux de culture de l'essai, incluant le milieu de préculture et le milieu d'expression, ne contenaient pas de soufre ajouté (0.5X M9-MM-S).

8 BIBLIOGRAPHIE DU CHAPITRE 7

- Chen B, Zuo X, Wang YX & Dayie TK (2012) Multiple conformations of SAM-II riboswitch detected with SAXS and NMR spectroscopy. *Nucleic Acids Res* 40(7):3117-3130.
- Chen J, Li Y, Zhang K & Wang H (2018) Whole-Genome Sequence of Phage-Resistant Strain *Escherichia coli* DH5alpha. *Genome Announc* 6(10).
- Cruz JA & Westhof E (2009) The dynamic landscapes of RNA architecture. *Cell* 136(4):604-609.
- Haller A, Rieder U, Aigner M, Blanchard SC & Micura R (2011) Conformational capture of the SAM-II riboswitch. *Nat Chem Biol* 7(6):393-400.
- Kim HS, Schell MA, Yu Y, Ulrich RL, Sarria SH, Nierman WC & DeShazer D (2005) Bacterial genome adaptation to niches: divergence of the potential virulence genes in three *Burkholderia* species of different survival strategies. *BMC Genomics* 6:174.
- Poiata E, Meyer MM, Ames TD & Breaker RR (2009) A variant riboswitch aptamer class for S-adenosylmethionine common in marine bacteria. *RNA* 15(11):2046-2056.
- Savinov A, Perez CF & Block SM (2014) Single-molecule studies of riboswitch folding. *Biochim Biophys Acta* 1839(10):1030-1045.
- Simon R, Priefer U & Pühler A (1983) A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. *Bio/Technology* 1(9):784-791.

9 DISCUSSION GÉNÉRALE ET CONCLUSION

9.1 Caractérisation de pVK-f-*lux* comme outil pour l'étude des séquences intergéniques

Un plasmide modulaire, pVK-f-*lux*, pour l'étude des IGR a été conçu avec les caractéristiques suivantes:

- Deux sites de clonages multiples (MCS) : un à chaque extrémité (5' et 3') de la région du promoteur, ainsi qu'un site RE de type IIS qui coupe à l'extérieure de la séquence de reconnaissance de l'enzyme de restriction AarI permettant une fusion traductionnelle des riboswitchs et d'autres types d'ARN régulateurs.
- Possibilité d'utiliser les régions MCS comme des régions standards de chevauchement pour le clonage de type Gibson, permettant de facilement échanger les différents promoteurs et éléments d'ARN régulateurs.
- Une construction de base comprenant le promoteur constitutif, P1, le riboswitch fluor de *P. syringae* et la cassette *luxCDABE*, qui comprends les gènes nécessaires pour produire la luciférase avec une expression réduite, mais inductible, pour réduire la toxicité potentielle de promoteurs forts.

Les caractéristiques de pVK-f-*lux* rendent cet outil moléculaire novateur pour le clonage en fusion traductionnelle et pour l'utilisation des promoteurs forts et/ou non-inductibles dans un contexte inductible. Ce plasmide permet de remédier, jusqu'à certaines limites, à la toxicité possible causée par la surexpression du gène rapporteur en régulant le niveau d'expression de ce dernier via le riboswitch fluor. Dans nos études, le riboswitch fluor a permis la détection du clonage réussi du promoteur constitutif P1, mais pas celui du P_{S7} indiquant qu'un riboswitch avec un plus grand pouvoir répressif serait nécessaire. On a aussi confirmé que l'utilisation du fluor jusqu'à des concentrations de 15 mM permettait quand même la transformation des cellules d'*E. coli* compétents, bien qu'avec des rendements inférieurs. Ce résultat suggère qu'il y a une possibilité de réguler l'expression de gènes de constructions synthétique sans utiliser de promoteurs inductibles et ceci même en état de stress comme en présence de concentrations de fluor.

En explorant la possibilité d'utiliser le riboswitch fluor comme un élément tampon pour le clonage des promoteurs forts, on a découvert une limite importante relié à la portabilité de ce riboswitch à travers les différents hôtes bactériens. Nous avons observé une inversion de régulation d'une espèce à l'autre— notamment le riboswitch fluor, provenant de *B. thailandensis* E264, cause une surexpression du gène dans l'organisme hôte en présence de fluor (tel qu'attendu), mais une répression dans les mêmes conditions dans *E. coli* DH5alpha et SM10λpir. À notre connaissance, c'est la première fois qu'une telle inversion de régulation est rapportée. Plus d'études sont nécessaires pour comprendre la cause de ce

phénomène, mais comme l'effet de la régulation par le fluor est présent chez les deux espèces, on pense que ce serait lié à un repliement différent de la plateforme d'expression d'une espèce à l'autre, causé soit par des différences de conditions chimiques à l'intérieur des cellules ou par des différences dans la cinétique des enzymes participant à la transcription.

Chez les bactéries, la polymérase d'ARN (RNAP) est constituée de cinq sous-unités qui généralement sont conservées à travers les bactéries, mais dont une, la sous-unité σ aussi nommé facteur sigma, permettant à la RNAP de reconnaître le promoteur (Parker, 2001) peut varier d'une espèce à l'autre. Il est connu qu'une espèce bactérienne peut contenir même plusieurs différents facteurs sigma qui reconnaissent différentes séquences appartenant aux promoteurs. La vitesse de transcription des RNAP procaryotes est en moyenne autour de 50 nts/seconde, mais varie selon l'occurrence de sites de pauses (Parker, 2001). Les sites de pauses des RNAP ont été démontrés comme étant des points critiques pour la régulation par les riboswitchs, incluant pour le riboswitch de thiamine pyrophosphate (Chauvier *et al.*, 2017; Uhm *et al.*, 2018) et pour le riboswitch FMN (Wickiser *et al.*, 2005). Les sites de pause ont été auparavant étudiés chez *E. coli* (Gabizon *et al.*, 2018; Larson *et al.*, 2014), mais pas chez *Burkholderia*, au meilleur de nos connaissances, donc il est difficile d'estimer l'ampleur des différences cinétiques de RNAP entre ces espèces. En théorie, si la reconnaissance pour les sites de pauses varie entre les RNAP de différentes espèces il est possible que la position de ces sites joue un rôle important dans le repliement de l'ARN au fur et à mesure qu'il est transcrit. Dépendamment de la conformation, une même séquence d'ARN régulateur, en théorie, pourrait avoir un mécanisme de régulation différent.

9.2 Étude des riboswitchs SAM chez *B. thailandensis* E264

À la suite de la confirmation de l'affinité pour SAM de plus de cinq aptamères des riboswitchs tirés chez diverses bactéries, trois instances de riboswitches SAM chez *B. thailandensis* E264 ayant toutes des structures différentes ont été étudiées plus en détail dans le but d'élucider leurs rôles dans le métabolisme de SAM.

Pour l'étude *in vivo* avec un gène rapporteur en bactérie, en temps réel, ces trois instances de riboswitchs SAM chez *B. thailandensis* E264—soit en amont des gènes *metK*, *metX*, et *metZ*—on a exploité le site de RE de type IIS, ainsi que l'autoinductibilité de l'opéron *lux* de pVK-f-*lux* qui contient tous les gènes d'enzymes nécessaires pour la production de la luciférase sans rajout des métabolites exogènes. Il est primordial pour l'étude des ARN régulateurs de conserver l'environnement structural en acides nucléiques tel que retrouvé dans l'hôte original. Souvent des construits créés à partir des plasmides non-adaptés pour le clonage en fusion traductionnel contiennent soit des bouts de séquences non-essentiels à l'expérience, ou des séquences omise du gène rapporteur dû à l'utilisation d'un site de restriction coupé. Notre plasmide pVK-f-*lux* est conçu pour le clonage en fusion traductionnelle

directement à partir du premier codon du premier gène de l'opéron *lux* ce qui permet non seulement de diminuer les chances d'un mauvais repliement des riboswitchs lors de la transcription/traduction simultanée, mais aussi de diminuer des effets artéfactuels causés par des assemblages non-idéaux.

On a découvert des dynamiques de régulation différentes pour chaque occurrence du riboswitch SAM chez *B. thailandensis* E264, indiquant l'importance des riboswitchs dans la synchronisation de la régulation à voies multiples pour la synthèse de ce métabolite et de la méthionine. On n'a pas réussi à atteindre une concentration de SAM suffisamment basse pour observer une régulation claire du riboswitch en amont de *metZ* probablement parce que même dans les conditions très limitantes imposées à la bactérie, cette dernière conserve un niveau de SAM, nécessaire pour sa survie, qui serait possiblement au-dessus des concentrations maximales pour que le riboswitch soit en conformation ON.

Pour mieux caractériser ce riboswitch, une stratégie potentielle pourrait être l'utilisation des lysats cellulaires pour exprimer des fusions de riboswitch et gènes rapporteurs, à la place des bactéries vivantes, dans un système *in vitro* de transcription/traduction. Ceci permettrait aussi de découpler les effets métaboliques de la production de SAM de l'effet de la régulation du riboswitch. Toutefois, comme démontré dans le chapitre 2, la portabilité des riboswitchs à travers les organismes n'est pas garantie, donc un lysat cellulaire provenant de *B. thailandensis* serait idéal. Comme les lysats cellulaires ont été préparés avec succès pour l'étude des organismes eucaryotes et procaryotes, incluant des organismes modèles tel que *Escherichia coli*, *Vibrio natriegens* et *Bacillus subtilis* (Tinafar *et al.*, 2019), mais aussi des organismes non-modèles tel que *Bacillus megaterium* (Moore *et al.*, 2018), *Streptomyces lividans* (Li *et al.*, 2017), et d'autres, il serait probable d'optimiser un protocole pour l'obtention d'un lysat provenant de *B. thailandensis*. Il faudrait cependant remédier au problème initial, soit baisser le niveau de SAM dans la réaction. Ceci pourrait être achevé soit par la purification des composantes du système, soit en introduisant dans la réaction une quantité d'ADN optimal d'un aptamère SAM en aval d'un promoteur, par exemple une de nos constructions d'aptamères SAM confirmés dans l'article du chapitre 5. De cette manière l'excès de ce métabolite sera lié en solution et ne se liera pas au riboswitch SAM dans le circuit du gène rapporteur utilisé.

L'élimination de toutes fonctions métaboliques à part la transcription et la traduction, lors de l'étude des riboswitchs SAM, permettrait aussi possiblement de supporter la régulation de type gradateur. Une des limites des expériences de *in-line probing* est qu'elles sont réalisées « à l'équilibre » durant 40 h et le plus souvent avec la partie aptamère seulement. Elles ne sont donc pas tout-à-fait représentatives de la modulation de l'expression. Cette limite ne sera plus pertinente dans les expériences utilisant les lysats cellulaires car les expériences seront faites à des échelles de temps similaires et avec les constructions représentant les riboswitchs complets, et non pas seulement la partie aptamère.

9.3 Conclusion

Nos études nous ont permis de mettre en place une stratégie de clonage en utilisant pVK-f-*lux* pour la construction modulaire des promoteurs et des éléments régulateurs d'ARN ainsi que pour le clonage des promoteurs forts, lorsque la toxicité de la surexpression du gène rapporteur peut affecter le succès du clonage. De plus on a utilisé ce plasmide pour confirmer la régulation *in vivo* de trois versions du riboswitch SAM chez *B. thailandensis* E264 et pour l'étude mutationnelle de la structure du riboswitch *metK* pour compléter nos résultats de *in-line probing*.

Nos études ont permis d'approfondir certaines de nos connaissances sur le riboswitch SAM, et de confirmer l'importance de la structure alternative pour la formation du pseudo-nœud, mais des questions importantes restent à explorer dans le domaine. Parmi celle-ci :

- Comment un ligand est reconnu par un riboswitch—est-ce le ligand qui induit le changement de conformation, ou est-ce le ligand qui fige une conformation transitoire qui est capable de le lier, ou serait-il un mélange des deux effets?
- Est-ce que c'est la vitesse de formation des structures ou la stabilité thermodynamique des structures qui orchestre le changement structural du *switch*?
- Comment les autres processus cellulaires affectent la formation et donc le fonctionnement du riboswitch?
- Comment différentes plateformes d'expression évoluent-elles pour des aptamères identiques?

Les réponses à ces questions aideront non-seulement à comprendre certains des fondements de l'évolution de la régulation et des ARNnc, mais pourraient aussi ouvrir la voie aux progrès pour l'utilisation efficace des riboswitchs en tant qu'outils pour de nouveaux systèmes en biologie synthétique.

10 BIBLIOGRAPHIE DU CHAPITRE 9

- Chauvier A, Picard-Jean F, Berger-Dancause JC, Bastet L, Naghdi MR, Dube A, Turcotte P, Perreault J & Lafontaine DA (2017) Transcriptional pausing at the translation start site operates as a critical checkpoint for riboswitch regulation. *Nat Commun* 8:13892.
- Gabizon R, Lee A, Vahedian-Movahed H, Ebright RH & Bustamante CJ (2018) Pause sequences facilitate entry into long-lived paused states by reducing RNA polymerase transcription rates. *Nat Commun* 9(1):2930.
- Larson MH, Mooney RA, Peters JM, Windgassen T, Nayak D, Gross CA, Block SM, Greenleaf WJ, Landick R & Weissman JS (2014) A pause sequence enriched at translation start sites drives transcription dynamics in vivo. *Science* 344(6187):1042-1047.
- Li J, Wang H, Kwon YC & Jewett MC (2017) Establishing a high yielding streptomyces-based cell-free protein synthesis system. *Biotechnol Bioeng* 114(6):1343-1353.
- Moore SJ, MacDonald JT, Wienecke S, Ishwarbhai A, Tsipa A, Aw R, Kyllis N, Bell DJ, McClymont DW, Jensen K, Polizzi KM, Biedendieck R & Freemont PS (2018) Rapid acquisition and model-based analysis of cell-free transcription-translation reactions from nonmodel bacteria. *Proc Natl Acad Sci U S A* 115(19):E4340-E4349.
- Parker J (2001) RNA Polymerase. *Encyclopedia of Genetics*, 10.1006/rwgn.2001.1135. p 1746-1747.
- Tinafar A, Jaenes K & Pardee K (2019) Synthetic Biology Goes Cell-Free. *BMC Biol* 17(1):64.
- Uhm H, Kang W, Ha KS, Kang C & Hohng S (2018) Single-molecule FRET studies on the cotranscriptional folding of a thiamine pyrophosphate riboswitch. *Proc Natl Acad Sci U S A* 115(2):331-336.
- Wickiser JK, Winkler WC, Breaker RR & Crothers DM (2005) The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. *Mol Cell* 18(1):49-60.

Supplementary Material

for the manuscript

Fluoride riboswitch-based dampening of gene expression for cloning potent promoters

Vesta Korniakova¹, Marie-Christine Groleau¹, Eric Déziel¹ and Jonathan Perreault¹

¹ INRS - Centre Armand-Frappier Santé Biotechnologie, 531 boulevard des Prairies, Laval, QC, H7V 1B7, Canada,

To whom correspondence should be addressed:

Phone: 450-687-5010 x4411, fax: 450-686-5301, e-mail: jonathan.perreault@iaf.inrs.ca

Key words: plasmid, reporter, promoter, regulatory region, fluoride riboswitch, *Burkholderia*, *Pseudomonas*, *E. coli*, luciferase, toxicity, cloning, induction, repression, translation fusion

Table of contents

Table S1. List of oligonucleotides used in this study.	81
Table S2. Plasmids used in this study.....	89
Table S3. Strains used in this study.	92
Figure S1: Inactive mutant transformants of the P_{S7} ribosomal promoter-lux construct may indicate lux overexpression lethality in <i>E. coli</i> DH5 α	96
Figure S2: Sequence of the optimized cloning area of pVK-f-lux.....	99
Figure S3: Transformation of heat-shocked <i>E. coli</i> is possible in high fluoride concentrations.	
100	
Table S4. P_{S7} promoter cloning sequencing results	101
Figure S4: Regulation thresholds of the <i>B. thailandensis</i> fluoride riboswitch in <i>B. thailandensis</i> E264 and in <i>E. coli</i> SM10 λ pir.	102
Table S5. pVK-f2-lux (P1 + <i>B. thai</i> F RS-lux) plasmid production and lux expression specifications in <i>E. coli</i> DH5 α and <i>E. coli</i> SM10 λ pir.	104
Quantifying plasmid production in <i>E. coli</i> strains DH5 α and SM10 λ pir	105
Potent promoter cloning bottleneck in lux hints at reporter gene toxicity	105
Figure S5: <i>P. syringae</i> fluoride riboswitch activity in <i>E. coli</i> DH5 α reveals oscillating lux expression and variability across transformant strains.	107
Quick User Manual for pVK-f-lux.....	108
REFERENCES	113

Table S1. List of oligonucleotides used in this study.

