

Institut Armand-Frappier**Recherche d'ARN non codants régulateurs senseurs de cations**

Par :
Mohammad Reza Naghdi

Thèse présentée pour l'obtention du grade
de Philosophiae Doctor (Ph.D.)
en sciences de biologie

Jury d'évaluation

Président du jury et examineur interne	Dr. Richard Villemur INRS-IAF
Examineur externe	Dre. Gertraud Burger Département de Biochimie Faculté de Médecine Université de Montréal
Examineur externe	Dr. Martin Pelchat Département de Biochimie, microbiologie et d'immunologie Faculté de médecine Université d'Ottawa
Examineur interne	Dr. Richard Villemur INRS-IAF
Directeur de recherche	Dr. Jonathan Perreault INRS-IAF

REMERCIEMENTS

Tout d'abord j'aimerais remercier mon directeur de recherche Pr. Jonathan Perreault qui m'a accueilli dans son laboratoire et m'a aidé au cours de ces années assez longues, pour apprendre petit à petit le processus de la réflexion, de la recherche, de la rédaction et enfin l'autonomie. J'aimerais également remercier mes membres de jury qui ont accepté d'évaluer ma thèse et leurs présences pour la soutenance. Je voudrais également remercier tous les membres anciens et actuels, du laboratoire Pr. Jonathan Perreault pour m'accompagner durant ces années avec leurs aides, amitiés et collaboration.

Je tiens à témoigner toute ma reconnaissance aux membres du laboratoire Pr. Eric Déziel surtout Marie-Christine Groleau qui m'aide pour une partie de mon projet. Je remercie mes chers amis, qui m'ont accompagné pendant ces années pour leurs présences, leurs aides, leurs sympathies et leurs encouragements. Un grand merci et reconnaissance à Dre. Elham Dianati, l'ex-étudiante du Pr. Isabelle Plante pour sa présence et ses aides, dès le premier examen jusqu'au dernier.

Je remercie ma famille mon oncle Dr. Khazalpour et sa conjointe Gisela, mes cousines Fati et Farimah Ghavamsadri pour leurs présences et l'encouragement durant ces dernières années. J'adresse ma reconnaissance également à mon oncle Pr. Paul M. Naghdi qui m'a inspiré pour la science. Il était professeur à l'université Berkeley et il était connu pour son équation, « Green-Naghdi equation ».

Enfin, J'aimerais dédier cette thèse à mes chers parents, Habib et Sabrieh qui m'ont souhaité toujours le bonheur et la réussite. Je remercie infiniment ma mère qui n'a jamais cessé de me soutenir pendant ces années et que la réalisation de ce travail lui rende heureuse.

RÉSUMÉ

Tous les êtres vivants et notamment les bactéries subissent des changements environnementaux tels que des variations de température, de nutriments, de pH, des métaux, etc. Ces changements peuvent être brusques et rapides et par conséquent les micro-organismes doivent s'adapter aux changements très rapidement. Les micro-organismes y compris des bactéries ont développé plusieurs mécanismes de régulation qui leur permettent de survivre dans l'environnement.

Les métaux ont une importance particulière pour tous les organismes vivants y compris les bactéries ; d'une part ils sont nécessaires pour les activités vitales de tous les êtres vivants, et d'autre part ils sont toxiques voire létaux dans les concentrations élevées pour les bactéries. Par exemple le fer est nécessaire pour des voies métaboliques telle le cycle de Krebs. Le zinc est nécessaire pour la structure de certaines protéines. Le magnésium est le cofacteur de plusieurs enzymes telles que des polymérases. Il est aussi nécessaire pour le repliement des acides nucléiques. Le manganèse est impliqué dans les enzymes peroxidase qui ont le rôle de la protection des bactéries.

L'homéostasie des métaux chez les bactéries dépend de transporteurs. Ces systèmes de transport de métaux sont sous le contrôle de régulations étroites par les bactéries. Les systèmes de régulation permettent de contrôler efficacement autant dans le cas de manque ou de l'excès des métaux.

Chez les bactéries les métaux sont régulés par deux couches (systèmes) connues de régulation. Un de ces systèmes est basé sur des protéines régulatrices des métaux et d'autres sont constitués d'ARN non codant. Il existe au moins deux groupes distincts d'ARN non codant qui sont impliqués dans la régulation des métaux : des petits ARN comme RhyB et des *riboswitchs* tels que les *riboswitchs* magnésium, manganèse et nickel/cobalt.

Le but de cette thèse est de découvrir des nouveaux ARN non codant, notamment des *riboswitchs* spécifiques aux métaux en utilisant des approches bioinformatiques et de biologies moléculaires.

Mot clés : ARN non codant (ARNnc), Bioinformatique, Riboswitch, Régulation de métaux.

ABSTRACT

All the living species including bacteria undergo several environmental changes such as changes in temperature, nutrition, pH, metals, etc. These changes normally occur suddenly and fast and therefore the microorganisms should be capable to adjust rapidly. The microorganisms as well as bacteria have developed several regulatory systems, which allow them to survive in such a fast changing environment.

The metals are especially important for bacteria. On one hand, metals are essential for vital activity of all the domains of life, including the bacteria, on the other hand they are toxic or lethal in high concentration. For instance, iron is necessary for metabolism pathways like Krebs cycle. Zinc is necessary in the structure of some proteins. Magnesium is a cofactor of several enzymes such as polymerases. It is needed also for nucleic acids folding. Manganese is involved in peroxidase enzyme, which is protecting the bacteria cells. In bacteria the metals are transported from exterior to interior or vice versa. Metal transporting systems are under tight control of regulation systems in bacteria. Bacteria have regulation systems which control the metal quantity necessary for the cell and which control efficiently in the case of shortage or excess of metals.

There are at least two regulatory systems for metals in bacteria. One system is protein based, metal regulatory proteins, and the other is based on noncoding RNA. There are at least two distinct groups of noncoding RNAs, involved in metal regulation. Small RNAs like RhyB and riboswitches such as magnesium, manganese and nickel/cobalt riboswitch. This thesis aims to discover *de novo* noncoding RNAs, especially riboswitches specific for metals, by using bioinformatics and molecular biology approaches.

Keywords : Noncoding RNA (ncRNA), Bioinformatics, Riboswitch, Metal Regulation.

TABLE DES MATIÈRES

REMERCIEMENTS.....	II
RÉSUMÉ.....	III
ABSTRACT.....	IV
Guide de lecteurs :.....	XIII
1. INTRODUCTION.....	1
1.1 La régulation chez les procaryotes :.....	1
1.2 L'ARN non codant (ARNnc) :.....	2
1.2.1 Les <i>riboswitchs</i> :.....	2
1.2.1.1 Les mécanismes de régulations :.....	2
1.2.1.2 Le repliement du <i>riboswitch</i> :.....	6
1.2.1.3 L'effet des cations sur les <i>riboswitchs</i> :.....	8
1.2.1.4 La structure des <i>riboswitchs</i> :.....	9
1.3 Le ribozyme <i>hammerhead</i> :.....	13
1.4 La régulation des métaux :.....	15
1.5 Le motif boucle E (<i>loop E</i>) :.....	24
1.5.1 Structure de la boucle E :.....	24
1.5.2 Le rôle de la boucle E et les métaux :.....	25
1.6 La prédiction d'ARN non codants :.....	26
1.7 Problématiques et hypothèse et objectifs.....	30
1.7.1 Problématique :.....	30
1.7.2 Hypothèse :.....	30
1.7.3 Objectifs :.....	30
2. Single mutation in hammerhead ribozyme favors cleavage activity with manganese over magnesium.....	32
Résumé :.....	34
Abstract :.....	35
2.1 Introduction.....	36
2.2 Results.....	38
2.2.1 Varying metal ion preference of a HHRz variant :.....	38
2.2.2 Effect of A6C mutation on another HHRz :.....	42
2.3 Discussion.....	45
2.4 Materials and Methods.....	47
2.4.1 PCR product of wild type and mutant ribozymes :.....	47
2.4.2 RNA transcription :.....	49
2.4.3 Kinetics of Cleavage :.....	49
2.5 Supplementary Materials.....	50
2.6 Author Contributions.....	50
2.7 Funding.....	50
2.8 Acknowledgments.....	51
2.9 Conflicts of Interest.....	51