Part name ¹	Oligonucleotide name	Sequence 5'-3' ^{2,3}	PCR oligonucleotides and template	Backbone used for cloning this part and other necessary parts ⁴
<i>AarI restriction site deletion in mini-CTX-lux</i>				
<i>luxC</i> AarI deletion	VK45_F	AACGAAGCAGGTGAGGAGCCCACCTGCGTGACTA AAAAAATTTTCATTCATTATTAACGG	PCR (170322 3. <i>luxC</i> tube 3): VK45_F/VK77_R/(<i>mini-CTX-lux</i>)	mini-CTX- <i>lux</i> digested with XcmI/SwaI (with SwaI fragment and P _{S7} + RBS strategy 1) or mini-CTX- <i>lux</i> digested with BsaAI/SwaI (with SwaI fragment and P _{S7} + RBS strategy 2)
	VK77_R	<u>CGACGTGATGAAGGTACACACATCTGCCAAGTGG</u> TTGATTAAATTCCACACCCGCATTTG		
SwaI fragment	VK43_F	<u>GGCAGATGTGTGTACCTTCATCACG</u>	PCR (170323 SwaI) : VK43_F/VK76_R/(<i>mini-CTX-lux</i>)	mini-CTX- <i>lux</i> digested with XcmI/SwaI (with <i>luxC</i> AarI deletion and P _{S7} + RBS strategy 1) or mini-CTX- <i>lux</i> digested with BsaAI/SwaI (<i>luxC</i> AarI
	VK76_R	<u>CACCTTTTAATGCTAACGCATCTCGATATTTAAATG</u> ATGAC		

				deletion and P _{S7} + RBS strategy 2)
<i>Backbone for P1 + 5'-AGGAGT-3' RBS</i>				
VK145	VK145	<u>AAAGTTACAGGCCAGGAACCAC</u> GTACCATGGCCT GCAGGAGTACTGCGAATCCTCCGTCGATCCGGGC GGCCGGCGCGATGAGAACAGGGTTGAAGGACTG CGGAAGGAGCAGCACGATGGACGCTCTGTGACAA TTCGAGCTCTGGAAGCTGGCTAC <u>GTGAACTTCAAC</u> <u>CGTAACCACAAC</u>	gBlock	Ppu21I digested 180222-3-3 (P1.2)
5' non transcribed + 5'UTR regions				
P _{S7} + RBS v1	VK47_F	<u>GAAATTACCCCATTA</u> AATGGATGGCAAATACAGG CTTATGTCTATACAGCTAGCC	PCR (170322 XcmlpMLS7): VK47_F/VK48_R/(pMLS7)	mini-CTX- <i>lux</i> digested with Xcml/Swal (with luxC Aarl deletion and Swal fragment)
	VK48_R	<u>CACGCAGGTGGGCTCCTCACCTGCTTCGTTTCTTC</u> CTTTAACTTTTCAGTTGGAGC		
P _{S7} + RBS v2	VK068_F	<u>CATAATTATGACGAAAGTTACAGGCCAGGAACCAC</u> GTAGTCAGAATCTG	PCR (1700203 Bs/pM): VK68_F/VK48_R/(P _{S7} + RBS strategy 1)	mini-CTX- <i>lux</i> digested with BsaAI/Swal (with luxC Aarl deletion and Swal fragment)

P1 + 5'- AGGAGT-3' RBS v1	VK110_F	<u>AAAGTTACAGGCCAGGAACCAC</u>	PCR (171201 5 D.): VK110_F/VK144_ R/(VK112)	G1C7 digested with BsaAI/AarI
	VK144_R	<u>CCGTTAATAATGAATGAAATTTTTTTAGTCATGCAG</u> GTGAGTACTCCTGCAGGC		
P1 + 5'- AGGAGT-3' RBS v2	VK162_F	<u>CCATGGCCTGCAGGAGTACTCACCTGCATGACTA</u> AAAAAATTTCAATTCATTATTAACGG	VK162_F + VK163_R	08143-6 digested with ScaI/AarI
	VK163_R	<u>CCGTTAATAATGAATGAAATTTTTTTAGTCATGCAG</u> GTGAGTACTCCTGCAGGCCATGG		
P1 + <i>B. thai</i> F RS	VK112	<u>AAAGTTACAGGCCAGGAACCAC</u> GTAGTCAGAATCT GATTTTCTATATATTTGTTATTTACATCGTCATAACA CAAAAATATAAGAAGCAAGTGTTGGTACGACCAGT TCGCAAGATAGTTAAACAGCAACTTAAGTTGAAAT TACCCCATTAATGGATGGCAAATACCAGGTCCT AGGGAGCTCGAATTCACGAACCCAGTTGACATAA GCCTGTTTCGGTTCGTAAACTGTAATGCAAGTAGCG TATGCGCTCACGCAACTGGTCCAGAACCTTGACC GAACGCAGCGGTGGTAACGGCGCAGTGGCGGTT TTCATGGCTTGTTATGACTGTTTTTTGTACAGTCT ATGCCTCGGGCATCCAAGCAGCAAGCGCGTTACG CCGTGGGTCGATGTTTGATGTTATGGAGCAGCAA CGATGTTACGCAGCAGGGCAGTCGCCCTAAAACA AAGTTAGGCAGCCGTTGTGCTGGTGCTTTCTGATA GTTGTTGTGGGGTAGGCAGTCAGAGCTCGATTTG	gBlock	G1C7 digested with BsaAI/AarI

		<p>CTTGTCGCCATAATAGATTCAACAAGAAGGATTCGA CCCAGGCCATGGCCTGCAGGAGTACTCGCCCGG CATCGCGTACAATCCGCGGCTACCGGAGATGGCA TGCCTCCGTACAACCGCCGGCGAGCCGGCTGATG ATGCCTACGCGTTCCTGGGTGCAGGAGGTCGTAG GCCATCCGTGACAAGCGGCGTCCTGCCGCCAG GTTTGATGTCCTGTCTGAATCTGGAAATTCATGAAA CGTCTGCACACTCTCGAGCCCGTCGCGGCGCTCG CGCATCTGTGCCGCTGGCTCGCGCTGTCCGCCGT CGTCGGCGTGCTCGCCGGCTCGGCATCCGCGCT CTTTCTGCGGGCGCTCGATCTCGCCACCGGCACG CGCGTCGCGCATCCGTGGCTGCTGTGGCTGCTGC CCGCAGCGGGCTTCACCTGCATGACTAAAAAATT <u>TCATTCATTATTAACGG</u></p>		
<i>B. thai metK</i> promoter + 5'- AGGAGC-3' RBS ⁵	VK149_F	<u>CGCTCTGTGACAATT</u> CGAGAAAGCCGCTATAATAC GGGCTTCCTCTGGAAGCTGGCTACG	VK149_F +VK150_R	P1.0 (pVK145) digested with Eco53KI
	VK150_R	CGTAGCCAGCTTCCAGAGGAAGCCCGTATTATAG CGGCTTTCTCGAATTGTCACAGAGCG		
<i>metK</i> IGR— from <i>B. thai</i>	VK126_F	<u>AAAGTTACAGGCCAGGAACCACG</u> TACCATGGCCT GCAGGAGTACTGCGAATCCTCCGTC	(PCR : 20171023 PCR tube 2): VK126_F/VK109_ R/(gDNA B.thai E264)	G1C7 digested with BsaAI/AarI
	VK109_R	<u>CTGGCCGTTAATAATGAATGAAATTTTTTTAGTATA</u> ATCGTTTGCCACGTTT		

<i>5' non transcribed region</i>				
P1 integron promoter	VK164_p4.6_F	<u>CCCATTAATGGATGGCAAATACCAGGTCCTAGG</u> <u>GAGCTCGAATTCACGAACCCAGTTG</u>	PCR (180724 P4.6 A): VK164_p4.6_F/VK165_p4.6_R/(VK112 gBlock)	08143-6 digested with Csil/Ncol
	VK165_p4.6_R	<u>AGTACTCCTGCAGGCCATGGCCCGGGTTCGAATCCTTC</u>		
P1 integron promoter—no additional RBS	VK110_F	<u>AAAGTTACAGGCCAGGAACCAC</u>	PCR (180724 P0.2 C): VK110_F/VK175_R/(VK112 gBlock)	G1C7 digested with BsaAI/AarI
	VK175_R	<u>CCGTTAATAATGAATGAAATTTTTTTAGTCATGCAGGTG</u> <u>GAGTACTGCAGCTGCCATGGCCCGGG</u>		
P _{S7} promoter	VK158_p4.2_F	<u>CCCATTAATGGATGGCAAATACCAGGTCCTAGG</u> <u>GAGCTCACAGGCTTATGTCTATACAG</u>	PCR (180724 P4.2 A): VK158_p4.2_F/VK159_p4.2_R/(plasmid pMLS7)	08143-6 digested with Csil/Ncol or P4.1_1 digested with Csil/Ncol
	VK159_p4.2_R	<u>AGTACTCCTGCAGGCCATGGGTTTCTTCCTTTAAC</u> <u>TTTTTCAGTTG</u>		
<i>5' UTR</i>				
<i>P. syri</i> F RS	VK151 (P4.1)	<u>CCATGGCCTGCAGGAGTACTTTTGGACAGACCTAG</u> <u>GCTAAGATCGGCGCATTGGAGATGGCATTCTCC</u> <u>ATTAACAAACCGCTGCGCCCGTAGCAGCTGATGAT</u> <u>TGCCTACAGAAACCTGATCAAACCCAGGCTGTAGG</u> <u>CGTTCGCGCTTAGAATCCCTTCTTTGGTCAGGCC</u> <u>ACTTATTTTTTGTGGCTGGCCAAATGTCTAAATTT</u>	gBlock	08143-6 digested with Scal/AarI

		GACGACCTGAACAACCTCGACTTACTGCCCTATATA GCGAAATGGCTTGCGCTTGCTGGTCTTGTAGCTCT TTTGGCAGGCTCTGCTTCTGCGTTATTCCTGCTTT CTTTGGATCATGCCACCCAGTGGCGAGAAACCCA TCCCTGGGTAATCTGGCTCCTGCCAGTGGCCGGC TTTGTCACCTGCATGACTAAAAAATTCATTCATT <u>ATTAACGG</u>		
<i>metX</i> SAM RS—from <i>B.</i> <i>thai</i>	VK169_F	<u>CCATGGCCTGCAGGAGTACTCGGCGATGCTCGGA</u> AAG	PCR (180724 P2.2 H): VK169_F/VK171_ R/(gDNA <i>B. thai</i> E264)	08143-6 digested with Scal/AarI
	VK171_R	<u>CCGTTAATAATGAATGAAATTTTTTTAGTCATGCCG</u> ATCGATTCCATTCG		
<i>E. coli thiM</i> TPP riboswitch	VK194_F	<u>CCATGGCCTGCAGGAGTACTCCTCTGCGATTTATC</u> ATCG	PCR (16.190105 P5.1 A): VK194_F/VK209_ R/(gDNA <i>E. Coli</i> K12)	08143-6 digested with Scal/AarI
	VK209_R	<u>CCGTTAATAATGAATGAAATTTTTTTAGTGATCAGG</u> TCGACTTGCATAG		
<i>B. thai thiC</i> RS	VK196_F	<u>CCATGGCCTGCAGGAGTACTATCGTGTGCGCTTG</u> C	PCR (17.190105 P5.2 D) : VK196_F/VK210_ R/(gDNA <i>E. Coli</i> K12)	08143-6 digested with Scal/AarI
	VK210_R	<u>CCGTTAATAATGAATGAAATTTTTTTAGTGATGGGG</u> TTGGCGTTCATG		

<i>B. thai</i> miniyKkC RS	VK198_F	<u>CCATGGCCTGCAGGAGT</u> ACTTACTTCCGCTGCCC GGGGACG	PCR (18.190105 P5.3 A): VK198_F/VK211_ R/(gDNA <i>B. thai</i> E264)	08143-6 digested with Scal/AarI
	VK211_R	<u>CCGTTAATAATGAATGAAATTTTTTTAGT</u> GATCCAC AGCCAAGGC		
<i>P. aeruginosa</i> PA14 yybP- ykoY RS	RN_0079_PA14 _rs_fw	<u>CCATGGCCTGCAGGAGT</u> ACTCGCCCCGTCCTAGG GGA	PCR:RN_0079_PA 14_rs_fw/RN_007 9.1_PA14_cds_rev /(gDNA <i>P.</i> <i>aeruginosa</i> PA14)	08143-6 digested with Scal/AarI
	RN_0079.1_PA1 4_cds_rev	<u>CTGGCCGTTAATAATGAATGAAATTTTTTTAGTTTC</u> CGCGAGGGCAACG		
<i>P. syr</i> F RS + <i>B. thai</i> F RS	VK180_F	<u>CCATGGCCTGCAGGAGT</u> CTATATGTAATAATTATA G	PCR: VK180_F/VK183_ R/(VK187)	08143-6 digested with Scal
	VK181_R	<u>GTACGCGATGCCGGGCGAGT</u> ACTCTCTTTAAATA GCTTGCTC		
	VK187	CTGCAGGAGTACTGCGCCGCGCGCGACCGAA AAGGCGTGATACGATGGGGCCTTGCGTGCATTTG ATTTTCAGCTTGCGGACGCGGGGCAACCCGAAACA GCTAAAGCGAAGGCCGCGAGCAGCGCCATGCC GGCCCGAGTCGATAGCTGCTCCGCACACCAAGCC CGCTGATGCCGACGCATGAGCGGGCAAAAAGTTG GTCTGCGCATTGCGCGCGCGCGCACGGCGGCG ATGCGCTTGAAAACGGACACCTGCATGACTAAAAA		

		AATTTCAATTCATTATTAACGGGATCCCATGGCCTG CAGGAGTCTATATGTAATAATTATAGGCGATGGAG TTCGCCATAAACGCTGCTTAGCTAATGACTCCTAC CAGTATCACTACTGGTAGGAGTCTATTTTTTTGAG CAAGCTATTTAAAGAGAGTACTTTTTGGACAGACCT AGCTAAGATC		
<i>B. thai metZ</i> SAM RS	VK178_F	<u>CGAAAGTTACAGGCCAGGAACCACGT</u> ACCATGGC CTGCAGGAGTACTGCGCCGCGCGCGCGAC	PCR (7.190102 A): VK178_F/VK179_ R/(gDNA B.thai E264)	G1C7 digested with BsaAI/AarI
	VK179_R	<u>CCGTTAATAATGAATGAAATTTTTTT</u> AGTCATGCAG GTGTCCGTTTTCAAGCGCATCG		

¹*P. syringae*, *B. thailandensis*, fluoride, and riboswitch are abbreviated to *P. syr.*, *B. thai*, F, and RS respectively.

²Gibson assembly overlaps are underlined.

³Mutations, in reference to the template, are in bold.

⁴Constructs were visualised and designed using Genome Compiler (<http://www.genomecompiler.com/>).

⁵AGGAGC RBS from the *metK* IGR was taken from CP000086 Burkholderia thailandensis E264 chromosome (1204819...204824).

Table S2. Plasmids used in this study.

Plasmid	Description and associated part name	Source	Cloning procedure
mini-CTX- <i>lux</i> (Becher & Schweizer, 2000)	Integration vector with a restriction-site-modified <i>lux</i> gene cluster (<i>luxCDABE</i>) from <i>Xenorhabdus luminescens</i> ; Tc ^R	AF251497	n/a
pMLS7 (Lefebvre & Valvano, 2002)	Expression vector with constitutive promoter of the S7 ribosomal protein gene from <i>Burkholderia</i> sp. strain LB400	Addgene plasmid # 32056	n/a
pG1C7	mini-CTX- <i>lux</i> with 56nt insertion to ribosomal promoter S7 from pMLS7, <i>luxC</i>	This study	GA from numerous parts including the S7 promoter sequence from pMLS7.

	ATG: GTG, and AarI site deletion in luxC.		
pP1.0	Backbone for P1 + 5'-AGGAGT-3' RBS	This study	GA of Ppu21I digested 180222-3-3 (P1.2) and VK145 gblock.
pVK-f2-lux (p08143-6)	P1 + <i>B. thai</i> F RS	This study	GA of BsaAI and AarI digested G1C7 backbone and and VK112 gblock.
pP1.1	<i>B. thai metK</i> promoter + 5'-AGGAGC-3' RBS	This study	GA of Eco53KI digested P1.0 (pVK145) backbone and oligonucleotides VK149_F +VK150_R
p180222-3-3 (pP1.2)	<i>metK</i> IGR— from <i>B. thai</i>	This study	GA of BsaAI and AarI digested G1C7 backbone and PCR product of VK126_F/VK109_R/(gDNA <i>B.thai</i> E264).
pVK-f-lux (pP4.1_1)	P1 + <i>P. syr</i> F RS	This study	GA of Scal/AarI digested 08143-6 backbone and VK151 gblock.
pP2.2-2	P1 + <i>metX</i> SAM RS	This study	GA of Scal/AarI digested 08143-6 backbone and PCR product of VK169_F/VK171_R/ (180724 P2.2 H).
pP5.6	<i>P. aeruginosa</i> PA14 yybP-ykoY RS	This study	GA of Scal/AarI digested 08143-6 backbone and PCR product of RN_0079_PA14_rs_fw/RN_0079.1_PA14/(gDNA <i>P.aeruginosa</i> PA14).
pP4.72	<i>P.syr</i> F RS + <i>B.thai</i> F RS	This study	GA of Scal digested 08143-6 backbone and PCR product of VK180_F/VK183_R/(VK187).

pP6.3	<i>B.thai metZ</i> SAM RS	This study	GA of BsaAI and AarI digested G1C7 backbone and PCR product of VK178_F/VK179_R/(gDNA <i>B.thai</i> E264).
-------	---------------------------	------------	-----------------------------------------------------------------------------------------------------------

Table S3. Strains used in this study.

Strain	Lab catalogue	Description	Source
<i>E. coli</i> DH5 α (Chen <i>et al.</i> , 2018), 2018)	B1A4	Used for transformation of plasmid constructs	NCBI:txid668369
<i>E. coli</i> SM10 λ pir (Simon <i>et al.</i> , 1983)	B1A5	Used for conjugating plasmid constructs with <i>Burkholderia</i> sp. strain E264 and for luminescence assays	BCCM: LMBP 3889
<i>B. thailandensis</i> E264 (Kim <i>et al.</i> , 2005)	B1A7	Used for transformation of plasmid constructs	NCBI:txid271848
<i>E. coli</i> DH5 α /pG1C7	170404GAC7	Obtained from GA transformation test	This study
<i>E. coli</i> DH5 α /P1 + <i>B. thai</i> F RS	B1B3; p08143-6	Reporter strain used for luminescence assays	This study
<i>E. coli</i> SM10 λ pir/P1 + <i>B. thai</i> F RS	B1B5; p08143-6	Reporter strain used for luminescence assays	This study
<i>B. thailandensis</i> E264/ P1 + <i>B. thai</i> F RS	B1B7; p08143-6	Reporter strain used for luminescence assays	This study

<i>E. coli</i> DH5 α / <i>B. thai metK</i> promoter + 5'-AGGAGC-3' RBS	B1G5;p P1.1-1	Obtained from GA transformation test	This study
<i>E. coli</i> SM10 λ pir/ <i>B. thai metK</i> promoter + 5'-AGGAGC-3' RBS	B3B9; pP1.1-1-1	Obtained from GA transformation test	This study
<i>B. thailandensis</i> E264/ <i>B. thai metK</i> promoter + 5'-AGGAGC-3' RBS	B3F9;p P1.1-1-1-2	Obtained from GA transformation test	This study
<i>E. coli</i> DH5 α / <i>metK</i> IGR	B1D3; p180222-3-3/pP1.2	Reporter strain used for luminescence assays	This study
<i>E. coli</i> SM10 λ pir/ <i>metK</i> IGR	B1D6; p180222-3-3/pP1.2	Reporter strain used for luminescence assays	This study
<i>B. thailandensis</i> E264/ <i>metK</i> IGR	B1E8; p180222-3-3/pP1.2	Reporter strain used for luminescence assays	This study
<i>E. coli</i> DH5 α /P1 + <i>P. syr</i> F RS strain 1	B1I6; pP4.1-1	Reporter strain used for luminescence assays	This study
<i>E. coli</i> DH5 α /P1 + <i>P. syr</i> F RS strain 2	B1I7; pP4.1-2	Reporter strain used for luminescence assays	This study
<i>E. coli</i> DH5 α /P1 + <i>P. syr</i> F RS strain 3	B1I8; pP4.1-3	Reporter strain used for luminescence assays	This study

<i>E. coli</i> SM10λpir/P1 + <i>P. syr</i> F RS	B3D5; pP4.1-2-1	Obtained from plasmid transformation test	This study
<i>B. thailandensis</i> E264/P1 + <i>P. syr</i> F RS	B3I2; pP4.1-2-1-1	Reporter strain used for luminescence assays	This study
<i>E. coli</i> DH5α/P1 + <i>metX</i> SAM RS	B3E2; pP2.2-2	Reporter strain used for luminescence assays	This study
<i>E. coli</i> SM10λpir/P1 + <i>metX</i> SAM RS	B4A4; pP2.2-2-1	Reporter strain used for luminescence assays	This study
<i>B. thailandensis</i> E264/P1 + <i>metX</i> SAM RS	B4B1; pP2.2-2-2-1	Reporter strain used for luminescence assays	This study
<i>E. coli</i> DH5α/ <i>P. aeruginosa</i> PA14 yybP-ykoY RS	B4G4; pP14-1	Obtained from GA transformation test	This study
<i>E. coli</i> SM10λpir/ <i>P. aeruginosa</i> PA14 yybP-ykoY RS	B5E1; pP5.6-1_1	Obtained from plasmid transformation test	This study
<i>B. thailandensis</i> E264/ <i>P. aeruginosa</i> PA14 yybP-ykoY RS	B7C1; pP5.6-1_1-1	Obtained from plasmid transformation test	This study
<i>E. coli</i> DH5α/ <i>P. syr</i> F RS + <i>B. thai</i> F RS	B4E4; pP4.72-1	Obtained from GA transformation test	This study

<i>E. coli</i> DH5 α / <i>B. thai metZ</i> SAM RS	B4C7; pP6.3-1	Obtained from GA transformation test	This study
<i>E. coli</i> SM10 λ pir/ <i>B. thai metZ</i> SAM RS	B5C9; pP6.3-1_1	Obtained from plasmid transformation test	This study
<i>B. thailandensis</i> E264/ <i>B. thai metZ</i> SAM RS	B7A9; pP6.3-1_1-1	Obtained from plasmid transformation test	This study

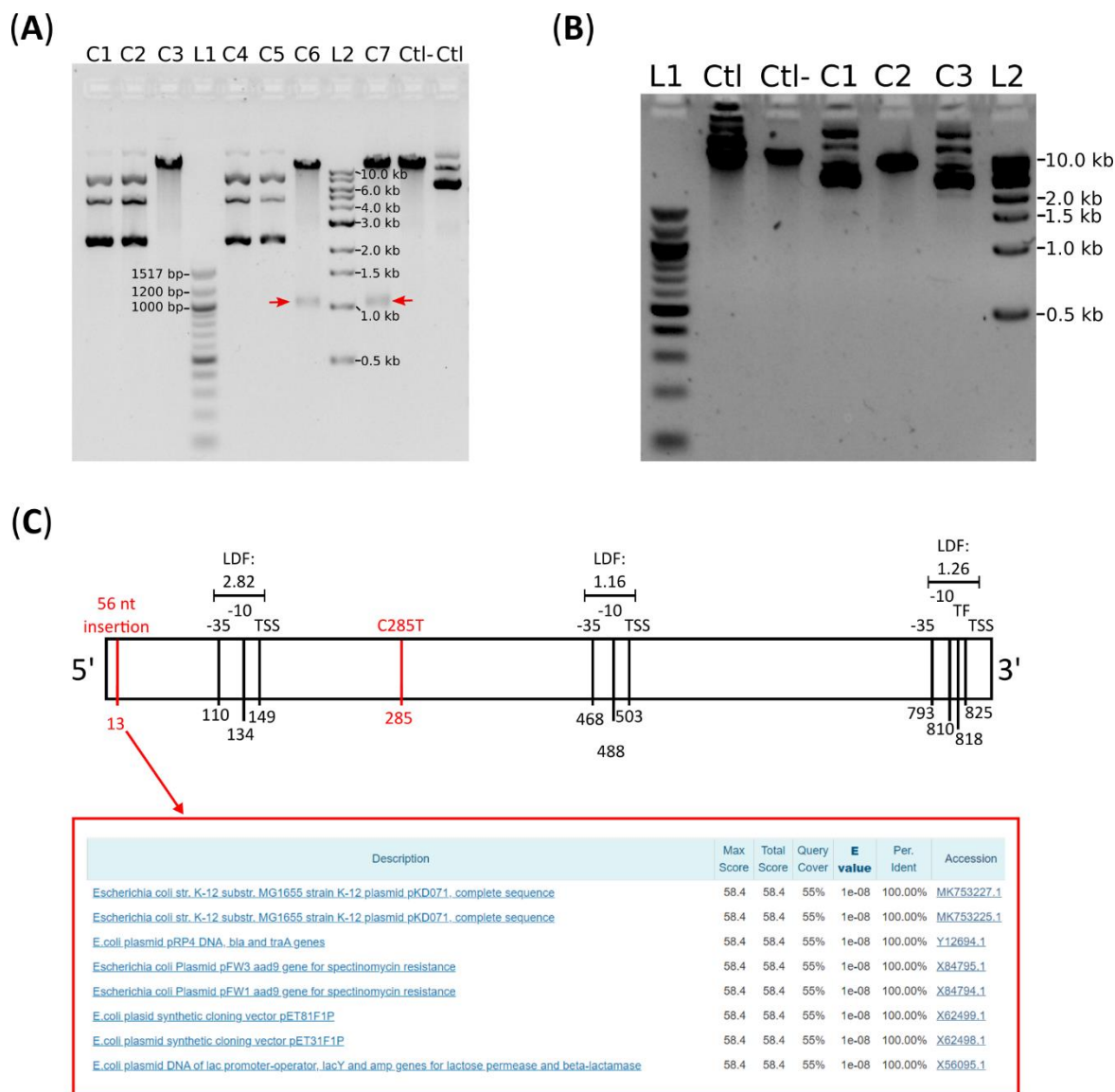


Figure S1: Inactive mutant transformants of the PS7 ribosomal promoter-lux construct may indicate lux overexpression lethality in *E. coli* DH5 α .

(A) Verification of a cloning and transformation experiment by PstI restriction enzyme digestion of plasmids extracted from *E. coli* clones transformed with a GA reaction of a XcmI/SwaI digested mini-CTX-*lux* backbone and three fragments forming a P_{S7}-RBS-*luxC* insert for downstream reporter assays. Ctl- is the original mini-CTX-*lux* plasmid digested with PstI (linearized plasmid, expected size: 12,538 bp) and Ctl is undigested mini-CTX-*lux*. Expected product sizes of a PstI digestion of a successful assembly should yield 12,439 bp and 973 bp products. PCR-amplified assembly fragments (with GA overlaps) were 907 bp (P_{S7} + RBS v1), 1162 bp (*luxC* Aarl deletion), and 139 bp (SwaI fragment) and

linearized backbone was 11311 bp (Table S1). The 7 clones obtained from the experiment are labelled C1-C7. Ladders are designated as L1 (100 bp DNA Ladder from NEB) and L2 (1 kb DNA Ladder from NEB) with indicated size for most relevant bands. Sequencing results were further obtained for C6 and C7 as they revealed expected sized bands (red arrows).

(B) Verification of a cloning and transformation experiment by PstI restriction enzyme digestion of plasmids extracted from *E. coli* clones transformed with a GA reaction of a BsaAI/SwaI digested mini-CTX-*lux* backbone and three fragments forming a P_{S7}-RBS-*luxC* insert for downstream reporter assays. PCR-amplified assembly fragments were 1057 bp (P_{S7} + RBS v2), 1162 bp (*luxC* AarI deletion), and 139 bp (SwaI fragment) and linearized backbone was 11178 bp (Table S1). Expected product sizes of a PstI digestion of a successful assembly should yield a 12439 bp and a 973 bp product. The three clones obtained from the experiment are labelled C1-C3. Ctl- is the original mini-CTX-*lux* plasmid digested with PstI (expected size: 12538 bp) and Ctl is undigested mini-CTX-*lux*. Ladders are designated as L1 (100 bp DNA Ladder from NEB), L2 (1 kb DNA Ladder from NEB). No sequencing results were further obtained as these results indicate misassembly of fragments. (C) Clones 6 and 7 of the GA attempts of the P_{S7}-RBS-*luxCDABE* yielded misassembled constructs for which the former had a point mutation of G334T of the *luxC* sequence (not shown), and the latter had a 56 nt insert in the P_{S7} promoter, as amplified from pMLS7, between C13 and T14, and a point mutation at position 285 (in red). The 850 bp promoter sequence was analysed using BPROM (Richardson & Watson, 2013) and the elements of the three predicted promoters are noted including the -35 and -10 conserved promoter boxes, the Transcription Start Site position (TSS), and position of the oligonucleotides from the crp Transcriptional Factor binding site (TF) (in black). The Linear discriminant function (LDF) of each predicted promoter is also noted. Nucleotide Blast alignment results of the 56 nt insert are also presented. Nucleotide BLAST search hits of the 56 nt insert sequence against *Escherichia coli* K-12 are presented in the red box.

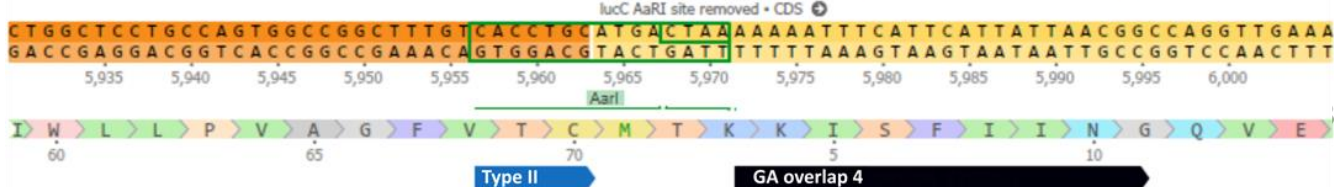
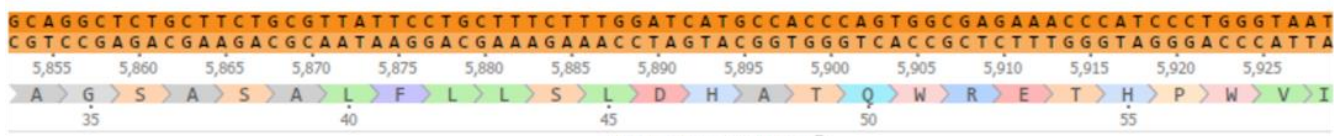
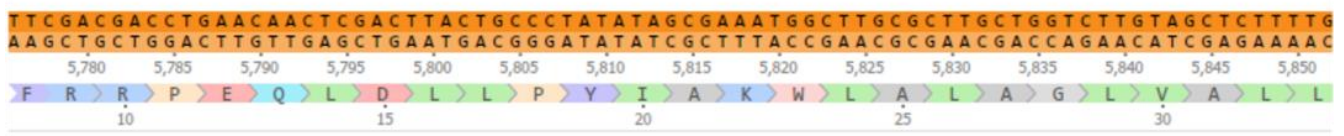
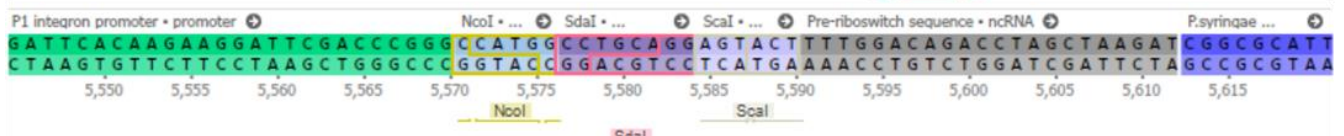
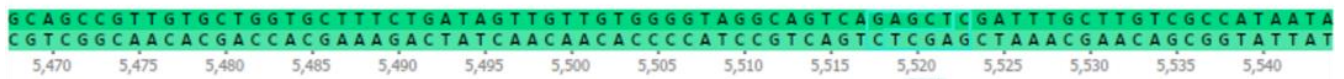
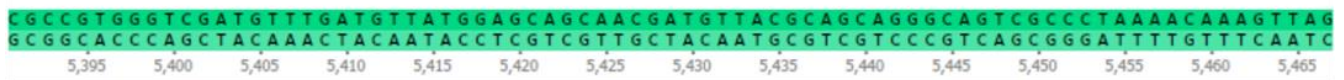
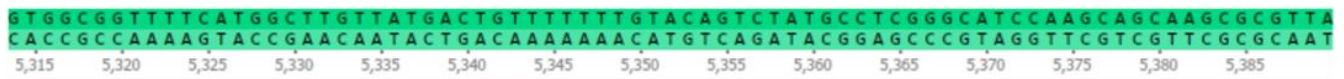
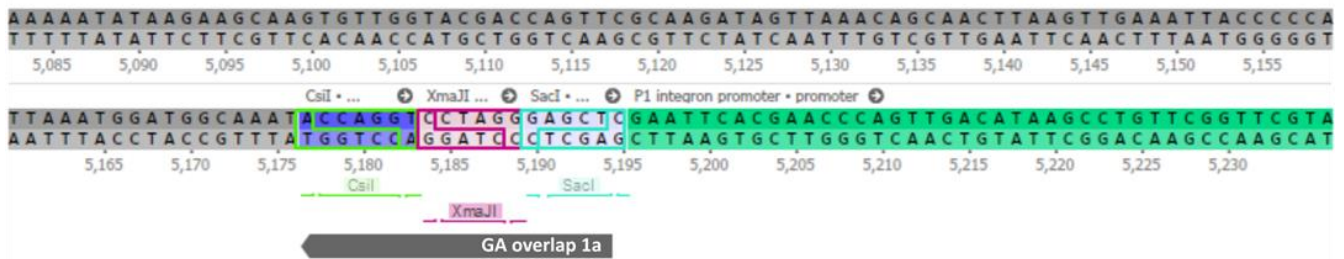


Figure S2: Sequence of the optimized cloning area of pVK-f-lux.