2.10 References.....	51
3. Search for 5'-leader regulatory RNA structures based on gene annotation aided by the RiboGap database.....	54
Résumé :.....	56
Abstract:.....	58
3.1 Introduction.....	59
3.2 Materials.....	60
3.2.1 Bioinformatics.....	60
3.2.1.1 RiboGap:.....	60
3.2.1.2 GraphClust:.....	63
3.2.1.3 Optional software:.....	63
3.2.2 Experimental methods:.....	63
3.3 Methods and results.....	63
3.3.1 Finding known RNA structures:.....	63
3.3.2 Finding sequence motifs:.....	66
3.3.2.1 Known sequence motifs:.....	66
3.3.2.2 New sequence motifs:.....	68
3.3.3 Finding new structures:.....	68
3.3.3.1 Obtaining intergenic sequences with RiboGap:.....	68
3.3.3.2 Finding conserved elements and predicting RNA secondary structures:.....	71
3.3.3.3 How to analyze results:.....	72
3.3.3.3.1 Analyzing the structure alignments:.....	72
3.3.3.3.2 Analyzing potential riboswitches:.....	75
3.3.3.3.3 Additional considerations:.....	75
3.3.3.3.4 Performing a global homology search of candidate motifs by Infernal:.....	76
3.4 <i>In-line</i> probing for experimental validation:.....	76
3.4.1 PCR to construct the template for RNA production:.....	77
3.4.2 RNA transcription:.....	77
3.4.3 Dephosphorylation and Labeling:.....	78
3.4.4 Determination of candidate RNA structure and potential modulation by <i>in-line</i> probing:.....	78
3.5 Discussion.....	81
3.6 Acknowledgements.....	82
3.7 References.....	82
4. A survey of <i>cis</i> regulatory non-coding RNAs involved in bacterial virulence.....	86
Résumé :.....	87
Abstract :.....	89
4.1 Introduction.....	90
4.2 Materials and methods.....	91
4.2.1 Bioinformatics :.....	91
4.2.2 tRNA searches :.....	92
4.2.3 Northern of co-transcribed tRNAs :.....	92
4.3 Result.....	93
4.3.1 <i>Cis</i> -regulatory RNA distribution upstream of virulence factors :.....	93
4.3.2 tRNAs upstream of virulence factors :.....	95

4.3.3 Rho-independent transcription terminators (RiTT) :.....	95
4.4 Discussion.....	96
4.5 Acknowledgement.....	99
4.6 Funding.....	99
4.7 References.....	100
5. ARNnc candidats potentiels de <i>riboswitchs</i> calcium.....	106
5.1 Mise en contexte :.....	106
5.2 Matériels et Méthodes :.....	106
5.2.1 Bioinformatique :.....	106
5.2.2 Validation fonctionnelle :.....	107
5.2.2.1 Détermination de la structure et de l'affinité <i>in vitro</i> par <i>in line</i> <i>probing</i> :.....	107
5.2.2.2 Clonage :.....	107
5.2.2.3 Mesure d'expression par luminescence :.....	110
5.3 Résultats :.....	111
5.3.1 Prédiction de nouvelles familles d'ARNnc :.....	111
5.3.1.1 Structure ressemblant à <i>yybP-ykoY</i> :.....	111
5.3.2 Caractérisation des ARNnc ressemblants à <i>yybP-ykoY</i> :.....	116
5.3.2.1 <i>in line-probing</i> :.....	116
5.3.2.2 Mesurer l'expression du gène rapporteur ayant le motif <i>yybP-</i> <i>ykoY</i> chez <i>Pseudomonas aeruginosa</i>	118
6. Discussion.....	125
6.1 L'ARN, les métaux et les origines :.....	125
6.2 Combien de <i>riboswitchs</i> existent-ils sur la terre ?.....	126
6.3 Le motif <i>yybP-ykoY</i> (<i>riboswitch manganèse</i>) :.....	127
6.3.1 Contexte génétique :.....	127
6.3.2 La boucle E dans le motif <i>yybP-ykoY</i> (<i>riboswitch manganèse</i>) :.....	129
6.3.3 Le motif « ressemblant à <i>yybP-ykoY</i> », (<i>yybP-ykoY-r</i>) :.....	131
6.3.4 Comparaison des motifs <i>yybP-ykoY</i> et <i>yybP-ykoY-r</i> :.....	131
6.3.5 L'effet du pH et le motif <i>yybP-ykoY</i> et <i>yybP-ykoY-r</i> :.....	132
6.3.6 Les données transcriptomiques vis-à-vis du motif <i>yybP-ykoY</i> :.....	132
6.4 L'effet de variation des nucléotides sur la structure et la régulation du ribozyme et du <i>riboswitch</i> :.....	134
7. Conclusion	136
8. Perspectives.....	138
8.1 Étudier l'expression du gène rapporteur avec une version mutante non- fonctionnelle du <i>riboswitch yybP-ykoY-r</i> :.....	138
8.2 L'effet du pH :.....	138
8.3 Le motif RAGUA :.....	139
8.4 Terminaison « Rho dépendant » <i>yybP-ykoY P. aeruginosa</i> :.....	139
8.5 Sous famille L3 :.....	140
9. Bibliographie.....	141
10. Annexe I (Des séquences utilisées pour différentes constructions.).....	157
11. Annexe II (La comparaison des modèles de Rfam et le cluster 1.1.).....	160
12. Annexe III (matériels supplémentaires pour l'article : Single mutation in hammerhead ribozyme favors cleavage activity with manganese over magnesium...)	164
13. Annexe IV (matériels supplémentaires pour l'article : Search for 5'-leader	

regulatory RNA structures based on gene annotation aided by the RiboGap database).	
.....	165
14. Annexe V (matériels supplémentaires pour l'article : Survey of non-coding RNAs involved in bacterial virulence).....	166
15. Annexe VI (Article Methods in Molecular Biology).....	167
16. Annexe VII (Article Nature Communications).....	168
17. Annexe VIII (Des résultats de ONE way ANOVA).....	169
18. Annexe IX (droit d'auteurs de journaux).....	170

LISTE DES FIGURES

1. Figure 1.1 Des mécanismes principaux de la régulation par le <i>riboswitch</i>	3
2. Figure 1.2 Des mécanismes alternatifs des <i>riboswitchs</i>	5
3. Figure 1.3 La représentation de la structure de certains <i>riboswitchs</i>	10
4. Figure 1.4 Représentation schématique du ribozyme <i>hammerhead</i> et la réaction du clivage.....	14
5. Figure 1.5 Les concentrations des métaux chez <i>E.coli</i> cultivée dans le LB et le milieu minimal avec glucose mesuré par spectromètre de masse couplé à un plasma inductif, ICP-MS.....	16
6. Figure 1.6 La comparaison des ratios de volume du magnésium et du calcium.....	18
7. Figure 1.7 L'interaction des ions avec l'ARN.....	19
8. Figure 1.8 Le consensus de la boucle E retrouvé chez l'ARN ribosomal 5S et 23S.....	24
9. Figure 1.9 L'interaction de magnésium avec la boucle E de l'ARNr 5S.....	25
10. Figure 2.1 The structure and sequence consensus of the hammerhead ribozyme (HHRz).....	37
11. Figure 2.2 Cleavage assay for the hammerhead ribozyme <i>Bcep176</i> (C6).....	39
12. Figure 2.3 Importance of C6 in the bacteriophage hammerhead ribozyme <i>Bcep176</i>	41
13. Figure 2.4 Effect of A6C mutation on self-cleavage activity for a pseudoknotted type II consensus core HHRz (A6) derived from the mouse gut metagenome (mouse gut HHRz).....	43
14. Figure 3.1 Schematic representation of tables and major applications of RiboGap.....	61
15. Figure 3.2 Screenshot of RiboGap interface from the Advanced search page.....	62
16. Figure 3.3 Results from RiboGap queries.....	65
17. Figure 3.4 Schematic presentation of the pipeline for novel RNA structures discovery.....	70
18. Figure 3.5 RNA methylation-related candidate "RNA-methyl-28".....	74
19. Figure 3.6 In-line probing of the RNA-methyl-28 candidate.....	80
20. Figure 5.1 La présentation schématique de la séquence à cloner dans le vecteur mini-CTX-Lux.....	108
21. Figure 5.2 Le vecteur mini-CTX-Lux et le <i>riboswitch yybP-ykoY</i> cloné.....	109
22. Figure 5.3 L'alignement obtenu à partir des séquences retrouvées en aval des gènes ayant un rapport au calcium.....	112
23. Figure 5.4 La comparaison des modèles du <i>riboswitch</i> du manganèse et le candidat du calcium.....	113
24. Figure 5.5 Deux candidats obtenus prédits, à partir de gènes liés au calcium.....	115
25. Figure 5.6 <i>in-line probing</i> obtenu à partir du motif retrouvé chez <i>Desulfovibrio magneticus</i> RS-1 en présence de manganèse et calcium.....	117
26. Figure 5.7 A Déterminer l'effet des agents chélateurs l'EDTA et l'EGTA dans le milieu M9 pour la bactérie <i>Pseudomonas aeruginosa</i>	120
27. Figure 5.7 B Déterminer l'effet des agents chélateurs l'EDTA et l'EGTA dans le milieu M9 pour la bactérie <i>Pseudomonas aeruginosa</i>	121
28. Figure 5.8 L'effet de l'agent chélateur sur le contrôle négatif (un <i>riboswitch</i> SAM) dans le milieu M9.....	123
29. Figure 5.9 La normalisation des résultats obtenus par rapport à ceux du <i>riboswitch</i> SAM chez <i>Pseudomonas aeruginosa</i>	124
30. Figure 6.1 La famille de réseaux TerC.....	128
31. Figure 6.2 L'interaction de la boucle E avec le manganèse dans le motif <i>yybP-ykoY</i>	130

LISTE DES TABLEAUX

1. Tableau 1.1 Les ligands qui peuvent être reconnus par les <i>riboswitchs</i>	11
2. Tableau 1.2 Certaines propriétés physico-chimiques des cations.....	20
3. Tableau 1.3 Le résumé de régulations des métaux chez les bactéries.....	21
4. Table 2.1 Cleavage rates of <i>Bcep176</i> (C6) variant and pseudoknotted type II mouse gut HHRz (mouse gut HHRz) with comparative Mn ²⁺ and Mg ²⁺ concentrations.....	44
5. Table 2.2 Wild type and mutant sequences for <i>Bcep176</i> HHRz and the primer sequences.....	48
6. Table 4.1 Virulence factors (VFs) with <i>cis-regulatory</i> RNAs.....	94

LISTE DES ABRÉVIATIONS

2'-dG : 2'-désoxyguanosine, 2'-deoxyguanosine.