Nucleotide parts are designated by different colours: grey areas are sequences of unknown function, including the pre-riboswitch sequence (5591...5612) and the post-riboswitch sequence (5686...5753) naturally found directly upstream and downstream, respectively, of the fluoride riboswitch in *P.syringae*.; turquoise is the P1 integron promoter (5196...5570); purple is the *P.syringae* fluoride riboswitch (5613...5570); orange is a portion of the gene encoded downstream of the *P. syringae* fluoride riboswitch (5754...5956) with an additional Aarl Type IIS recognition site (5957...5963 with blue arrow below sequence); and yellow is the original luxC sequence (starting at position 5964). Restriction enzyme recognition sites intended for adding or swapping parts are annotated. GA overlaps for parts design are annotated in dark grey for promoter swapping, using either MCS1 (GA overlap 1a) or an upstream BsaAI RE site (GA overlap 1b) and MCS2 (Gibson overlap 2), and in black for 5'UTR swapping, using MCS2 (GA overlap 3) and Aarl RE site (GA overlap 4). Image was put together using Genome Compiler software.

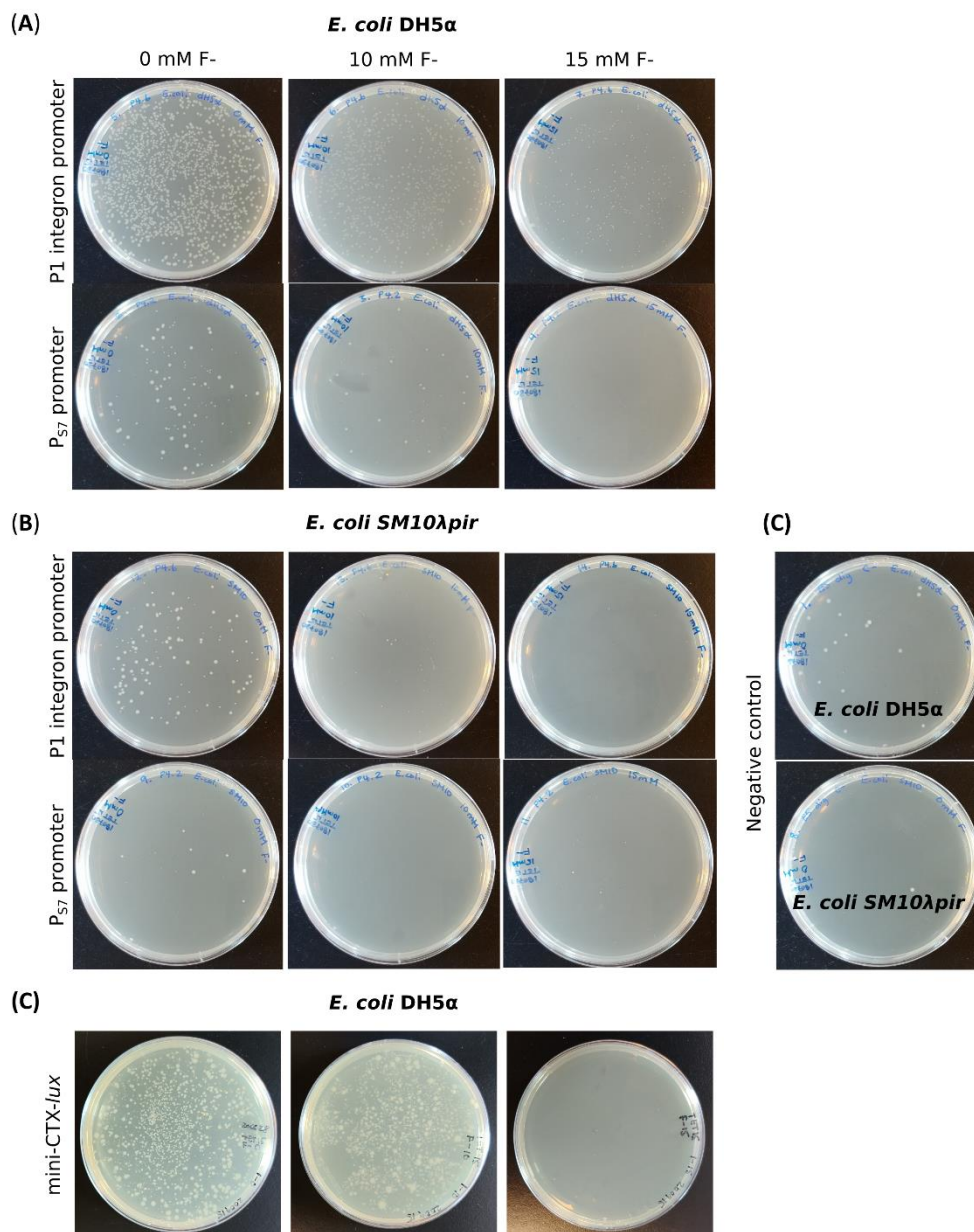


Figure S3: Transformation of heat-shocked *E. coli* is possible in high fluoride concentrations.

Transformation plates of GA reactions of the P1 integron promoter or the P_{S7} promoter upstream of the *B. thailandensis* fluoride riboswitch controlled lux operon (Table S1: P1 integron promoter and P_{S7} promoter, respectively), performed at different concentrations of fluoride into *E. coli* DH5α (A) or *E. coli* SM10λpir (B) after an overnight incubation at 37°C. Selection plates are of LB agar and are supplemented with 15 μg/mL tetracycline. Negative controls, of transformations with digested backbone only are shown in (C) and positive controls of transformations with the original mini-CTX-lux into *E. coli* DH5α performed at different concentrations of fluoride are shown in (D).

Table S4. P_{S7} promoter cloning sequencing results

Transformation fluoride conditions (mM)	Strain	Number of successfully sequenced clones	Number of clones containing 56 nt addition and C: T mutation in the P _{S7} promoter	Number of clones with a gap instead of the promoter sequence	Number of clones with an unrelated sequence instead of P _{S7}
0	<i>E. coli</i> SM10λpir	7	3	3	1
0	<i>E. coli</i> DH5α	9	2	2	5
10	<i>E. coli</i> DH5α	10	1	6	3
15	<i>E. coli</i> DH5α	4	4	0	0

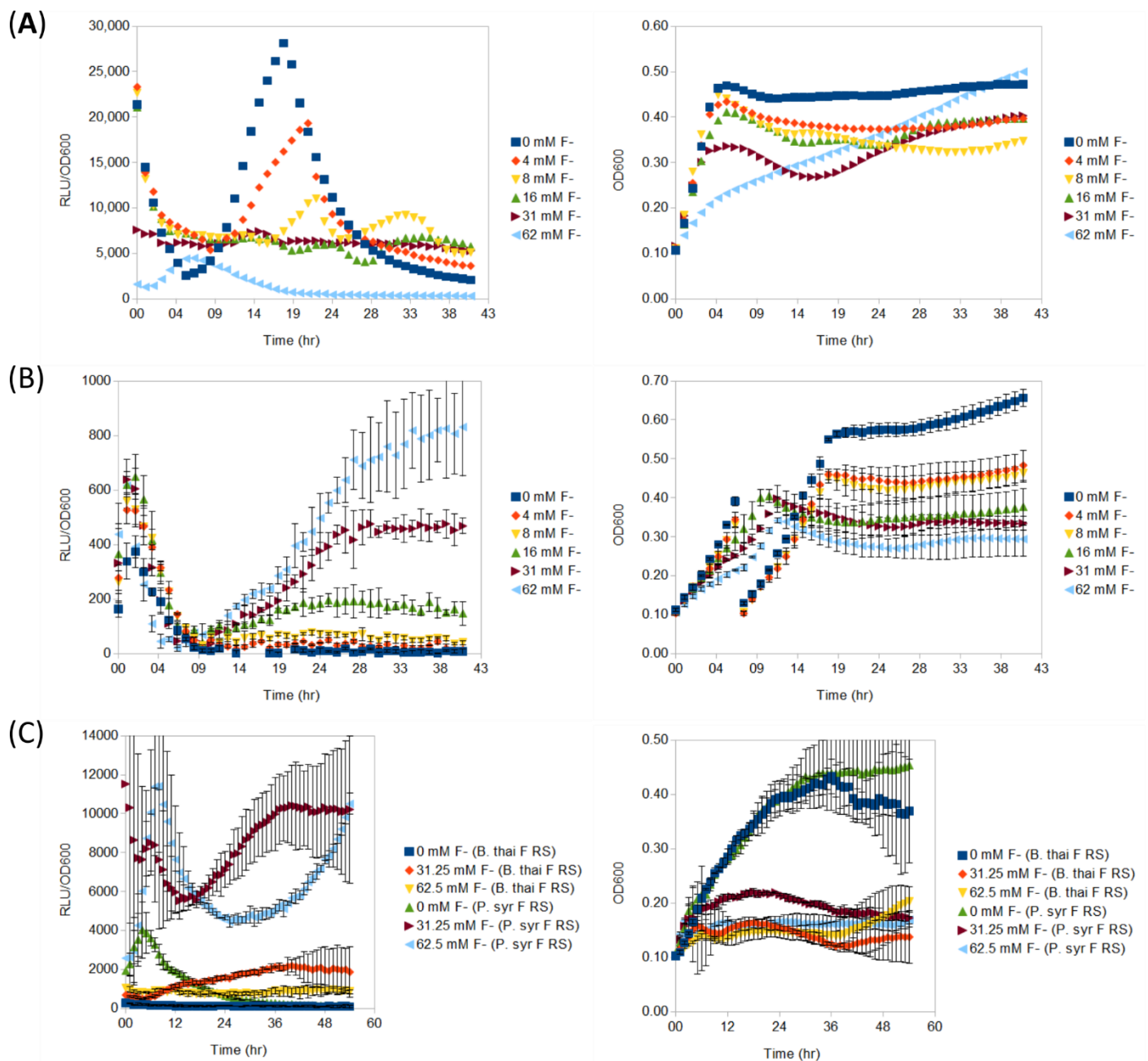


Figure S4: Fluoride riboswitch regulation and fluoride tolerance thresholds of *B. thailandensis* E264 and *E. coli* SM10 λ pir .

(A) *Lux*/OD600 over a 42h time-course regulation by the *B. thailandensis* fluoride riboswitch in *E. coli* SM10 λ pir for a range of fluoride concentrations between 0 and 62.5mM. Pre culture (LB) and expression media (0.5X MM) was supplemented with tetracycline (15 μ g/ml).

(B) *Lux*/OD600 over a 42h time-course regulation by the *B. thailandensis* fluoride riboswitch in *B. thailandensis* E264 for a range of fluoride concentrations between 0 and 62.5mM. Pre-culture (LB) and expression media (0.5X MM) was supplemented with tetracycline (25 μ g/ml), gentamicin (50 μ g/ml), and polymyxin (15 μ g/ml).

(C) *Lux*/OD600 over a 52h time-course regulation by either the *B. thailandensis* fluoride riboswitch (B. thai F RS) or the *P.syringae* fluoride riboswitch (P. syr F RS) in *B. thailandensis* E264 for fluoride concentrations of 0, 31.25 and 62.5mM. Pre-culture (LB) and expression media (0.5X MM) was supplemented with tetracycline (25 µg/ml), gentamicin (50 µg/ml), and polymyxin (15 µg/ml).

Right hand graphs represent growth curves of strain cultures. The data points in (A) represent single values while the data points in (B) and (C) represent the means and standard deviations of triplicate values. Samples (A) and (B) were measured on the same 96-well microplate run. Triplicate values for the *B. thailandensis* fluoride riboswitch construct in *E. coli* SM10λpir exist on other microplate runs and yield similar results (not shown).

Table S5. pVK-f2-*lux* (P1 + *B. thal* F RS-*lux*) plasmid production and *lux* expression specifications in *E. coli* DH5 α and *E. coli* SM10 λ pir.

Luciferase activity values and growth data were obtained from an 41h microplate incubation of 200 μ L cultures containing the P1+ *B. thal* F RS-*lux* construct (pVK-f2-*lux*) in presence of 3.91 mM fluoride, previously shown to be too low to induce riboswitch-mediated regulation. Values represent the means and standard deviations of triplicate results.

Strain	Plasmid production (total ng) ¹	Peak Lum (RLU) ²	OD ₆₀₀ at Peak Lum	Peak Lum/OD ₆₀₀	Average Lum	Average OD ₆₀₀	Average Lum/OD ₆₀₀	Fitness (OD ₆₀₀ Plateau value)
<i>E. coli</i> SM10 λ pir	4910 \pm 175	10,400 \pm 100	0.31 \pm 0.01	33,700 \pm 400	5100 \pm 100	0.33 \pm 0.014	125 \pm 2	0.305 \pm 0.003
<i>E. coli</i> DH5 α	6820 \pm 403	2,120 \pm 90	0.39 \pm 0.03	5,400 \pm 300	1500 \pm 20	0.38 \pm 0.03	32 \pm 2	0.405 \pm 0.018

¹Extracted from a 4.5 mL culture of 0.7 OD₆₀₀, details below in “Quantifying plasmid production in *E. coli* strains DH5 α and SM10 λ pir”

²Both strains were tested on the same microplate during the same run.

Quantifying plasmid production in *E. coli* strains DH5 α and SM10 λ pir

20 mL of liquid Luria Broth supplemented with 15 μ g/mL tetracycline was inoculated in a 50 mL falcon tube with *E. coli* DH5 α /P1 + *B. thai* F RS or *E. coli* SM10 λ pir/P1 + *B. thai* F RS and incubated O/N at 37°C on a shaker. OD₆₀₀ measurements were taken from each inoculation after vortexing and both cultures were diluted to an OD₆₀₀ of 0.7 in a volume of 15 mL. 4.5 mL of each culture was used for plasmid extraction with the Genaid kit and plasmid concentrations were measured using a Nanodrop. Each extraction was repeated three times.

Potent promoter cloning bottleneck in lux hints at reporter gene toxicity

Cloning of the *B. thailandensis* E264 *metK* RBS sequence, 5'-AGGAGC-3', was attempted downstream of the S7 ribosomal protein gene promoter (P_{S7}) from *Burkholderia xenovorans* strain LB400 (recently renamed *Paraburkholderia xenovorans*), using two different three-fragment GA design strategies, one of which had the P_{S7} +RBS fragment adapted for an XcmI/SwaI digested mini-CTX-*lux* backbone and the other which had the P_{S7} +RBS fragment adapted for a BsaAI/SwaI digested mini-CTX-*lux* backbone (Table S1). Vector assembly restriction enzyme digestion results suggested that 7 of the 9 obtained clones contained misassembled constructs (Figure S1A and C). The two vector assemblies containing correctly sized inserts were sequenced and revealed mutations which are suspected to have rendered clones non-luminescent.

A 56 nt sequence consisting of 5'-AACTGGGTTTCGTGCGAGCTCATCGATTTTCGTTCCACTGAGCGTCAGACCCCGTAGA-3', inserted between C13 and T14 of the P_{S7} promoter from pMLS7 (Lefebvre & Valvano, 2002) and a point mutation of C285T of the promoter sequence was present in one of the clones, while the other clone contained a point mutation of G334T of the *luxC* sequence, modifying a glycine to a cysteine. Nucleotide BLAST analysis of the 56 nt insert against *Escherichia coli* K-12 revealed many possible origins (Figure S1B). We suspect this insertion and point mutation somehow blocked transcription from the P_{S7} promoter, or inactivated *luxC* as other clones obtained in experiments described in this paper also contained the same sequence characteristics (Table S4). It is important to note, that the P_{S7} sequence used for our cloning experiments is identical to the one annotated in pMLS7, however the beginning 25 bp of this sequence is absent from the originating host species *Paraburkholderia xenovorans* strain LB400, and thus the reoccurring insertion site in our experiments for the 56 additional nucleotides may not be

present in the host species. Because only clones with mutations/insertions were obtained from the GA, we believe that the correct clones were not viable due to reporter gene toxicity.

Moreover, GA cloning into the same *lux*-containing backbone was attempted for another potent promoter + RBS combination. The P1 integron promoter, previously shown to be an optimal potent constitutive promoter for single copy expression of fluorescent reporter genes in *Burkholderia* spp (Su *et al.*, 2014) was amplified with the RBS sequence of 5'-AGGAGT-3' directly downstream with appropriate GA overhangs which had been tested to work for other constructs (Table S1: P1 + 5'-AGGAGT-3' RBS v1 and v2). However, no clones containing the correct assembly were obtained. An attempt to clone P1 alone, which contains a potential RBS (5'-AGGATT-3') 11 nts upstream of *luxC* in-frame ATG, was also unsuccessful (Table S1: P1 integron promoter—no additional RBS).

The *B. thailandensis* E264 *metK* RBS sequence was successfully cloned into the *lux* backbone with the full original 5' non-transcribed region, containing the native *metK* promoter, as well as the 5'UTR which included a SAM riboswitch (Table S1: *metK* IGR— from *B. thai*). Equally 5'-AGGAGC-3' was successfully cloned into the *lux* backbone in combination with only the native 5' non transcribed region of *metK* (Table S1: *B. thai metK* promoter + 5'-AGGAGC-3' RBS).

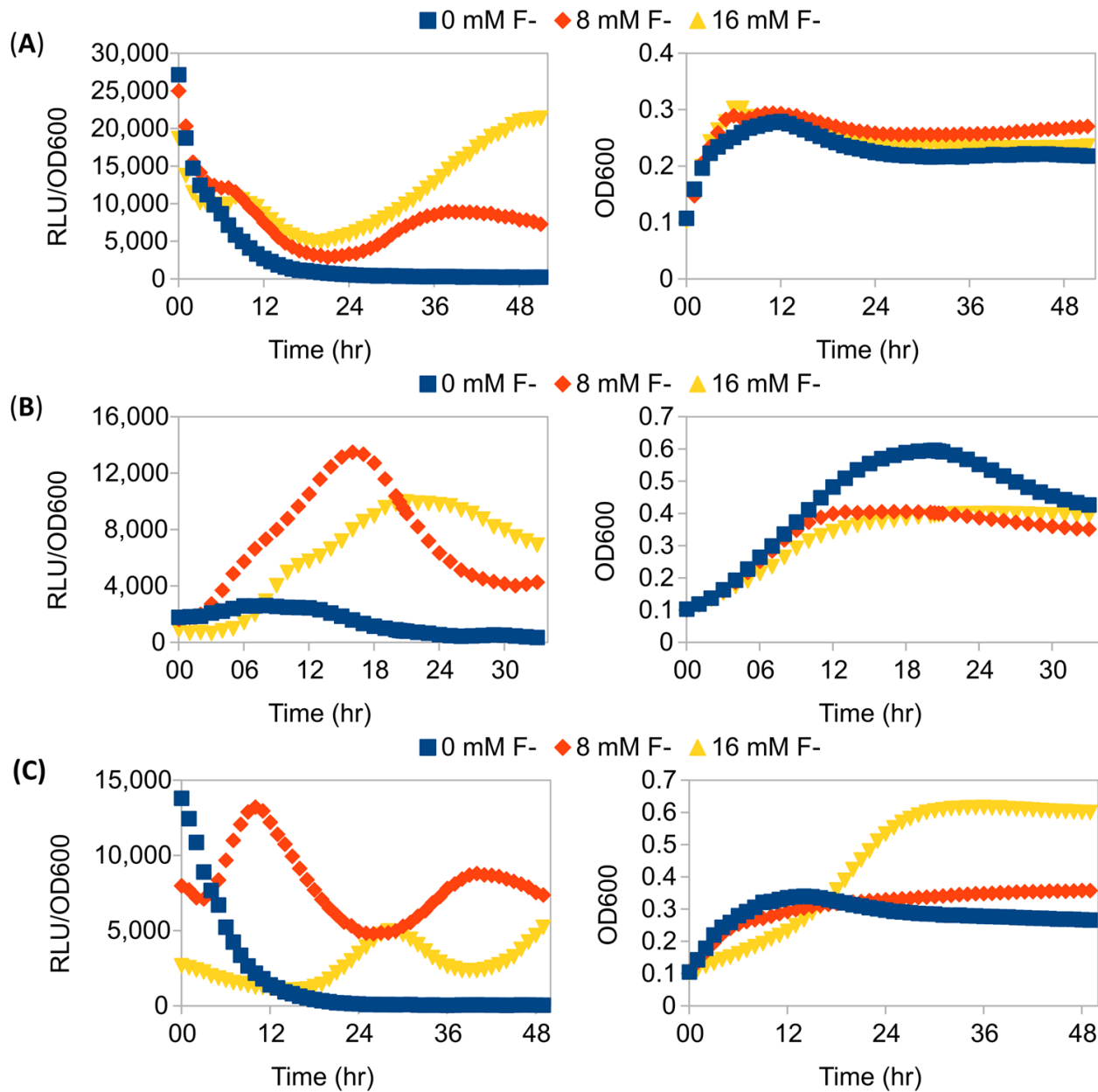


Figure S5: *P. syringae* fluoride riboswitch activity in *E. coli* DH5 α reveals oscillating *lux* expression and variability across transformant strains.

RLU/OD₆₀₀ (left) and growth curves (right) for a 52h, a 34h and 49h time-course expression assay of pVK-f-*lux* constructs in *E. coli* DH5 α /P1 + *P. syr* F RS strain 1 (A), *E. coli* DH5 α /P1 + *P. syr* F RS strain 2 (B), and in *E. coli* DH5 α /P1 + *P. syr* F RS strain 3 (C) respectively. Luciferase expression portrays the variability of regulation across transformants. *Lux* expression is shown in presence of 0, 7.8 and 15.6 mM fluoride. Pre-culture (LB) and expression media (0.5X M9-MM) was supplemented with tetracycline (15 μ g/ml). The data points represent the means and standard deviations of triplicate values measured on the same 96-well microplate run.

Quick User Manual for pVK-f-*lux*

(P1 + *P. syringae* fluoride riboswitch + *lux*)

Updated 2020-07-01

Vesta Korniakova

Intro: Why Use This Plasmid?

This plasmid is useful for down-regulating the *lux* operon through translation regulation via the fluoride riboswitch from *Pseudomonas syringae*. This is useful when you do not know why cloning/transformation or conjugation has been unsuccessful for your construct. A possible reason for your problem may be that the promoter which you are attempting to clone upstream of the *luxCDABE* cassette is too strong and is causing toxicity in your host bacteria, due to overexpression of *lux*. This modified version of the integration vector mini-CTX-*lux*, named pVK-f-*lux* contains an extra layer of regulation (the fluoride riboswitch) between the promoter and *luxCDABE*. In the absence of fluoride, *lux* is down-regulated by at least a factor of 2 in *Escherichia coli* DH5 α . If maximal up-regulation is required, fluoride can be added to the growth media. In *E. coli* DH5 α , fluoride concentrations between 7.8 and 15.6 mM, caused the strongest up-regulation effect while in *Burkholderia thailandensis* E264, the strongest effect was achieved with a 31 mM fluoride concentration. Growth curves at these concentrations, both in *E. coli* DH5 α and in *B. thailandensis* E264 were moderately affected and should be monitored for individual experiments.

Plasmid Contents

This mini-CTX-*lux* modified plasmid contains:

1. P1 integron promoter flanked by 2 added multiple cloning sites (3 restriction sites on each side of promoter; MCS1: Csil, XmaJI and SacI; MCS2: NcoI, SdaI and Scal)
2. *P. syringae* fluoride riboswitch, flanked by 22 nt of upstream sequence and 271 nt of downstream sequence containing the beginning of the gene naturally downstream of the riboswitch (encoding a voltage-gated chloride channel family protein) in fusion with *luxC*, including a type IIS restriction enzyme site (AarI) between the two (without interrupting the fused ORFs).
3. *luxC* has an optimized codon for alanine for *B. thailandensis* E264, at position 7044...7046 nt. This was done to remove the AarI restriction site present in the original mini-CTX-

lux plasmid, in order to relocate the AarI site upstream of *luxC*. Preliminary tests show no side effects of this mutation.

The key features and the pVK-f-*lux* map are presented in Figure 1. Additionally, the nucleic acid and amino acid sequence is shown in Figure S2.

Cloning Strategies

For swapping out the P1 integron promoter for a different promoter

Digest plasmid upstream of P1 integron promoter with either: XmaJI or SacI.

Digest plasmid downstream of P1 integron promoter with: NcoI, SdaI or Scal.

For swapping the riboswitch out for a different RNA regulatory element:

Digest plasmid upstream of the fluoride riboswitch with: Scal, SdaI or NcoI.

Digest plasmid downstream of the fluoride riboswitch with: AarI.

For swapping out the P1 integron promoter and the fluoride riboswitch using one RE:

Digest plasmid upstream of the P1 integron promoter and downstream of the fluoride riboswitch with: CsiI. Note that with this option translational fusion is not possible.

*Important detail: AarI will cut after the 4th nt. of the original *luxC*, so make sure to add back 1 nt to re-establish the correct reading frame of *luxC*. Compare your design and make sure it contains the beginning of the original *luxC* amino acid sequence: M-T-K-K-I-S-F and ends with a stop codon. For translational fusion, the M start codon should be omitted however, to instead use the first codons of the coding sequence relevant to the regulatory element under study. This design was intentional as during Gibson Assembly cloning it allows screening for colonies containing plasmids that have re-ligated without the insert, as the exonuclease will chew back any 5' overhangs thus rendering original sequence re-ligation unlikely. The re-ligated plasmid colonies which do not contain the insert will have lost the required reading frame of *luxC* and will not be luminescent.

Useful tips

1. All of these enzymes use the same digestion buffer for promoter swapping: CsiI, XmaJI, SacI, NcoI and SdaI (10X FastDigest Buffer).

2. All of these enzymes use the same digestion buffer for RNA regulatory element swapping: Scal and AarI (Fisher 10X Buffer AarI, AjiI, Bpu10I, Pasi, Scal).

3. If you want to mix and match different promoters and regulatory RNA elements (5'UTRs) for different modular constructs see below for tried-and-true Gibson Assembly (GA) overlaps.

Note that the underlined sequence is the recognition site of the restriction enzyme (RE) used corresponding either to the left or right cutter and bold sequence is not part of the GA overlap but is a recommended addition in order to re-introduce either the RE recognition site, the *luxC* original sequence, or both in the final construct. The *luxC* start codon is in bold and in italics. Spacing indicates the frame to match for proper *luxC* expression. If primer length is limiting the right GA overlap does not require the suggested re-insertion of the AarI RE site nor the beginning nucleotides of *luxC*. Do make sure to add 1 nt back instead of the latter in order to ensure the correct reading frame of *luxC*. For translational fusions, it is preferable to delete the start codon to avoid the possibility that translation initiation might occur at this site rather than at the intended start codon more relevant for the system being studied.

- a. Use left cutter SacI RE, and right cutter NcoI RE, for swapping promoters. The left and right GA overlaps to use, as they would appear in a 5' to 3' PCR-amplified insert fragment are shown on the left and right side of the insert sequence (N_n):

5'-ACCAGGTCCTAGGGGAGCTC[N_n]CCATGGCCTGCAGGAGTACT-3'.

If needed, a unique BsaAI RE recognition site is located further upstream of the P1 promoter and may also be used as a left cutter. For a BsaAI RE (outside the MCS1)-cut backbone, the left GA overlap to use is shown below:

5'-AAAGTTACAGGCCAGGAACCACGTA[N_n]-3'.

- b. Use left cutter Scal RE and right cutter AarI RE for swapping regulatory RNA elements. The left and right GA overlaps to use, as they would appear in a 5' to 3' PCR-amplified insert fragment are shown on the left and right side of the insert sequence (N_n):

5'- CCATGGCCTGCAGGAGTACT[N_n]CC ACC TGC ATG ACT AAA AAA ATT TCA TTC ATT ATT AAC GG-3'.

The above recommendations will allow to use the same GA insert design for all constructs, using the same GA overlaps instead of re-designing the overlap sequences necessary for each construct if cloning in a non-sequential manner.

4. Easy restriction digestion verification was envisioned with this plasmid. Cutting the P1 integron promoter out with SacI and NcoI will yield a 328 and a 49 bp band because there is a SacI site within the P1 promoter. Cutting out the fluoride riboswitch with Scal and AarI will yield a 380 bp band. Bands (328 bp or 380 bp band), although faint, should be visible as long as 150 ng of digestion reaction is loaded into a small well (15-well comb for small 6x10cm agarose gel tray).

References

- Becher A & Schweizer HP (2000) Integration-proficient *Pseudomonas aeruginosa* vectors for isolation of single-copy chromosomal lacZ and lux gene fusions. *Biotechniques* 29(5):948-950, 952.
- Chen J, Li Y, Zhang K & Wang H (2018) Whole-Genome Sequence of Phage-Resistant Strain *Escherichia coli* DH5alpha. *Genome Announc* 6(10).
- Kim HS, Schell MA, Yu Y, Ulrich RL, Sarria SH, Nierman WC & DeShazer D (2005) Bacterial genome adaptation to niches: divergence of the potential virulence genes in three *Burkholderia* species of different survival strategies. *BMC Genomics* 6:174.
- Lefebvre MD & Valvano MA (2002) Construction and evaluation of plasmid vectors optimized for constitutive and regulated gene expression in *Burkholderia cepacia* complex isolates. *Appl Environ Microbiol* 68(12):5956-5964.
- Richardson EJ & Watson M (2013) The automatic annotation of bacterial genomes. *Brief Bioinform* 14(1):1-12.
- Simon R, Priefer U & Pühler A (1983) A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. *Bio/Technology* 1(9):784-791.
- Su S, Bangar H, Saldanha R, Pemberton A, Aronow B, Dean GE, Lamkin TJ & Hassett DJ (2014) Construction and characterization of stable, constitutively expressed, chromosomal green and red fluorescent transcriptional fusions in the select agents, *Bacillus anthracis*, *Yersinia pestis*, *Burkholderia mallei*, and *Burkholderia pseudomallei*. *Microbiologyopen* 3(5):610-629.