A

ARN : Acide ribonucléique.

ARNm : ARN messenger.

ARNnc : ARN non codant.

ATP : adénosine triphosphate.

AdoCbl : Adénosylcobalamine, (Adenosylcobalamin).

ADP : Adénosine-5'-diphosphate, (Adenosine-5'-diphosphate).

B

BLASTn : Basic Local Alignment Search Tool nucleotide.

C

c-AMP-GMP : cyclic adénosine monophosphate, guanosine monophosphate.

CarS : Calcium regulated Sensor.

Car R : Calcium regulated Regulator.

c-di-GMP : cyclic diguanylate.

c-di-AMP : cyclic di-adénosine monophosphate.

CF : Cystic Fibrosis, fibrose kystique.

CDP : Cytidine-5'-diphosphate.

D

DO : Densité optique.

DtxR : Diphteria toxin repressor.

E

EDTA : Éthylènediaminetétraacétique, Ethylenediaminetetraacetic acid.

EGTA : Éthylène glycol-bis(β -aminoethyl ether)-N, N, N', N'-tetraacetic acid.

F

fM: femtomolaire.

FMN : Flavine mononucléotide, (Flavin mononucleotide).

Fur : Ferric uptake protein.

G

GlcN6P : glucosamine-6-phosphate.

H

HHRz : Hammerhead ribozyme.

HMP-PP : 4-amino-5-hydroxymethyl-2-methylpyrimidine diphosphate

I

ICP-MS : Spectrométrie de masse couplée à un plasma inductif, (Inductively Coupled Plasma Mass Spectrometry).

J

K

L

LB : Luria-Bertani medium.

M

M9 : Minimal salt medium.

μM : micromolaire.

MFE : Énergie minimum libre, (Minimum Free energy).

MoCo : Cofacteur à molybdène, (Molybdenum cofactor).

N**O****P**

PHB/polyP : poly-(R)-3-hydroxybutyrate/polyphosphate.

ppGpp : Guanosine tétraphosphate, (Guanosine tetraphosphate).

PRPPP : Phosphoribosyl pyrophosphate.

PreQ1 : Pre-queuosine1.

Q**R**

RBS : Ribosome Binding Site.

REGEX : Regular Expression.

S

SAM : S-adénosylméthionine, (S-Adenosyl methionine).

SAH : S-adénosylhomocystéine, (S-adenosylhomocysteine).

SD : Shine-Dalgarno.

SCFG : Stochastic context free grammar.

T

TCA : Cycle de l'acide citrique, (Tri Carboxylic Acid).

THF : Tétrahydrofolate, (Tetrahydrofolate).

TPP : Pyrophosphate de thiamine, (Thiamine pyrophosphate).

U

UTR : UnTranslated Region.

V**W**

Wco : tungsten cofactor.

X**Y****Z**

ZTP : N10-formyl-THF.

Guide de lecteurs :

Voir en bref les chapitres de cette thèse :

Le chapitre 1 consiste en une introduction qui survole la littérature des différents sujets abordés : ARNnc (*riboswitchs* et ribozymes), structure d'ARN et bioinformatique et régulation de l'homéostasie des métaux.

Le chapitre 2 (article re-soumis) décrit un premier exemple d'un nouvel ARNnc qui aurait un potentiel comme senseur de cation métallique : un variant du ribozyme *hammerhead* qui coupe préférentiellement en présence de Mn^{2+} .

Le chapitre 3 (article publié) décrit la base de données RiboGap et un pipeline de découverte de nouveaux ARNnc. Ce chapitre est central pour les travaux qui suivront.

Le chapitre 4 (article à soumettre) utilise RiboGap pour établir une liste d'ARNnc qui pourraient être impliqués dans la virulence, incluant des senseurs de métaux.

Le chapitre 5 (travaux en cours) décrit de nouveaux ARNnc potentiels senseurs de Ca^{2+} , découverts par le pipeline décrit dans le chapitre 3.

Le chapitre 6 est une discussion générale sur la possibilité de régulation des métaux par des ARNnc et revient sur les résultats des chapitres précédents avec une vue plus globale sur le sujet.

Le chapitre 7 récapitule les éléments de cette thèse.

Le chapitre 8 décrit quelques perspectives et des propositions pour faire suite aux travaux de cette thèse.

Le chapitre 9 est la bibliographie.

1. INTRODUCTION

1.1 La régulation chez les procaryotes :

Maintenir l'homéostasie de cellules vivantes y compris des bactéries est un processus actif et complexe qui a besoin d'expression ou de répression de plusieurs gènes (Mandal et al., 2004a). Les bactéries subissent plusieurs formes de stress tels que la variation de pH, de nutriments, de température, d'ions, etc. Afin de maintenir la stabilité cellulaire, les bactéries ont plusieurs niveaux de régulation telles que des protéines régulatrices (Osman et al., 2010) et l'ARN non codants (Waters et al., 2009). Parmi des éléments mentionnés à réguler, la régulation des métaux chez les bactéries sera étudiée davantage dans cette thèse.

D'une part, les métaux sont essentiels pour plusieurs activités vitales y compris des bactéries, d'autre part en concentration élevée ils sont toxiques pour les cellules, par conséquent les bactéries ont aussi besoin de les éliminer. L'activité oxydoréductase des métaux transitoires leur confère la caractéristique à la fois de toxicité et à la fois catalytique (Palmer et al., 2016). Des métaux comme le magnésium, le manganèse, le zinc et le fer sont utilisés davantage comme cofacteur par certains enzymes ; par exemple le fer est impliqué dans l'activité oxydoréductase (Waldron et al., 2009; Chandrangu et al., 2017). Les métaux sont particulièrement importants dans la virulence des bactéries (Turner et al., 2015), par exemple le magnésium Mg^{2+} , joue un rôle important dans la virulence chez *Salmonella typhimurium* (Garcia Vescovi et al., 1996) et chez *Yersinia pestis* (Perry et al., 2015). De même le fer est impliqué dans la virulence chez *Listeria monocytogenes* (Pi et al., 2016).

L'homéostasie des métaux chez les bactéries dépend de transporteurs (Chandrangu et al., 2017). Quant à l'homéostasie des métaux, il faut considérer les points suivant : **a)** le quota des métaux : ce qui indique le besoin total des cellules (atome/cellule). **b)** la fraction labile des métaux qui change et qui est partiellement accessible pour les protéines et l'ARN pendant la traduction et la transcription, et finalement **c)** la concentration libre dans la cellule bactérienne au point de vue thermodynamique (Chandrangu et al., 2017).

1.2 L'ARN non codant (ARNnc) :

Bien que les définitions puissent varier légèrement selon les sources, pour cette thèse nous entendons par ARNnc des structures d'ARN définies qui ne codent pas pour des protéines. Plusieurs de ces ARN pourraient plutôt avoir un rôle de régulation dans les bactéries (Eddy, 2001). Quelques exemples d'ARNnc bactériens incluent : **T-Box** (Henkin et al., 2006), une partie de l'ARNm qui régule le gène en amont en présence ou absence des ARNt chargés ; **ribozymes**, l'ARN avec l'activité catalytique (Altman, 1990; Cech, 1990) ; **riboswitchs** partie de l'ARNm qui régule le gène en aval en présence ou absence des métabolites ou des ions (Mironov et al., 2002; Nahvi et al., 2002; Winkler et al., 2002a,b) et des petits ARN (small RNA, **sRNA**) qui régule en s'appariant avec leur l'ARNm cible à l'aide de certaines protéines chaperonnes (Mizuno et al., 1984). Il y a plusieurs types d'ARNnc régulateurs chez les bactéries, mais au moins deux grands groupes hétérogènes, les *riboswitchs* et les *sRNA* dont plusieurs classes ont des rôles dans la régulation des métaux. Parmi ces ARNnc les *riboswitchs* et le ribozyme *hammerhead* seront abordés dans cette thèse.