12 ANNEXE II : MATÉRIEL SUPPLÉMENTAIRE RELIÉ AU CHAPITRE 3

Supplementary material for:

Structural diversity of new SAM riboswitch variants extends the scope of the SAM-II/V clan

Balasubramanian Sellamuthu^{1,2#}, Vesta Korniakova^{1#}, Xiaoling Yang¹, Mohammad Reza Naghdi¹, Radia Benmiloud, Michelle Mageau-Dubuc and Jonathan Perreault^{1*}

¹Institut National de la Recherche Scientifique (INRS), Centre Armand-Frappier Santé Biotechnologie, 531 boul. des Prairies, Laval, Québec, H7V 1B7, Canada;

²Centre de Recherche du CHUM, Université de Montréal, 900 rue st-Denis, Montréal, QC, H2X 0A9, Canada

#Authors had equal contribution

*Corresponding author. Email: Jonathan.Perreault@iaf.inrs.ca

1 TABLE OF CONTENTS FOR THIS SUPPLEMENTARY MATERIAL

METHODS ONLINE	117
<i>Discovery of new variant SAM-II/SAM-V riboswitch in β-Proteobacteria by bioinformatics.</i>	
117	
<i>Chemicals and oligonucleotides.....</i>	117
<i>DNA amplification and assembly.....</i>	117
<i>RNA-preparation and In-line analysis.....</i>	118
<i>Bacterial strains and transformation.....</i>	119
<i>lux Reporter assay.....</i>	119
<i>In vivo results analysis</i>	120
<i>Calculations</i>	121
<i>DNA oligonucleotides and plasmids</i>	123
<i>Supplementary Table S1. List of oligonucleotides used in this study.....</i>	123
<i>Supplementary Table S2. Plasmids used in this study.....</i>	125
FIGURE S1: INITIAL DISCOVERY OF A NOVEL RNA STRUCTURE.....	128
RNAROBO DESCRIPTORS.....	129
FIGURE S2: SELECTION OF INSTANCES EVALUATED EXPERIMENTALLY.....	137
FIGURE S3: ALTERNATIVE STEM COMPETING WITH THE APTAMER PSEUDOKNOT.....	140
FIGURE S4: EXPRESSION PROFILE OF <i>METX</i> AND <i>METK</i>	142
FIGURE S5: GENE EXPRESSION MODULATION ACCORDING TO METHIONINE CONCENTRATION.....	144

1.1 Methods online

1.1.1 Discovery of new variant SAM-II/SAM-V riboswitch in β -Proteobacteria by bioinformatics.

While analyzing sequences for another project (Chapalain *et al.*, 2013), we noticed a high level of conservation for sequences upstream of *metK* in species of the genus *Burkholderia*. Because *metK* is well known to be often regulated by SAM riboswitches, we then looked at this more carefully using approaches similar to those described in (Naghdi *et al.*, 2017). In short, we looked for conserved structures in relevant intergenic regions with CMfinder (Yao *et al.*, 2006), later extending the model to many other Betaproteobacteria by using Infernal. The structure initially found with this approach is pictured in Figure S1.

1.1.2 Chemicals and oligonucleotides.

After bioinformatic analysis of the putative riboswitch found in the 5'UTR of *metK* in β -proteobacteria, oligonucleotides were custom synthesized for PCR amplification of the aptamer domain from *Burkholderia thailandensis* E264 genome as a first model organism. Many different constructs were designed and the construct yielding a clear modulating pattern obtained from in-line probing experiments, an aptamer portion of 100 nt, was amplified using the chosen primer regions (Supplementary Table S1). Similarly, other SAM aptamer candidates tested were custom synthesized and all forward primers were flanked by a T7 promoter sequence. The T7 promoter sequence alone used as a forward primer in combination with a short reverse primer were used in assembly-PCR reactions along with longer primers (Supplementary Table S1) to obtain high concentration of desired amplicon in the subsequent amplification cycles.

1.1.3 DNA amplification and assembly.

All short oligonucleotides and gene fragments used in this study can be found in Supplementary Table 1. Oligonucleotides and gene fragments were obtained from Integrated DNA Technologies (25 nmole DNA Oligo and 500ng of gBlocks® Gene Fragments).

Aptamer constructs for the *in vitro* inline probing assays were amplified using the reagent kit and corresponding protocol from New England Biolabs (Taq DNA Polymerase with Standard Taq Buffer (M0273)). PCR oligonucleotide primers and template for each construct are found in Supplementary Table 1.

DNA amplification of inserts for cloning was carried out using the reagent kit and corresponding protocol from New England Biolabs (Q5® High-Fidelity DNA Polymerase (M0491S)).

Backbones for cloning were linearized using appropriate restriction enzymes and cloning was carried out using Gibson Assembly as per instructions provided with the reagent kit from New England Biolabs (NEBuilder® HiFi DNA Assembly Master Mix, #E2621S). DNA inserts and backbones for reporter plasmid constructs are described in Supplementary Table. 2.

1.1.4 RNA-preparation and In-line analysis.

The 5' UTR of *metK* genes were amplified using specific primers from genomic DNA of *B. thailandensis* E264 and *Burkholderia* CCGE1001. Bacterial strains (*B. thailandensis* E264, *Burkholderia* CCGE1001 and *Pseudomonas*) were obtained from our institute. Other bacterial strains 5' UTR (riboswitch aptamer domain) region was obtained by performing assembly-PCR using the custom synthesized primers (Supplementary Table S1). All forward were flanked by the T7 promoter sequence. Amplified products were transcribed and gel purified for use in subsequent studies. RNAs obtained from beta-proteobacteria were dephosphorylated using antarctic phosphatase and labeled with (γ -³²P) ATP using T4 polynucleotide kinase (New England Biolabs). Labeled RNA (5' ³²P) was gel purified on 6% PAGE. For the inline probing assay, radiolabeled RNA (\approx 30000 CPM) was added to the master mix and incubated in 20 μ L reactions with different concentrations of SAM (800 μ M – 0.1 nM) at room temperature for 40h. The 2X inline buffer consists of 40 mM MgCl₂, 200 mM KCl and 100 mM Tris-HCl pH 8.3. Alkaline lysis (10 μ L reaction) of labeled RNA was performed using 1 μ L of Na₂CO₃ (1M) and incubated at 90°C for 2 min. Similarly, the T1 reaction (10 μ L volume) was performed using 1 μ L of formamide (>99.5 %) and sodium citrate (0.25 M pH 5) with labeled RNA incubated at 56°C for 2 min and then RNase T1 enzyme (1/100 dilution) was added and incubated for 5 more minutes at 56°C. The cleaved RNA products of the inline reaction were fractionated on 10% denaturing PAGE (8M urea). Gels were exposed onto a phosphorImager membrane overnight and scanned using Typhoon FLA 7000 (GE Healthcare Life Sciences). The band intensities were quantified using Image QuaNT software. The band intensities were normalized to a range of 0 (lowest ligand concentration) to 1 (highest ligand concentration). Obtained values were plotted in log scale of SAM molar concentrations.

1.1.5 Bacterial strains and transformation.

E. coli strain DH5 α (F⁻, ϕ 80d*lacZ* Δ M15, Δ (*lacZYA-argF*)U169,*deoR*, *recA1*, *endA1*, *hsdR17*(rk⁻, mk⁺),*phoA*, *supE44*, λ ⁻, *thi-1*, *gyrA96*, *relA1* from our Lab's collection) was used for transforming Gibson-Assembly-cloned plasmid constructs for sequencing verification. *E. coli* strain SM10 λ *pir* (*thi-1*, *thr-1*, *leuB6*, *tonA21*, *lacY1*, *supE44*, *recA1::RP4-2-Tc::Mu* [λ *pir*]; Km^R, obtained from Éric Déziel's lab) was used for conjugation of sequence-confirmed vector constructs with *B. thailandensis* strain E264(Brett *et al.*, 1998) (Wild Type, obtained from Éric Déziel's lab). During transformation procedures all bacteria were grown at 37°C on a rotary agitator in liquid Luria Broth (Alpha Biosciences L12-112) or statically on petri dishes containing Luria agar (Alpha Biosciences L12-111).

When required, selection and/or growing conditions in liquid or solid media for both *E. coli* strains were supplemented with 15ug/mL tetracycline. For transformed *B. thailandensis*, selection, growing liquid and solid media conditions were 25 ug/ml tetracycline, 50 ug/ml gentamycine, and 15 ug/ml polymyxin.

Construction of *B. thailandensis* reporter strains was accomplished by first transforming *E. coli* DH5 α with the cloning mix as per the NEBuilder® HiFi DNA Assembly Chemical Transformation Protocol (E2621). Transformed colonies obtained from the selection plates were cultured overnight and culture samples were pre-screened for luminescence signal using a microplate reader (Cytation 3; BioTek Instruments, Inc., Winooski, VT).

Luminescent cultures were then used for plasmid extraction with a kit from Geneaid (Presto™ Mini Plasmid Kit (PDH300)). Reporter constructs were confirmed by Sanger sequencing at Genome Quebec CES (Montreal, Canada). The confirmed reporter constructs were used for transformation into competent *E. coli* SM10 cells by the same protocol as previously used. Conjugation transformation of *B. thailandensis* was achieved by mating reporter plasmid carrying *E. coli* SM10 cells with *B. thailandensis* E264. transformants were selected on antibiotic supplemented plates as previously mentioned.

1.1.6 *lux* Reporter assay.

To assess gene expression regulation by the putative *metk*, *metX* and *metZ* SAM riboswitches, *B. thailandensis* reporter strains containing chromosomally integrated riboswitch

translational fusions with the bacterial operon *luxCDABE* (Becher & Schweizer, 2000) as well as their corresponding mutants were cultured overnight in either liquid Luria Broth (LB) (Alpha Biosciences L12-112), M9 Minimal Medium (MM; as per Cold Spring Harbor Protocol 2010. 2010: pdb.rec12295), or M9 Minimal Medium with no sulfur (MM-S), where 1 M MgSO₄ was substituted with 1 M MgCl₂. When required, overnight culture media was supplemented with antibiotics. In some cases cycloleucine, an inhibitor of methionine adenosyltransferase, the enzyme which catalyzes the generation of SAM (Lombardini & Talalay, 1970) was added for a final concentration of 100 mM. 1.5 mL of overnight cultures were pelleted by centrifugation at 14.8 rpm for 3 minutes and washed twice with 750 µL of either MM, or MM-S. The washed pellet was resuspended in either MM or MM-S. The assay was conducted on a 96 well microplate from Greiner Bio-One (Microplate, 96 well, PS F-bottom (chimney well), white, med. binding Ref: 655095). A total culture volume of 200 µL containing 0.5X dilution of either MM or MM-S supplemented with 50 µg/ml gentamycin and 15 µg/ml polymixin. When required culture was additionally supplemented with 25 µg/ml tetracycline and/or 100 mM cycloleucine. Initial optical density of the culture at 600nm (OD₆₀₀) was approximately 0.06 and cultures were initially induced with concentrations of L-methionine (Met) between 0 and 10 mM, purchased from Sigma Aldrich (M2893). Luminescence and OD₆₀₀ readings were automatically recorded using a multimode microplate reader (Cytation 3; BioTek Instruments, Inc., Winooski, VT) at 20-minute intervals for assay runtimes between 50 and 60 hours.

1.1.7 *In vivo* results analysis

Initial readings of different assay runs were set to starting points where the OD₆₀₀ reading was between 0.15 and 0.3. Time-point corresponding background blank values were subtracted from each time point of each sample reading. Luciferase expression was measured in relative light units (RLU) and normalised to the OD₆₀₀. To compare samples the luminescence activity signal strength was compared between reads originating from the same microplate. When required to compare samples across different microplate runs, the samples were normalized to the *metK* mut. construct in absence of supplementary Met. For analysis of riboswitch time-course dynamics from the same microplate, the signals from *metK* and *metX* constructs were normalized to 1. For easier visualisation of the time-course luciferase expression curve, hourly time points were selected from the raw set of 20-minute interval time-points.

1.1.8 Calculations

Fold Induction (FI) of total luciferase activity:

$$FI of lum_{total} = \frac{\sum_n^N f^{+Met}(n)}{\sum_n^N f^{-Met}(n)}$$

where:

$$f^{+Met}(n) = \frac{lum^{+Met}(n)}{OD^{+Met}(n)},$$

$$f^{-Met}(n) = \frac{lum^{-Met}(n)}{OD^{-Met}(n)}$$

$lum^{+Met}(n)$ is the luminescence reading for the *B. thailandensis* E264 transformant culture containing the chromosomally integrated SAM RS-*lux* construct of interest in presence of supplementary Met (+Met) for the time point n.

$lum^{-Met}(n)$ is the luminescence reading for the *B. thailandensis* E264 transformant culture containing the chromosomally integrated SAM RS-*lux* construct of interest in absence of supplementary Met (-Met) for the time point n.

$OD^{+Met}(n)$ is the OD_{600} reading for the *B. thailandensis* E264 transformant culture containing the chromosomally integrated SAM RS-*lux* construct of interest in +Met condition for the time point n.

$OD^{-Met}(n)$ is the OD_{600} reading for the *B. thailandensis* E264 transformant culture containing the chromosomally integrated SAM RS-*lux* construct of interest in -Met condition for the time point n.

Fold Induction (FI) of maximum luciferase activity:

$$FI of lum_{max} = \frac{\max(f^{+Met}(n))}{\max(f^{-Met}(n))}$$

For each replicate, the maximum luminescence/ OD_{600} value for +Met was divided by the average maximum luminescence/ OD_{600} value for -Met.

Basal level of total luciferase activity inhibition:

$$Basal FI of lum_{total} = \frac{\sum_n^N f^{metK;-Met}(n)}{\sum_n^N f^{metKmut;-Met}(n)}$$

where:

$$f^{metK;-Met}(n) = \frac{lum^{metK;-Met}(n)}{OD^{metK;-Met}(n)},$$

$$f^{metKmut;-Met}(n) = \frac{lum^{metKmut;-Met}(n)}{OD^{metKmut;-Met}(n)}$$

$lum^{metK;-Met}(n)$ is the luminescence reading for the *B. thailandensis* E264 transformant culture containing the chromosomally integrated *metK* SAM RS-*lux* construct of interest in -Met condition for the time point n.

$OD^{metK;-Met}(n)$ is the OD_{600} reading for the *B. thailandensis* E264 transformant culture containing the chromosomally integrated *metK* SAM RS-*lux* construct of interest in -Met condition for the time point n.

$lum^{metK mut;-Met}(n)$ is the luminescence reading for the *B. thailandensis* E264 transformant culture containing the chromosomally integrated *metK* mut SAM RS-*lux* construct of interest in -Met condition for the time point n.

$OD^{metK mut;-Met}(n)$ is the OD_{600} reading for the *B. thailandensis* E264 transformant culture containing the chromosomally integrated *metK* mut SAM RS-*lux* construct of interest in -Met condition for the time point n.

Basal level of maximum luciferase activity inhibition:

$$Basal FI of lum_{max} = \frac{\max(f^{metK;-Met}(n))}{\max(f^{metKmut;-Met}(n))}$$

1.1.9 DNA oligonucleotides and plasmids

1.1.10 Supplementary Table S1. List of oligonucleotides used in this study.

Construct name	Oligonucleotide name	Sequence 5'-3' ^{a,b}	PCR oligonucleotides and template
<i>Parts for in vitro inline probing assays</i>			
<i>metK</i> wt	BS018-F	taatacgcactcactataggctataatacgggcttc	BS018-F/BS019_R/(gDNA <i>B.thai</i> E264)
	BS019_R	ggcgacgcttagcggtat	
<i>metK</i> mut.	JP216_M4-F(c-alt)	taatacgcactcactataggctataatacgggcttcgccgagAaactgacaTcttgccggggcgaagccggcg	PCR 150916 M4: JP216_M4-F(c-alt)/BS019_R/(gDNA <i>B.thai</i> E264)
<i>metX</i> wt	BS0052 F-T7	taatacgcactcactatagggtgctcggaaagcgcgccgattgctc	BS0052 F-T7/BS0053 R/(gDNA <i>B.thai</i> E264)
	BS0053 R	ggcaggctgacgacctttag	
<i>metZ</i> wt	BS0054 F-T7	taatacgcactcactataggatcacgatggggccttgctcgatt	BS0054 F-T7 / BS0055 R/(gDNA <i>B.thai</i> E264)
	BS0055 R	gctcgcggccttgcttagctgtt	
<i>Parts for constructing backbone for cloning</i>			
P1.0 backbone part for cloning all mutant constructs	VK145 gblock	aaagttacaggccaggaaccacgtaccatggcctgcaggagtactgcaatcctccgtcgatccgggcccgcggcgcgatgagaacagggttgaaggactgcggaaggagcagcacgatggacgctctgtgacaattcga gctctggaagctggctacgtgaactcaaccgtaaccacaac	n/a
<i>Parts for in vivo bioluminescence assays</i>			
<i>metK</i> wt	VK126_F	aaagttacaggccaggaaccacgtaccatggcctgcaggagtactgcaatcctccgtc	VK126_F/VK109_R/(gDNA <i>B.thai</i> E264)
	VK109_R	ctggccgtaataatgaatgaaatttttagtataatcgttgccacgttc	
<i>metK</i> mut.	VK146	acgctctgtgacaattcgagaaagccgctataatcggcct	VK146/VK147/(PCR 150916 M4)
	VK147	acgctctgtgacaattcgagaaagccgctataatcggcct	

<i>metX</i> wt	VK176	<u>ttacaggccaggaaccacgtaccatggcctgcagg</u> <u>agtactcgcggaagcggggcggaac</u>	PCR 5.181110P6.1 C: VK176_P6.1_F/VK177_R/(gDNA <i>B.thai</i> E264)
	VK177_R	<u>gtaataatgaatgaaatTTTTtagtcatgcaggtgac</u> <u>gccgatcgattcattcgctc</u>	
	VK207	<u>gtaataatgaatgaaatTTTTtagtgatgcaggtg</u>	PCR 5.190102 C: VK176_P6.1_F/VK207_P6.1_R/(5.181110P6.1 C)
<i>metX</i> mut.	VK188 gBlock	<u>agtactcgcggaagcggggcggaacaagcatccg</u> <u>gcacgcatcgacgcgcgggcgacgccggacgcgc</u> <u>ggctccccctccgtcgccgcggcgcggaatTTTc</u> <u>cgttatactgcccggctatcgccgcttcgcggcg</u> <u>atgctcggaagcgcgccgatatgctgactcgattgc</u> <u>gggcgcggaaccggccgggctcgttcccggcg</u> <u>ggttactacaaatcgggctaaagaggtcgtcagcct</u> <u>gccccgcttcccgcgcacgccgacaaccatttag</u> <u>ccgccagtcataaggcgggaacgaatggaatcgat</u> <u>cgcgctcacctg</u>	PCR 6.181031P6.2 C: VK176_F/VK177_R/(VK188 gBlock) VK176_P6.1_F/ VK207_R/(PCR 6.181031P6.2 C)
P1 + <i>metZ</i>	VK172_F	<u>ccatggcctgcaggagtagtggcgtgatacgatggg</u>	VK172_F/VK174_R/(gDNA <i>B.thai</i> E264)
	VK174_R	<u>ccgtaataatgaatgaaatTTTTtagtcatcgtttcaa</u> <u>gcgcatcg</u>	

^a T7 promoter sequence is underlined

^b Mutations are capitalized

1.1.11 **Supplementary Table S2. Plasmids used in this study.**

Plasmid	Description	Source	Cloning procedure
mini-CTX- <i>lux</i> (Becher & Schweizer, 2000)	Integration vector with a restriction-site-modified <i>lux</i> gene cluster (<i>luxCDABE</i>) from <i>Xenorhabdus luminescens</i> ; Tc ^R	Obtained from E.Deziel's lab, INRS-IAF.	n/a
pMLS7(Lefebvre & Valvano, 2002)	Expression vector with constitutive promoter of the S7 ribosomal protein gene from <i>Burkholderia</i> sp. strain LB400	Obtained from E.Deziel's lab, INRS-IAF but is a gift from Miguel Valvano (Addgene plasmid # 32056 ; http://n2t.net/addgene:32056 ; RRID:Addgene_32056)	n/a
pG1C7	mini-CTX- <i>lux</i> with 56nt insertion to ribosomal promoter S7 from pMLS7, <i>luxC</i> ATG:GTG, and AarI site deletion in <i>luxC</i> .	Article du chapitre 3	n/a
pVK-f2- <i>lux</i>	P1 + <i>B. thai</i> F RS	Article du chapitre 3	n/a
pP2.2-2	P1 + <i>metX</i> SAM RS	Article du chapitre 3	n/a
pP2.3-1-1-1	P1 + <i>metZ</i> SAM RS	This study	Gibson Assembly of ScaI/AarI digested pVK-f2- <i>lux</i> backbone and PCR product of VK172_F/VK174_R/(gDNA <i>B.thai</i> E264)
180222-3-3 (P1.2)	<i>metK</i> IGR	This study	Gibson Assembly of BsaAI/AarI digested pG1C7 backbone and PCR product of VK126_F/VK109_R/(gDNA <i>B.thai</i> E264).
180611-2-2 (P1.0 or pVK145)	Backbone for parts 1.1 to 1.8. plasmid	This study	Gibson Assembly of Ppu21I digested 180222-3-3 backbone and and VK145 gblock.
P1.6_1	<i>metK</i> mut.	This study	Gibson Assembly of Eco53KI digested 180611-2-2 backbone and and PCR product of VK146/VK147/(150916 M4).

190106 P6.1-3	metX IGR	This study	Gibson Assembly of BsaAI/AarI digested pG1C7 backbone and PCR product of VK176_P6.1_F/VK207_P6.1_R/ (5.181110P6.1 C).
190106 P6.2-1	metX IGR with SAM RS BP mutation	This study	Gibson Assembly of BsaAI/AarI digested pG1C7 backbone and PCR product of VK176_P6.1_F/VK207_R/(PCR 6.181031P6.2 C).

1.1.12 Supplementary Table S3. Strains used in this study.

Strain	Lab catalogue	Description	Source
<i>E. coli</i> DH5 α (Chen et al., 2018)	B1A4	Cloning and transformation strain	NCBI:txid668369
<i>E. coli</i> SM10 λ pir (Simon et al., 1983)	B1A5	Conjugal donor strain for conjugating <i>Burkholderia</i> sp. strain E264	BCCM: LMBP 3889
<i>B. thailandensis</i> E264 (Kim et al., 2005)	B1A7	Transformation strain for chromosomal integration of <i>lux</i> plasmids	NCBI:txid271848
<i>B. thailandensis</i> E264/ <i>metK</i> IGR	B1E8; p180222-3/pP1.2	Reporter strain	Article du chapitre 3
<i>B. thailandensis</i> E264/ <i>metK</i> mut.	B3G9; pP1.6-1-1-2	Negative control reporter strain	This study
<i>B. thailandensis</i> E264/ <i>metX</i> IGR	B7A5; pP6.1-1_1-1	Reporter strain	This study

<i>B. thailandensis</i> E264/ <i>metX</i> mut	B7A7; pP6.2-1_1-1	Negative control reporter strain	This study
<i>B. thailandensis</i> E264/ <i>metZ</i> IGR	B7A9; pP6.3-1_1-1	Reporter strain	This study
<i>B. thailandensis</i> E264/P1 + <i>metX</i> SAM RS	B4B1; pP2.2-2-2-1	Reporter strain	<i>Article du chapitre 3</i>
<i>B. thailandensis</i> E264/P1 + <i>metZ</i> SAM RS	B3H8; pP2.3-1-1-1	Reporter strain	This study

1.3 RNArobo descriptors

Several RNArobo descriptors were used to find more instances of SAM riboswitches homologous to the SAM-II clan. The main feature that makes RNArobo more advantageous over Infernal in this particular case is its ability to search for pseudoknots. On the other hand, RNArobo requires more precise descriptions when compared to Infernal and for this several descriptors need to be formulated to accommodate potential mispairs, bulges, extra bases in certain positions, etc. If too much flexibility is allowed in a single descriptor results will be overpopulated with too many hits.

```
#SAM-v06-1_bulge+1-2ss.des
#s4 preferentially starts with R (G/A) (and can be up to 30)
#s7 = up to 200

#motif map for consensus SAM riboswitch
s1 h1 s2 s3 s4 h2 s5 h3 h4 s6 h1' s7 h4' s8 h3' s9 h2' s10

s1 0      *
h1 0:0:1 *****NNNN:RNNN*****
s2 0      N***
s3 0      UU
s4 0      N[30]
h2 0:0    ***NN:NN***
s5 0      NNV
h3 0:0    UU:AA
h4 0:0    GN****:****NN
s6 0      [2]
s7 0      NNN[200]
s8 0      U
```

s9 0 ANN

s10 0 *



```

#SAM-v06-2_misp+1-2ss).des

#s4 preferentially starts with R (G/A) (and can be up to 30)
#s7 = up to 200

#motif map for consensus SAM riboswitch
s1 h1 s2 s3 s4 h2 s5 h3 h4 s6 h1' s7 h4' s8 h3' s9 h2' s10

s1 0      *
h1 0:1:0 *****NNNN:RNNN*****
s2 0      N***
s3 0      UU
s4 0      N[30]
h2 0:0    ***NN:NN***
s5 0      NNV
h3 0:0    UU:AA
h4 0:0    GN****:****NN
s6 0      [2]
s7 0      NNN[200]
s8 0      U
s9 0      ANN
s10 0     *

```

```

#SAM-v06-3_misp+0ss).des

#s4 preferentially starts with R (G/A) (and can be up to 30)
#s7 = up to 200

```

#motif map for consensus SAM riboswitch

s1 h1 s2 s3 s4 h2 s5 h3 h4 h1' s6 h4' s7 h3' s8 h2' s9

s1 0 *

h1 0:1:0 *****N***R***N*****

s2 0 N***

s3 0 UU

s4 0 N[30]

h2 0:0 ***NN:NN***

s5 0 NNV

h3 0:0 UU:AA

h4 0:0 GN****:****NN

s6 0 NNN[200]

s7 0 U

s8 0 ANN

s9 0 *

```

# SAM-v06-4_bulge+0ss).des

#s4 preferentially starts with R (G/A) (and can be up to 30)

#s7 = up to 200

#motif map for consensus SAM riboswitch

s1 h1 s2 s3 s4 h2 s5 h3 h4 h1' s6 h4' s7 h3' s8 h2' s9

s1 0      *
h1 0:0:1 *****NNNN:RNNN*****
s2 0      N***
s3 0      UU
s4 0      N[30]
h2 0:0    ***NN:NN***
s5 0      NNV
h3 0:0    UU:AA
h4 0:0    GN****:****NN
s6 0      NNN[200]
s7 0      U
s8 0      ANN
s9 0      *

-----

#SAM-v07-1_bulge+1-2ss.des

#s4 preferentially starts with R (G/A) (and can be up to 30)

#s7 = up to 200

```

#motif map for consensus SAM riboswitch

s1 h1 s2 s3 s4 h2 s5 h3 h4 s6 h1' s7 h4' s8 h3' s9 h2' s10

s1 0 NNN

h1 0:0:1 *****NNNN:RNNN*****

s2 0 N***

s3 0 UU

s4 0 N[30]

h2 0:0 ***NN:NN***

s5 0 NNV

h3 0:0 UU:AA

h4 0:0 GN****:****NN

s6 0 [2]

s7 0 NNN[200]

s8 0 U

s9 0 ANN

s10 0 NNN

```

#SAM-v07-4_bulge+0ss.des

#s4 preferentially starts with R (G/A) (and can be up to 30)

#s7 = up to 200

#motif map for consensus SAM riboswitch

s1 h1 s2 s3 s4 h2 s5 h3 h4 h1' s6 h4' s7 h3' s8 h2' s9

s1 0      NNN
h1 0:0:1 *****NNNN:RNNN*****
s2 0      N***
s3 0      UU
s4 0      N[30]
h2 0:0    ***NN:NN***
s5 0      NNV
h3 0:0    UU:AA
h4 0:0    GN****:****NN
s6 0      NNN[200]
s7 0      U
s8 0      ANN
s9 0      NNN

```

```

# SAM-v08-2_1-2ss_stringent.des

#s4 preferentially starts with R (G/A) (and can be up to 30)

#s7 = up to 120

```

#motif map for consensus SAM riboswitch

s1 h1 s2 s3 s4 h2 s5 h3 h4 s6 h1' s7 h4' s8 h3' s9 h2' s10

s1 0 ***

h1 0:0 *****NNNNN:RNNNN*****

s2 0 N***

s3 0 UU

s4 0 N[20]

h2 0:0 ***NN:NN***

s5 0 NNV

h3 0:0 UU:AA

h4 0:0 GNNN**:**NNNN

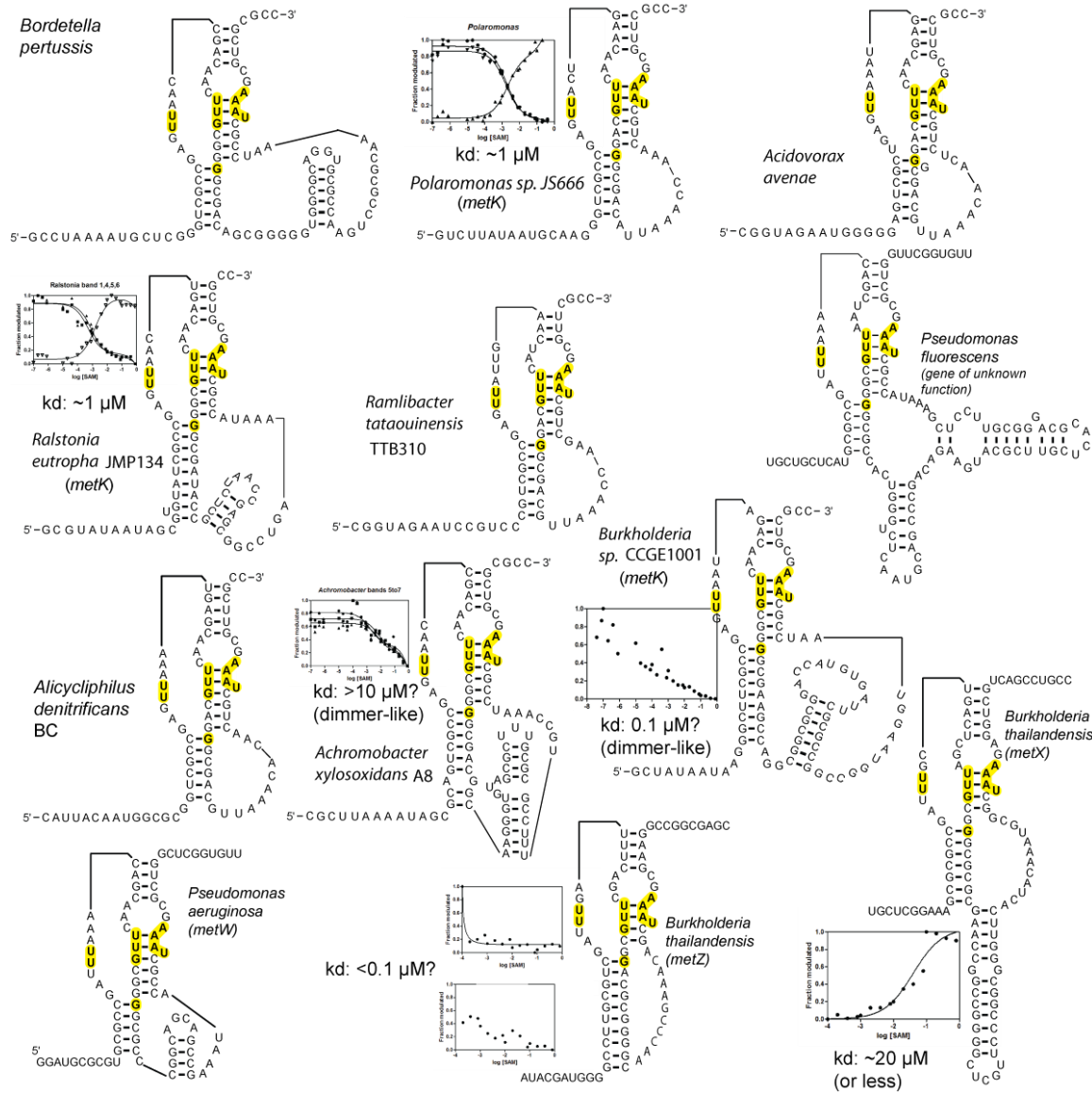
s6 0 [1]

s7 0 NNN[120]

s8 0 U

s9 0 ANN

s10 0 ***



1.4 Figure S2: Selection of instances evaluated experimentally.

Hundreds of novel instances of riboswitches corresponding to the SAM-II clan were discovered. The initial searches permitted such discoveries mostly within the class of Betaproteobacteria, but many more were found later on. The structure drawings were hand-made to highlight similarities with the SAM-II clan, with the binding pocket highlighted in yellow. The stem often present within loop 3 is only a prediction, with some cases fitting a particular consensus, but not always. A few example curves are shown next to the corresponding structure, the X axis is in log [μM] and Y axis is the normalized intensity of chosen modulating bands.

Some of the modulation results obtained from in-line probing of these riboswitches show an unusual profile where the concentration at which modulation occurs corresponds to a range of 3, 4 or even 7 logs, instead of the typical 2 logs, thus potentially acting more like a dimmer, rather than a simple switch.

Regarding this, we noted a trend where the instances had what we called an “accessory stem”. In several cases, the K_d could only be roughly estimated because of the atypical modulation pattern where plateaus could not be determined. More details below.

The p4 stem possessing the SAM aptamer from *Achromobacter xylosoxidans* strain A8 and *Bordetella pertussis*, revealed a linear and widespread modulating pattern (data not shown).

Polaromonas sp. JS666 revealed a typical sigmoidal curve but with an apparent K_d value in the range of 10 μ M of SAM in all three selected bands (Figure 3 and map the gel). Furthermore, SAM aptamer secondary structure from *Polaromonas* sp. p4 stem is absent, and its structure is similar to SAM-II riboswitch reported (Corbino *et al.*, 2005; Meyer *et al.*, 2009; Poiata *et al.*, 2009). Thus, the absence of this stem might allow the aptamer to modulate in an ON/OFF switch pattern.