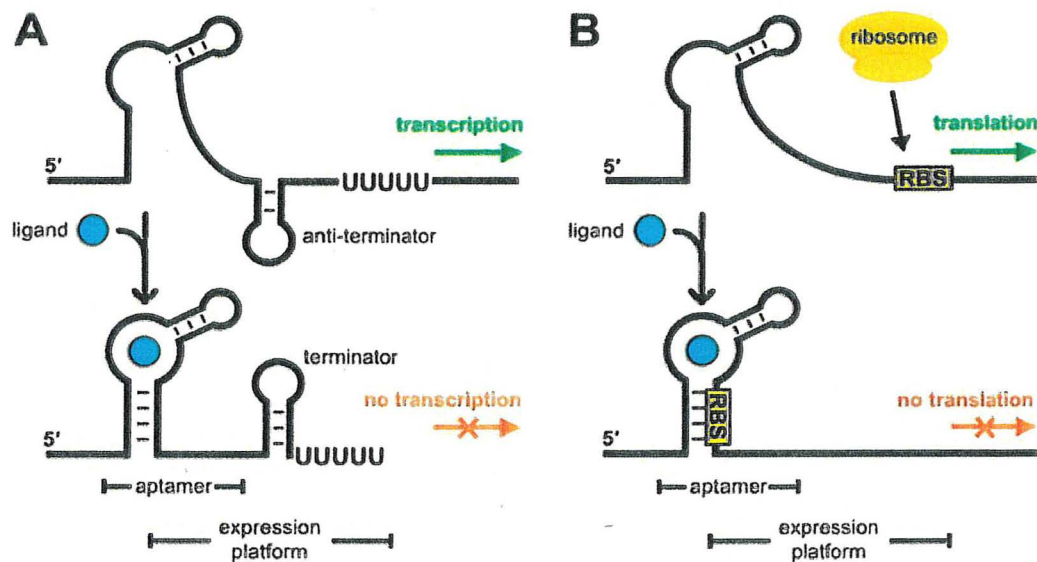
1.2.1 Les *riboswitchs* :

1.2.1.1 Les mécanismes de régulations :

Les *riboswitchs* sont des ARNnc qui jouent un rôle très important dans la régulation. Les *riboswitchs* font partie de la structure de l'ARNm et sont capables de reconnaître une molécule particulière telle que des métabolites ou des ions sans avoir besoin de protéine (Nahvi et al., 2002). Les *riboswitchs* chez les bactéries se retrouvent souvent dans la partie 5'-UTR (**UnTranslatedRegion**), et ils exercent leurs rôles de régulation sur des gènes en aval. Le nom « *switch* » indique le mécanisme de cette structure d'ARN qui peut avoir deux états, « ON » et « OFF » (Breaker, 2012a). Le *riboswitch* thiamine pyrophosphate (Winkler et al., 2002a), se retrouve également chez les plantes (Wachter et al., 2007) et les champignons (Li et al., 2017).

Chaque *riboswitch* consiste en deux parties : **a)** l'*aptamère*, qui reconnaît de manière spécifique le ligand, et **b)** la plate-forme d'expression qui module l'expression de gènes en aval (Breaker, 2011; Garst et al., 2011). Il existe une partie de séquence commune entre l'*aptamère* et la plate-forme d'expression, ce qui permet au *riboswitch* de changer de la configuration « ON » à « OFF » et vice-versa. Ces conformations sont

mutuellement exclusives (Breaker, 2012a). Les *riboswitchs* ont deux mécanismes principaux pour la régulation des gènes (Figure 1.1). Ces deux mécanismes jouent sur la transcription (Mironov et al., 2002; Sudarsan et al., 2003), et/ou aussi sur la traduction des gènes en aval (Nahvi et al., 2002; Winkler et al., 2002a). Dans le cas de la transcription (Figure 1.1 A), en absence ou en présence du ligand, l'arrêt de transcription « Rho indépendant » suivi par des uridines (Farnham et al., 1981) pourrait se former ou se défaire selon le *riboswitch*. De même manière, pendant l'initiation de la traduction, (Figure 1.1 B), le site de reconnaissance du ribosome (RBS, *Ribosome Binding Site*), ou Shine-Dalgarno, (SD) (Shine et al., 1975), peut être séquestré ou être libéré, selon le *riboswitch* (Breaker, 2012a).



1. Figure 1.1 Des mécanismes principaux de la régulation par le *riboswitch*.

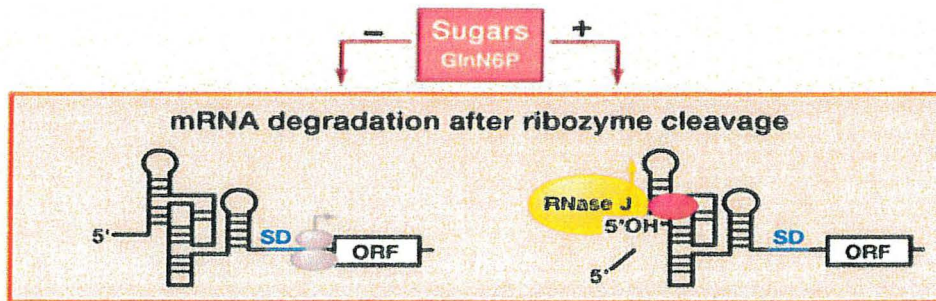
A) La régulation par la modulation de l'arrêt de transcription : En absence du ligand la tige anti-terminatrice se forme, et le gène se transcrit et donc le *riboswitch* est en état « ON ». En présence du ligand, l'*aptamère* stabilise la tige du terminateur de transcription, suivi par la queue de poly U, ce qui entraîne par conséquent l'arrêt de la transcription et donc le *riboswitch* est dans l'état « OFF ». B) La régulation par la modulation de la traduction : En absence du ligand, le ribosome peut se lier au site Shine-Dalgarno (RBS) et par conséquent, le gène se traduit. Alors le *riboswitch* est en état « ON ». En présence du ligand, l'*aptamère* séquestre le site Shine-Dalgarno, empêchant ainsi le ribosome de se lier, et donc la traduction ne peut pas initier, et donc le *riboswitch* est en état « OFF ». Tiré de (Kim et al., 2008).

Les *riboswitchs*, pour un ligand spécifique, peuvent avoir différents mécanismes de régulation selon l'espèce. Par exemple, le *riboswitch* adénine a deux modes d'action tels qu'expliqués plus haut (Figure 1.1) (Mandal et al., 2004b). Ce mécanisme de régulation est dû à une séquence commune entre l'*aptamère* et la plate-forme d'expression (Rieder et al., 2010).

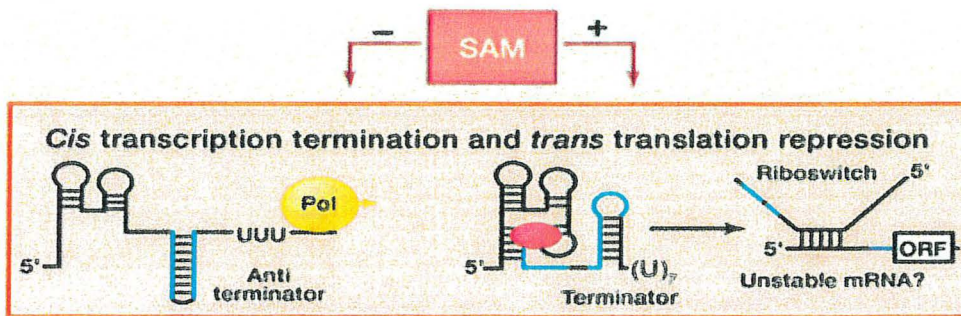
Des mécanismes alternatifs de régulation ont aussi été découverts. Un de ces mécanismes est la régulation « Rho dépendante ». Certains *riboswitchs* régulent par le mécanisme « Rho dépendant » en induisant la terminaison de la transcription. Ce mécanisme paraît répandu au sein des bactéries (Hollands et al., 2012). Le mécanisme « Rho dépendant » peut aussi réguler la transcription chez certains *riboswitchs* qui fonctionnent aussi sur l'initiation de la traduction. Par exemple, il a été suggéré que le mécanisme « Rho dépendant » du *riboswitch* TPP commence après la répression de la traduction chez *E. coli* afin d'assurer une double contrôle pour l'expression du gène (Bastet et al., 2017). D'autres mécanismes d'inhibition ont aussi été démontrés, comme la dégradation de l'ARNm via l'ARNase E ; par exemple lorsque la lysine se lie au *riboswitch* *LysC* (Caron et al., 2012).

Il existe d'autres mécanismes complexes, mais moins fréquents. La régulation en tandem pour un même ligand particulier, permet de lier de façon coopérative la glycine (Mandal et al., 2004c) et d'améliorer la réponse à TPP (Welz et al., 2007) ou encore deux *riboswitchs* pour deux ligands différents, SAM et, adénosylcobalamine, (B₁₂) qui permet d'analyser deux « entrées » (*input*, la concentration de ces deux ligands) pour produire une « sortie » (*output*, l'expression du gène) (Sudarsan et al., 2006). Un autre mode de régulation est utilisé par le *riboswitch-ribozyme glmS* (Winkler et al., 2004). Lorsque l'ARN *glmS* lie son ligand, la Glucosamine-6-phosphate (GlcN6P), ce dernier sert de cofacteur pour catalyser la coupure de l'ARN, la partie 5'-OH libre sera ensuite dégradée par l'ARNase J (Collins et al., 2007) et par conséquent l'ARNm sera dégradé (Figure 1.2 A). Bien que le mécanisme *cis-acting* des *riboswitchs* soit répandu, un mécanisme *trans-acting* de *riboswitch* (appariement avec un ARN cible autre que son propre ARNm afin de réprimer l'expression de ce dernier) (Loh et al., 2009) et aussi un mécanisme d'antisens (l'ARN qui se transcrit dans le sens inverse et s'apparie avec l'ARNm du gène à réprimer) (Mellin et al., 2013) ont été identifiés chez *Listeria monocytogenes* (Figure 1.2 B).

A



B



2. Figure 1.2 Des mécanismes alternatifs des *riboswitchs*

A) La liaison de glucosamine-6-phosphate (GlcN6P) change la conformation du *riboswitch* et le ribozyme se clive. La partie 5'-OH sera dégradée par l'ARNase. B) En absence de SAM chez *L. monocytogenes*, les *riboswitchs* *SreA* et *SreB* forment un anti-terminateur et la transcription se réalise. En présence de SAM, l'arrêt de la transcription se produit et l'ARN messager (l'ARNm) fonctionnera comme les petits ARN *en trans* et déstabilisera sa cible. Tiré de (Serganov et al., 2013).

1.2.1.2 Le repliement du *riboswitch* :

La fonction de régulation des *riboswitchs* commence par la transcription de l'*aptamère* qui reconnaîtra ensuite le ligand (Garst et al., 2009). Par exemple, le *riboswitch* adénine, « *add* » chez *Vibrio vulnificus*, qui fonctionne en affectant la traduction, se replie même en absence du ligand mais la liaison du ligand stabilise la conformation de l'*aptamère* et permet d'augmenter l'expression (Rieder et al., 2007). De même manière, il a été montré que sans ligand, le *riboswitch* fluor, adopte sa conformation active pendant la transcription avec une durée de vie très courte (3 ms), ce qui mène par conséquent à l'arrêt de la transcription ; Par contre en présence de fluor, la conformation de l'*aptamère* change pour empêcher la formation de l'arrêt de transcription et ainsi la transcription se continue (Zhao et al., 2017).

L'*aptamère* du *riboswitch* peut reconnaître son ligand avec un K_d variant du picomolaire jusqu'au micromolaire (Breaker, 2012a). Deux états mutuellement exclusifs peuvent se former au point de vue d'équilibre thermodynamique pour les *riboswitchs* (Furtig et al., 2015). En rajoutant le ligand, le *riboswitch* se rend dans l'état qui favorise la structure de l'*aptamère*, ce qui affectera la structure de la plate-forme d'expression pour activer ou réprimer le gène (Badelt et al., 2015).

Il a été montré que pour certains *riboswitchs*, la concentration du ligand requise est supérieure d'au moins un ou plusieurs ordres de grandeur à la constante de dissociation apparente, K_d , du *riboswitch* (Breaker, 2012a). Ce type de *riboswitch* fonctionne sous le modèle cinétique et le changement d'un état à l'autre n'est pas réversible (Badelt et al., 2015).

Dans le modèle cinétique, le ligand doit lier avant que la plate-forme d'expression se replie. Il existe des sites de pause qui ralentissent la transcription et par conséquent l'*aptamère* aura plus de temps pour se replier. L'importance des sites de pause pour le bon fonctionnement de *riboswitchs* a aussi été démontrée pour le *riboswitch* *btuB*, B_{12} chez *Escherichia coli* (Perdrizet et al., 2012) le *riboswitch* de la guanine, *xpt-pbuX* chez *Bacillus subtilis* (Steinert et al., 2017) et les *riboswitchs* *ThiC* et *ThiM* chez *Escherichia coli* (Chauvier et al., 2017; Uhm et al., 2018).

Dans le modèle thermodynamique, le ligand peut se lier en même temps que le *riboswitch* complet se replie (Coppins et al., 2007). Par exemple, il a été démontré que dans le modèle cinétique le *riboswitch* flavine mononucléotide (FMN) n'atteint pas l'équilibre au point de vue de l'énergie libre minimale, et par conséquent, il faut plus de concentration de

ligand au moment où la transcription se continue ou s'arrête (Wickiser et al., 2005). La concentration du ligand et la vitesse de transcription sont importantes pour le repliement de l'*aptamère* fonctionnant sous le modèle cinétique (Wickiser et al., 2005). Dans le cas de certains *riboswitchs* comme celui d'adénine, « *pbuE* » chez *Bacillus subtilis* les conditions comme la température et la vitesse de la transcription, peuvent porter la transcription vers l'équilibre thermodynamique ou bien vers un modèle cinétique (Wickiser et al., 2005). Un des facteurs les plus importants est la transcription elle-même, qui a un rôle déterminant pour le repliement et la fonction de certains *riboswitchs*. Le *riboswitch* adénine « *pbuE* » chez *Bacillus subtilis*, fonctionne sous le modèle cinétique, il dépend de la transcription (Lemay et al., 2006) et son repliement co-transcriptionnel selon la concentration du ligand est déterminant (Frieda et al., 2012; Perdrizet et al., 2012; Chauvier et al., 2017; Steinert et al., 2017; Uhm et al., 2018).

Pour mieux comprendre ce qui peut changer de l'état « OFF » à l'état « ON », pour le *riboswitch* sous le modèle cinétique, il a été suggéré que des structures méta-stables indépendantes du ligand se forment, ce qui permet de continuer la transcription sans rentrer dans l'équilibre (Helmling et al., 2017). De plus, pour certains *riboswitchs*, comme celui de 2'-désoxyguanosine (2'-dG), la liaison du ligand favorise le ralentissement du repliement de l'*aptamère*, permettant à la polymérase de dépasser la majorité des structures méta-stables (Helmling et al., 2018).

1.2.1.3 L'effet des cations sur les *riboswitchs* :

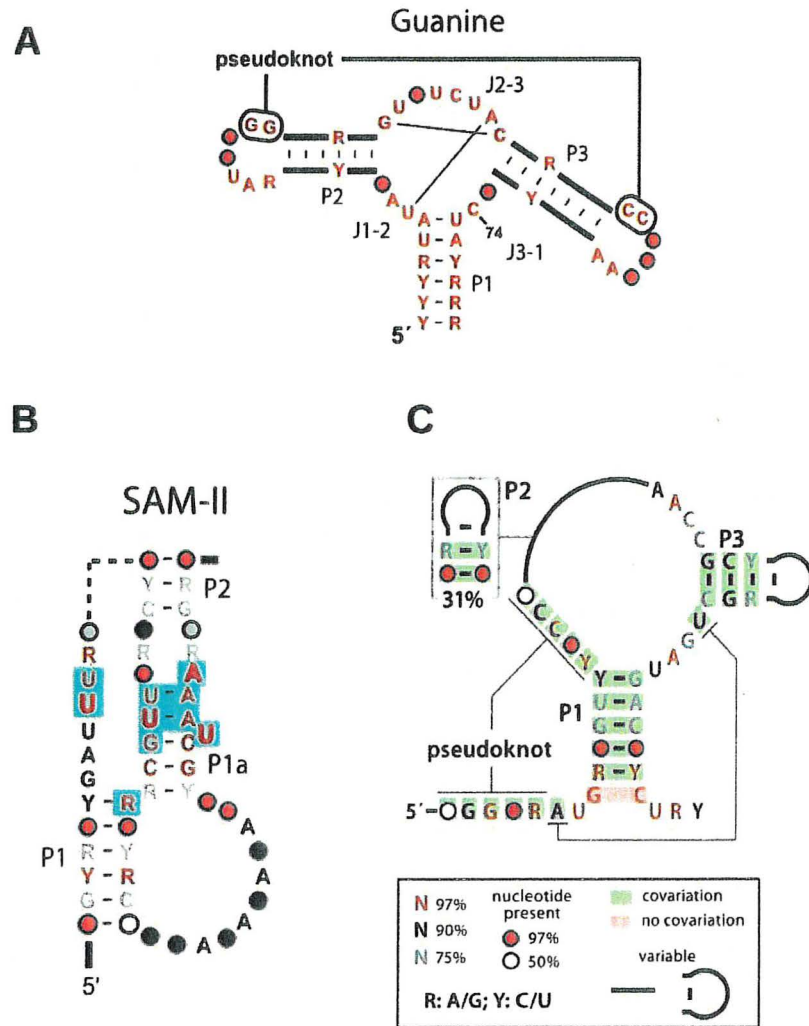
Tout ARN est chargé négativement dû à ses groupements phosphates, ce qui cause une répulsion réduisant la spontanéité de la formation des paires de bases. Par conséquent, tout cation peut diminuer la charge négative de l'ARN pour permettre son bon repliement (Draper, 2004). Les métaux, et surtout le magnésium, stabilise la structure 3D de l'ARN (Draper, 2004). Le magnésium et le potassium sont les cations les plus abondants dans le milieu intercellulaire et ainsi ils font des liaisons non spécifiques avec tous les acides nucléiques (Ferre-D'Amare et al., 2011).

Des ligands ayant des groupements phosphates peuvent être reconnus par le *riboswitch* via des interactions avec le magnésium, Mg^{2+} (Ferre-D'Amare et al., 2011). Par exemple, le *riboswitch* thiamine pyrophosphate (TPP) peut reconnaître son ligand via l'aide de Mg^{2+} qui interagit avec le groupement de phosphate dans le ligand (Noeske et al., 2006). Ce *riboswitch* peut se lier à son ligand en présence de 1 mM de Mg^{2+} tandis qu'il ne peut pas se lier à son ligand en présence de cations uniquement monovalents tel que Na^+ et K^+ , même en concentration élevée (Yamauchi et al., 2005). Le *riboswitch* adénine, « *pbuE* » devient partiellement organisé pour accepter le ligand en présence du magnésium (Noeske et al., 2007). Un autre exemple montre que la concentration d'au moins 0,5 mM de magnésium est nécessaire pour le repliement du *riboswitch* cobalamine afin de reconnaître son ligand et afin de permettre la transition d'un état à l'autre (Choudhary et al., 2014). Le magnésium permet de replier *riboswitch* LysC pour être correctement accessible pour son ligand (McCluskey et al., 2019). Le magnésium n'est pas le seul cation nécessaire pour les *riboswitchs*. Pour certains *riboswitchs* le magnésium pourrait être remplacé par d'autres cations. Par exemple le *riboswitch* lysine a plus d'affinité pour l'ion potassium (K^+) que magnésium, Mg^{2+} (Serganov et al., 2008).