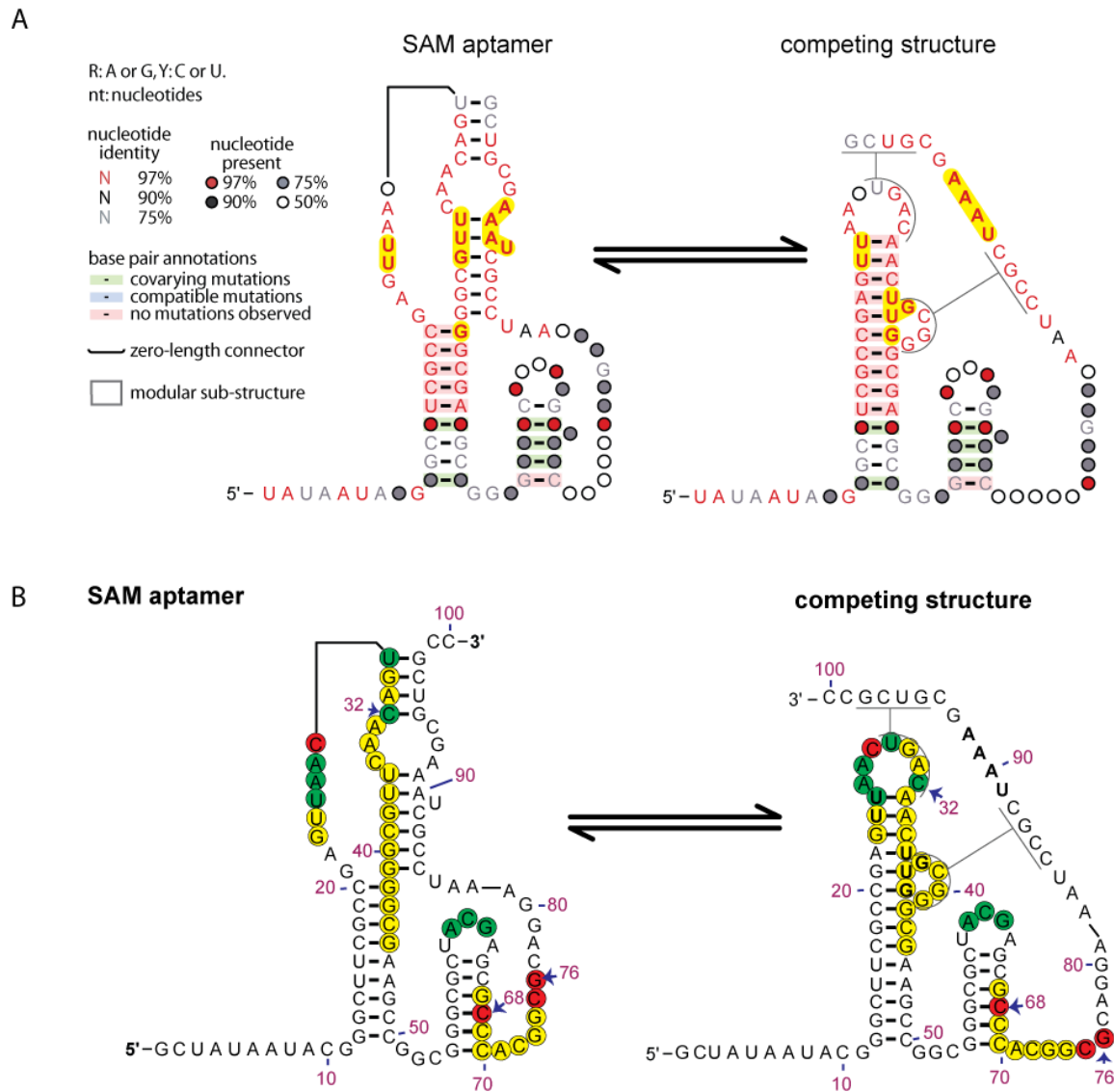
Ralstonia eutropha JMP134 SAM aptamer (with p4 stem) revealed sigmoidal modulating pattern over 3 logs. *Ramlibacter* species, lacking the p4 stem, demonstrated a relatively typical modulating pattern (over 3 logs). Whereas *Acidovorax* and *Alicyclophilus* strains did not reveal a clear modulating pattern, suggesting that the construct used was not ideal to observe conformational changes and/or allow proper folding of the aptamer and thus binding of SAM. Each experiment was conducted in triplicates.

We surveyed the literature for previously published riboswitches and the majority of them have a linear range of approximately two logs for the sigmoidal curve (with ranges from 0.5 and 1 for the cooperative riboswitches, up to 2.5).

	0.5	1	1.5	~1.8	2	2.5	3	3.5	4
Number of logs									
Linear range	1	3	6	2	20	4			
Between plateaus			3		2	7	14	6	2

Note: These values were qualitatively determined from graphs with log scales illustrating different types of experiments (generally in-line probing assays, but also SHAPE, fluorescence quenching and transcription termination assays) with varying degrees of quality for data points. Note that it was impossible to determine these ranges without raw data when published graphs did not use a log scale, which limited our survey to 33 publications.

Our work does not solve why these *Burkholderia* species (and others) require a wide concentration of SAM for their biological functions, sometimes with modulation still observable beyond 10-100 μM , i.e. much higher than the typical $\sim 0.1\text{-}1 \mu\text{M}$ Kd usually seen for SAM riboswitches.

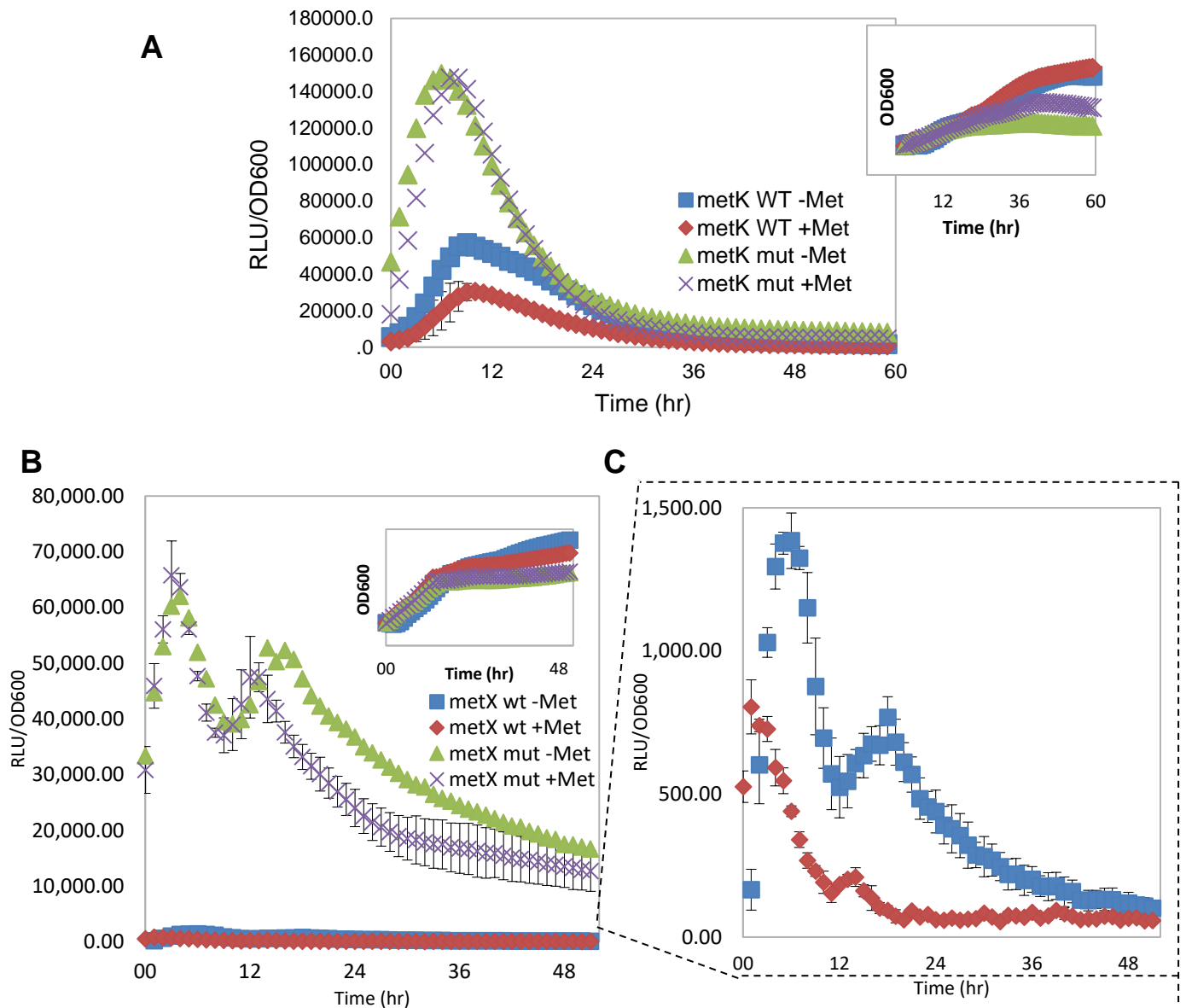


1.5 Figure S3: Alternative stem competing with the aptamer pseudoknot.

A) Secondary structure found upstream of *metK* in Burkholderiales species as in Figure 1 with a putative alternative structure. The alternative structure was initially inferred from a search for novel RNA structures (see Figure S1). B) Same structure as in (A), but for the riboswitch instance in *B. thailandensis* with annotations of modulation observed from in-line probing. Green, yellow and red correspond to greater, same or lesser band intensity, respectively.

We hypothesize that some of the differences observed between the riboswitches upstream of *metK* and *metX* (both with regards to in-line probing and reporter assays), could be in part

explained by the presence of competing structures that would affect the folding dynamics of this riboswitch.

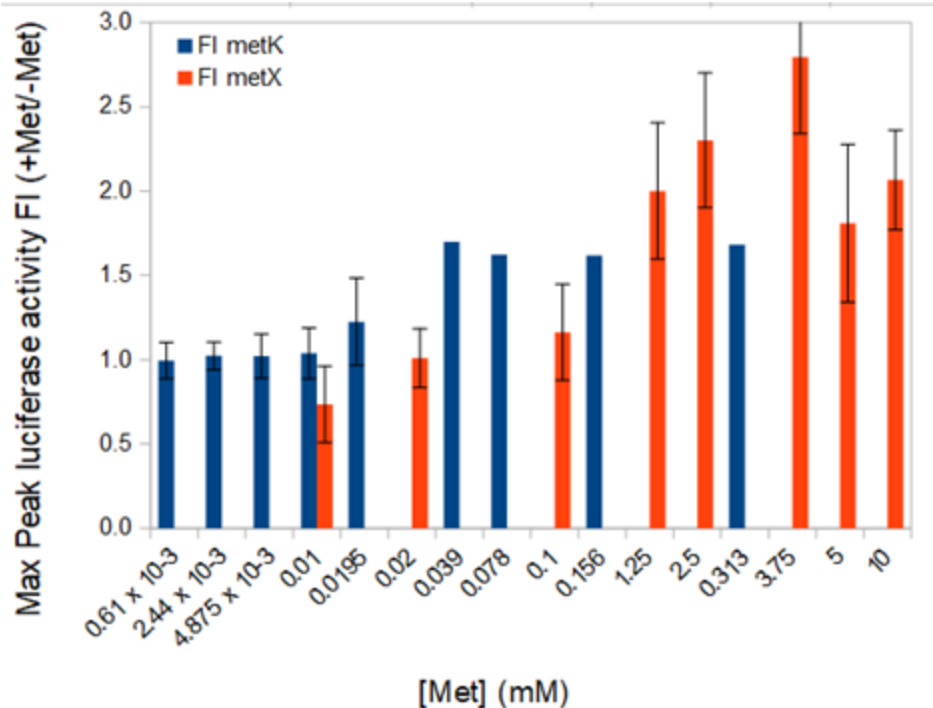


1.6 Figure S4: Expression profile of *metX* and *metK*.

A) Reporter assays of the riboswitch found upstream of *metK* in *B. thailandensis*. Inset shows the growth curves. B) Reporter assays of the riboswitch found upstream of *metX* in *B. thailandensis*. Inset shows the growth curves. C) Zoom on the curves of the WT-*metX* constructs.

It is worth noting that, for the *metK* reporters, the curves all have roughly the same shape, albeit with expression levels varying widely, suggesting that the threshold for switching is reached over the entire time course. On the other hand, for *metX* reporters, not only do the inactive mutants

have a ~100-fold higher expression, but the shape of the curves of this inactive riboswitch is similar to that of the WT-reporter in absence of methionine, while supplementation of methionine leads to a very different curve. Initial expression (first 3 hours) of the WT-reporter is similar whether methionine is added or not, presumably to allow time for expression and production of SAM, but then drops compared to medium without added methionine. This suggests that, contrary to the *metK* riboswitch, the *metX* riboswitch might act more as a dimmer since variations of SAM concentration over time apparently affect levels of repression.



1.7 Figure S5: Gene expression modulation according to methionine concentration.

Data (error bars correspond to deviation calculated from triplicates) is represented as fold induction as calculated from the maximum peak of expression ($[\text{RLU}/\text{OD}]_{-\text{Met}}/[\text{RLU}/\text{OD}]_{+\text{Met}}$). Modulation of gene expression seems to occur from lower concentrations of methionine for *metK* (although some data points at 0.039 mM and 0.078 mM might be missing to ascertain this). However, more interestingly, gene modulation appears to rapidly plateau in these conditions, while it continues to increase up to 3.75 mM for *metX* in these conditions. This is in line with the relatively high k_d value observed for the *metX* riboswitch and could be seen as further support for the “dimmer” vs “switch” hypothesis.

1.8 References

- Becher A & Schweizer HP (2000) Integration-proficient *Pseudomonas aeruginosa* vectors for isolation of single-copy chromosomal *lacZ* and *lux* gene fusions. *Biotechniques* 29(5):948-950, 952.
- Brett PJ, DeShazer D & Woods DE (1998) *Burkholderia thailandensis* sp. nov., a *Burkholderia pseudomallei*-like species. *Int J Syst Bacteriol* 48 Pt 1:317-320.
- Chapalain A, Vial L, Laprade N, Dekimpe V, Perreault J & Deziel E (2013) Identification of quorum sensing-controlled genes in *Burkholderia ambifaria*. *Microbiologyopen* 2(2):226-242.

- Chen J, Li Y, Zhang K & Wang H (2018) Whole-Genome Sequence of Phage-Resistant Strain *Escherichia coli* DH5alpha. *Genome Announc* 6(10).
- Corbino KA, Barrick JE, Lim J, Welz R, Tucker BJ, Puskarz I, Mandal M, Rudnick ND & Breaker RR (2005) Evidence for a second class of S-adenosylmethionine riboswitches and other regulatory RNA motifs in alpha-proteobacteria. *Genome Biol* 6(8):R70.
- Kim HS, Schell MA, Yu Y, Ulrich RL, Sarria SH, Nierman WC & DeShazer D (2005) Bacterial genome adaptation to niches: divergence of the potential virulence genes in three Burkholderia species of different survival strategies. *BMC Genomics* 6:174.
- Lefebvre MD & Valvano MA (2002) Construction and evaluation of plasmid vectors optimized for constitutive and regulated gene expression in Burkholderia cepacia complex isolates. *Appl Environ Microbiol* 68(12):5956-5964.
- Lombardini JB & Talalay P (1970) Formation, functions and regulatory importance of S-adenosyl-L-methionine. *Adv Enzyme Regul* 9:349-384.
- Meyer MM, Ames TD, Smith DP, Weinberg Z, Schwalbach MS, Giovannoni SJ & Breaker RR (2009) Identification of candidate structured RNAs in the marine organism 'Candidatus Pelagibacter ubique'. *BMC Genomics* 10:268.
- Naghdi MR, Smail K, Wang JX, Wade F, Breaker RR & Perreault J (2017) Search for 5'-leader regulatory RNA structures based on gene annotation aided by the RiboGap database. *Methods* 117:3-13.
- Poiata E, Meyer MM, Ames TD & Breaker RR (2009) A variant riboswitch aptamer class for S-adenosylmethionine common in marine bacteria. *RNA* 15(11):2046-2056.
- Simon R, Prierer U & Pühler A (1983) A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. *Bio/Technology* 1(9):784-791.
- Yao Z, Weinberg Z & Ruzzo WL (2006) CMfinder--a covariance model based RNA motif finding algorithm. *Bioinformatics* 22(4):445-452.