1.2.1.4 La structure des *riboswitchs* :

Différentes classes de structures de *riboswitchs* ont été identifiées. Les *riboswitchs* varient en taille de 30 à 200 nucléotides (Lippa et al., 2012). Deux formes principales d'architecture pour les *riboswitchs* ont été proposées l'une est multi-jonctionnelle comme le *riboswitch* purine (Figure 1.3 A), alors que la deuxième comporte un pseudonœud comme les *riboswitchs* S-adénosylméthionine-II (SAM-II) (Figure 1.3 B) et fluor (Figure 1.3 C) (Serganov et al., 2013). La forme multi-jonctionnelle pourrait avoir trois jonctions, comme les *riboswitchs* qui reconnaissent la famille des purines, ou quatre jonctions comme le *riboswitch* SAM-I (Schwalbe et al., 2007). Les *riboswitchs* Lysine et flavine mononucléotide (FMN) ont cinq jonctions (Serganov et al., 2013). Il a été suggéré que les *riboswitchs* ayant une structure plus complexe sont plus anciens au point de vue d'évolution et les *riboswitchs* ayant les structures plus simples sont des nouveaux *riboswitchs*, ce qui permet de penser qu'il y aurait encore de nombreux *riboswitchs* non identifiés (Breaker, 2012a).

La spécificité d'un *aptamère* pour son ligand suggère la conservation de la structure tridimensionnelle au sein de l'évolution (Barrick et al., 2007). La conservation de la structure est cruciale pour la reconnaissance du ligand, mais dans certains cas, une mutation dans la séquence peut changer le ligand reconnu. Par exemple, le *riboswitch* qui reconnaît la guanine (Mandal et al., 2003) et l'adénine ont la même structure, mais avec un nucléotide différent. L'ARN qui serait devenu le *riboswitch* adénine aurait subi une mutation du C → U (Mandal et al., 2004b). De même manière, quelques mutations de la séquence de même structure permettent de reconnaître la 2'-désoxyguanosine au lieu de la guanine (Kim et al., 2007). Inversement, un ligand peut être reconnu par plusieurs structures de *riboswitchs*. Par exemple, il existe plusieurs classes de *riboswitchs* pour le SAM (Epshtein et al., 2003; McDaniel et al., 2003; Winkler et al., 2003; Serganov et al., 2013).



3. Figure 1.3 La représentation de la structure de certains *riboswitchs*.

A) La structure multi-jonctionnel du *riboswitch* guanine B, C) La structure des *riboswitchs* SAM-II et fluor avec le pseudonœud Tiré de (Breaker, 2011; Breaker, 2012b).

Il existe parfois plusieurs classes de *riboswitchs* pour un seul ligand ou bien des structures similaires pour des ligands différents. Autrement dit, il faut s'attendre à retrouver encore plusieurs nouveaux *riboswitchs*. Il existe à date près de 40 classes de *riboswitchs*, selon des données cristallographiques et la reconnaissance de ligands, le plus répandu est le *riboswitch* TPP dont on retrouve plusieurs milliers d'exemplaires dans les génomes séquencés de tous les domaines du vivant alors que tout juste une dizaine d'exemplaires

sont connus pour d'autres *riboswitchs*, tels que le *riboswitch* nickel-cobalt (Furukawa et al., 2015) *NiCo*, flavine mononucléotide, FMN variant (Blount et al., 2015) ou 2'-désoxyguanosine (Kim et al., 2007).

Le tableau ci-après présente tous les *riboswitchs* connus à cette date adapté de (Breaker, 2011).

1. Tableau 1.1 Les ligands qui peuvent être reconnus par les *riboswitchs*.

Coenzymes	Nucléotides et dérivés	Acides Amines	Autres métabolites	Ions
AdoCbl ¹	Guanine ⁹	Glycine ¹⁹	GlcN6P ²² (sucre)	Mg ²⁺ ^{25, 26}
TPP ²	Adénine ¹⁰	Lysine ²⁰	Azaromatic ²³	F ⁻ ²⁷
FMN ³	PreQ1 ¹¹	Glutamine ²¹	Guanidine ²⁴	Mn ²⁺ , 28, 29, 30
THF ⁴	2' – dG ¹²			Ni ²⁺ /Co ²⁺ ³¹
SAM ⁵	c-di-GMP ¹³			H ⁺ ³²
SAH ⁶	c-di-AMP ¹⁴			
MoCo ⁷	c-AMP-GMP ¹⁵			
WCo ⁸	ADP, CDP ¹⁶			
HMP – PP ³³	ZTP ¹⁷			
	ppGpp ¹⁸			

1. (Nahvi et al., 2002)
2. (Mironov et al., 2002; Winkler et al., 2002b)
3. (Gelfand et al., 1999; Mironov et al., 2002; Winkler et al., 2002b)
4. (Ames et al., 2010)
5. (Epshtein et al., 2003; McDaniel et al., 2003; Winkler et al., 2003)
6. (Wang et al., 2008)
- 7, 8. (Regulski et al., 2008a)
- 9, 10. (Mandal et al., 2003; Mandal et al., 2004b)
11. (Roth et al., 2007; Meyer et al., 2008)
13. (Sudarsan et al., 2008; Lee et al., 2010)
14. (Nelson et al., 2013)
15. (Nelson et al., 2015)
16. (Sherlock et al., 2019)
17. (Kim et al., 2015)
18. (Sherlock et al., 2018)
19. (Mandal et al., 2004c)
20. (Grundy et al., 2003; Sudarsan et al., 2003)
21. (Ames et al., 2011)
22. (Winkler et al., 2004)
23. (Li et al., 2016)
24. (Nelson et al., 2017)
- 25, 26. (Cromie et al., 2006; Dann et al., 2007)
27. (Baker et al., 2012)
- 28, 29, 30. (Shi et al., 2014; Dambach et al., 2015; Price et al., 2015)
31. (Furukawa et al., 2015)
32. (Nechooshtan et al., 2009)
33. (Atilho et al., 2019)

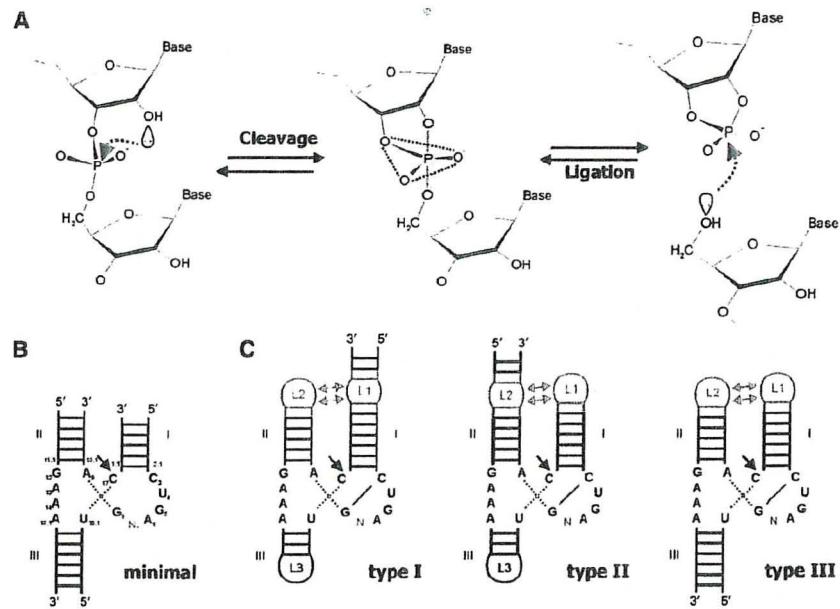
1.3 Le ribozyme *hammerhead* :

Avant que nous ne connaissions l'existence des *riboswitchs*, la découverte de ribozymes (des enzymes faites d'ARN) avait déjà démontré la capacité de l'ARN à adopter des structures complexes. L'un des ribozymes les plus utilisés comme modèle pour l'étude du lien entre la structure et la fonction de l'ARN est le ribozyme *hammerhead*.

La première évidence du ribozyme *hammerhead* a été rapportée par (Prody et al., 1986) et suivi de la découverte d'au moins neuf autres familles d'ARN ayant la capacité d'auto clivage y compris le ribozyme *hammerhead* (HHRz) (de la Pena et al., 2017). Le ribozyme *hammerhead* coupe sa propre séquence par une réaction de trans estérification selon le mécanisme de S_N2 (Figure 1.4 A) (Hammann et al., 2012). Le cœur catalytique minimal est constitué de nucléotides conservés qui sont flanqués de trois hélices (Figure 1.4 B) (Hammann et al., 2012). Le HHRz a été divisé dans trois groupes selon l'ouverture de la boucle de chaque tige (Figure 1.4 C) (Hammann et al., 2012).

Il a été montré que le HHRz existe dans toutes les domaines de la vie (de la Pena et al., 2010; Perreault et al., 2011). En utilisant des méthodes de recherche par homologie, plusieurs nouveaux HHRz ont été découverts (Perreault et al., 2011). La formation d'un pseudonœud ou autre type d'interaction tertiaire naturelle entre les tiges I et II forme la structure tertiaire qui est nécessaire pour l'activité catalytique du HHRz (Perreault et al., 2011).

Le HHRz a besoin d'ions métalliques, surtout le magnésium, et un pH physiologique pour l'activité catalytique (Dahm et al., 1993). Il a été aussi montré que le HHRz coupe *in vitro* avec une concentration élevée de 10 mM de Mg^{2+} (Hertel et al., 1994). Le HHRz peut avoir une activité catalytique en présence d'autres ions positivement chargés y compris les ions monovalents ou l'ammonium (NH_4^+) (Murray et al., 1998). En conséquence, il a été suggéré que les ions chargés positivement agissent de manière non spécifique avec le HHRz et aide au repliement plutôt que de participer à la catalyse (Klein et al., 2006). Pour certains HHRz, comme celui présent chez *Schistosoma*, la cinétique du clivage change en fonction de différents métaux tandis que ces métaux replieraient l'ARN plus ou moins de la même manière (Boots et al., 2008).



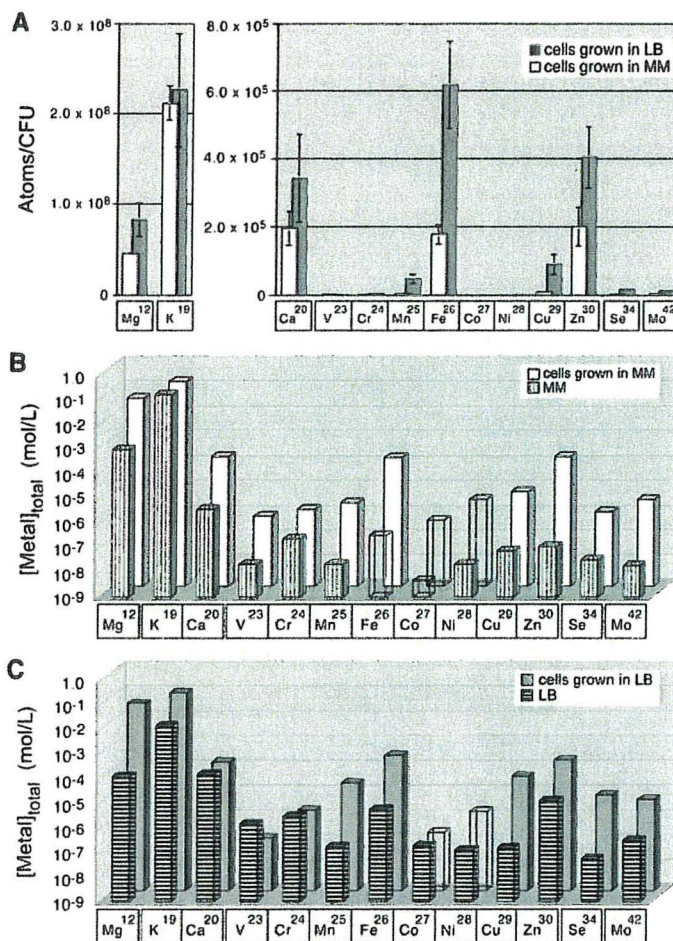
4. Figure 1.4 Représentation schématique du ribozyme *hammerhead* et la réaction du clivage.

A) La réaction de clivage du ribozyme *hammerhead*. Le mécanisme s'appelle S_N2 dans laquelle le groupe d'oxygène du carbone 2' attaque le voisinage 3'-5' de liaison phosphodiester. Le 2', 3' cyclique phosphate et 5'-hydroxyle se produisent dans les fragments 5' et 3' respectivement. B) La structure minimale et la partie de coupure du HHR. C) Les trois différents types d'ouverture des tiges avec l'interaction des boucles L1, et L2. Tiré de (Hammann et al., 2012).

1.4 La régulation des métaux :

Les métaux jouent plusieurs rôles essentiels dans les cellules. Plusieurs aspects du métabolisme dépendent des cations. Certains métaux sont essentiels pour la structure de protéines ou peuvent agir comme cofacteur (Waldron et al., 2009). Par exemple l'ADN et l'ARN polymérase ont besoins de Mg^{2+} comme cofacteur (Vernon, 1988). Le repliement de l'ARN a besoin du magnésium (Stein et al., 1976a; Misra et al., 2001). Le Mg^{2+} s'attache aux acides nucléiques et par conséquent permet de les replier (Hu et al., 2017). Le nickel est nécessaire pour le repliement de l'uréase (Hu et al., 2017), et aussi pour son activité (Hawtin et al., 1991; Carter et al., 2009; Ge et al., 2013) et par conséquent il a un rôle de virulence chez *Helicobacter pylori*, car cette enzyme permet de coloniser l'estomac en augmentant le pH environnant (Benoit et al., 2013). Le zinc se retrouve quant à lui dans la structure des protéines à doigt de zinc (Chou et al., 1998; Baglivo et al., 2009). Le fer est très impliqué dans la réaction de transfert d'électrons (Ferre-D'Amare et al., 2011); Et il joue un rôle important dans la virulence des bactéries (Palmer et al., 2016). Le manganèse joue un rôle important pour l'activité de la superoxide dismutase (Archibald et al., 1981). Le manganèse peut être remplacé par le fer dans certaines réactions d'oxydoréductases (Martin et al., 2015). Le rôle du manganèse dans la virulence de bactéries a été bien démontré (Ferre-D'Amare et al., 2011; Troxell et al., 2013a,b; Turner et al., 2015; Palmer et al., 2016). Ce sont quelques exemples parmi d'autres pour montrer le rôle et l'importance des métaux.

Le magnésium et le potassium sont présents chez plusieurs bactéries à une concentration intracellulaire de plus de 10 mM tandis que Fe^{2+} , Mn^{2+} et Ca^{2+} sont maintenus dans l'ordre de micromolaire, μM (Finney et al., 2003); La concentration du zinc est maintenue dans l'ordre de femtomolaire, fM (Outten et al., 2001). Il faut savoir que la détermination de la concentration accessible est difficile à déterminer (Chandrangsu et al., 2017). La Figure 1.5 présente différentes concentrations des métaux chez *E. coli* (Spinelli et al., 2008).



5. Figure 1.5 Les concentrations des métaux chez *E.coli* cultivée dans le LB et le milieu minimal avec glucose mesuré par spectromètre de masse couplé à un plasma inductif, ICP-MS.

A) Les nombres d'atome par bactérie. Les valeurs sont des moyennes de trois mesures. Les barres d'erreurs sont l'écart-type (SD). B, C) Le contenu des métaux dans la cellule de bactérie *E. coli*, moles par cellule volume pour les bactéries dans un milieu minimal (MM) ou LB. Les colonnes claires (non remplies) présentent la limite de détection pour les métaux faiblement abondants. Tiré de (Outten et al., 2001).

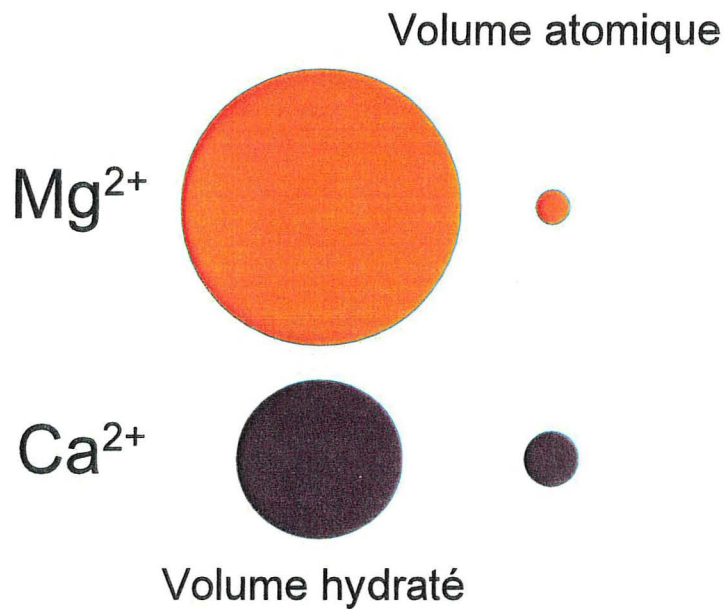
La régulation des métaux est très importante pour tous les êtres vivants y compris les bactéries. La régulation et l'homéostasie des métaux se font par plusieurs couches de régulations y compris des métaloprotéines et des ARNnc (Spinelli et al., 2008).

L'affinité et la spécificité des métalloprotéines et aussi de certains ARNnc (comme certains *riboswitchs*) aux métaux est cruciale, sinon la régulation sera perturbée si un métal est utilisé à la place de l'autre, c'est-à-dire par une *mismetallation* (Waldron et al., 2009).

Pour l'ARN qui interagit avec les métaux, l'affinité spécifique au magnésium est plus énigmatique, car premièrement, la concentration du magnésium est très élevée dans la cellule bactérienne (Finney et al., 2003) par rapport aux autres métaux. Deuxièmement, les ARN ont besoin de Mg^{2+} pour leur repliement (Romer et al., 1975; Stein et al., 1976b; Grilley et al., 2006). Finalement, les ARN sont négativement chargés, ce qui facilite la liaison avec tous les cations. Pour creuser ces notions, certaines propriétés physico-chimiques des ions métalliques pouvant aider à comprendre la régulation et le repliement de l'ARN sont abordées ci-dessous.

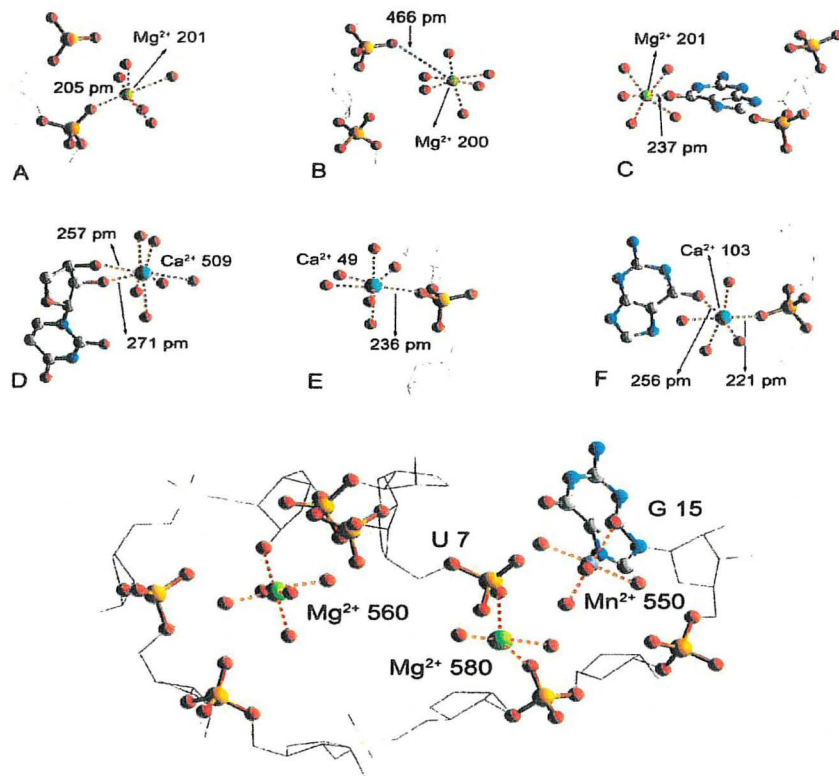
Il est intéressant de savoir que le magnésium a des propriétés physico-chimiques différentes par rapport aux autres métaux (Maguire, 2006). La taille du Mg^{2+} est très petite comparée à d'autres métaux, ce qui permet au magnésium d'avoir une forte liaison électrostatique avec l'eau (Draper, 2004). La liaison hydrogène entre le magnésium et l'eau lui procure le plus grand volume hydraté, environ 400X son volume non-hydraté (le ratio du volume). Cette proportion pour d'autres ions comme le calcium est 25X ou tout juste 4X pour le potassium (Figure 1.6 et Tableau 1.2) (Maguire, 2006). En comparant le numéro de coordination, (le nombre de liaisons possibles pour un atome avec d'autres atomes) du magnésium et d'autres ions comme le calcium, on comprend que le magnésium et le sodium sont hexacoordonnés, mais le calcium peut avoir différentes coordinations 6,7, 8 ou même 9 (Figure 1.7 et Tableau 1.2) (Maguire, 2006). Chez plusieurs eucaryotes, le calcium est une molécule de signalisation et il doit faire partie de différentes modulations, mais le magnésium se lie aux ATP ou aux phosphates des acides nucléiques (Maguire et al., 2002). Contrairement au magnésium, le manganèse a des électrons *d*, qui font une liaison covalente avec l'ARN, qui est plus forte que le magnésium (Draper, 2004).

Comment le *riboswitch* pourrait distinguer entre deux ions divalents dans un environnement qui contient une concentration élevée de magnésium ? Les sections suivantes récapitulent la régulation de certains métaux notamment ceux qui pourraient être régulés par des ARNnc dans un tableau (Tableau 1.3).



6. Figure 1.6 La comparaison des ratios de volume du magnésium et du calcium.

La comparaison du ratio de volume hydraté (**gauche**) vs le volume atomique (**droite**) du magnésium et du calcium. Le ratio du volume hydraté du magnésium vs le volume atomique est 400 fois plus grand tandis que le ratio du volume hydraté du calcium est 25 fois plus grand que son volume non hydraté. Adapté de (Maguire, 2006).



7. Figure 1.7 L'interaction des ions avec l'ARN.

Haut : A-C) L'interaction du magnésium avec l'ARN. (D-F) L'interaction du calcium avec l'ARN. Le numéro de coordination du magnésium est toujours 6 tandis que pour le calcium est 6 et 8. Bas : Mg²⁺ 560, Mg²⁺ 580 et Mn²⁺ 550 interagissent avec U7 et G15. Le magnésium interagit avec les groupements phosphates ou via des molécules d'eaux (groupement OH) ou directement avec la base nucléique. PDB1HEZ (Shi et al., 2000). Manganèse interagit directement avec l'azote de la guanine dans cette position. Des ions portent un numéro venant de PDB. Tiré de (Kolev et al., 2018).

A, B (Correll et al., 1997)

C (Pallan et al., 2005)

D (Deng et al., 2001)

E (Ippolito et al., 1998)

F (Tanaka et al., 2009)

2. Tableau 1.2 Certaines propriétés physico-chimiques des cations.

Ion	Rayon (Å)	Rayon hydraté (Å)	Volume proportion : (hydraté/non hydraté)	Coordination	Pk _a d'ion
Mg²⁺	~0.65	4.3	400X	6	11.42
Ca²⁺	1	4.1	25X	6, 7, 8, ou 9	12.70
Mn²⁺	0.67	ND	ND	6, 7	10.6
Na⁺	1	3.6	25X	6	14.48
K⁺	1.3	3.1	4X	6	?

Adapté de (Feig et al., 1999; Kehres et al., 2003; Maguire, 2006; Bachas et al., 2018).

(ND) : Non défini. (?) : pas retrouvé.

3. Tableau 1.3 Le résumé de régulations des métaux chez les bactéries.

Ion	Rôle	Régulation	Concentration
Fe²⁺	Le cycle de Krebs, la fixation de l'azote, le transfert d'électron (Ferre-D'Amare et al., 2011) et la photosynthèse (Andrews et al., 2003) et les réactions oxydatives (Cornelis et al., 2011). Rôle dans la virulence de bactéries (Palmer et al., 2016).	Protéines : Fur (Fer uptake protein) : Chez <i>E. coli</i> lorsque le fer est en excès, Fur empêche la transcription des gènes producteurs de sidérophores (Bagg et al., 1987). DtxR (Diphtheria toxin repressor) : Chez les bactéries à gram-positif la toxine diphtérique est exprimée lorsque le fer est en carence (Boyd et al., 1990; Tao et al., 1994), ce qui est régulé par DtxR (Schmitt et al., 1991). ARNnc : RyhB : Chez <i>E. coli</i> lorsque le fer est en manque RyhB réprime des gènes dépendants du fer comme des gènes du cycle de Krebs comme la succinate déshydrogénase, l'aconitase et la fumarase (Masse et al., 2005), RyhB est réprimé par Fur lorsque le fer est en excès (Masse et al., 2007; Vecerek et al., 2007).	~0.1mM (Finney et al., 2003); Concentration totale de μM à 0.5 mM (Helmann, 2014). La concentration libre du fer est très faible.
Zn²⁺	Le zinc a plutôt un rôle structural, par exemple des protéines à doigt de zinc (Chou et al., 1998; Baglivo et al., 2009). L'ADN et l'ARN polymérase ont besoin de zinc. Le ribosome est associé au zinc (Waldron et al., 2009). Rôle de virulence <i>via</i> régulation par Zur chez <i>E. coli</i> (Velasco et al., 2018). Facteur de virulence chez <i>Streptococcus pyogenes</i> (Makthal et al., 2017).	Protéines : Zur (zinc uptake repressor) : Chez la plupart des bactéries lorsque le zinc est en excès, Zur empêche la transcription des gènes importateurs du zinc, znuABC (Gaballa et al., 1998; Patzer et al., 1998). En même temps ZntR , un régulateur d'efflux chez <i>E. coli</i> , contrôle plusieurs exportateurs comme ZntA (Capdevila et al., 2016).	~0.1 mM totale (Finney et al., 2003). La concentration libre femtomolaire (fm) (Outten et al., 2001).
Mg²⁺	Rôle de cofacteur pour l'ADN et l'ARN polymérase (Vernon, 1988). Le repliement des acides nucléiques comme l'ARN (Stein et al., 1976a; Misra et al., 2001). L'action pour flagelles de certaines bactéries (Imazawa et al., 2016).	Protéines : Transporteur de magnésium : Chez <i>Salmonella typhimurium</i> CorA , (importateur) (Kehres et al., 2003), mgtA et mgtB (exportateur) (Maguire, 2006). Système à deux composants : PhoP/PhoQ : Chez <i>Salmonella typhimurium</i> (Garcia Vescovi et al.,	1-5 mM concentration libre (Alatossava et al., 1985), plus de 10 mM (Finney et al., 2003), voire 30 mM (Maguire et al., 2002).

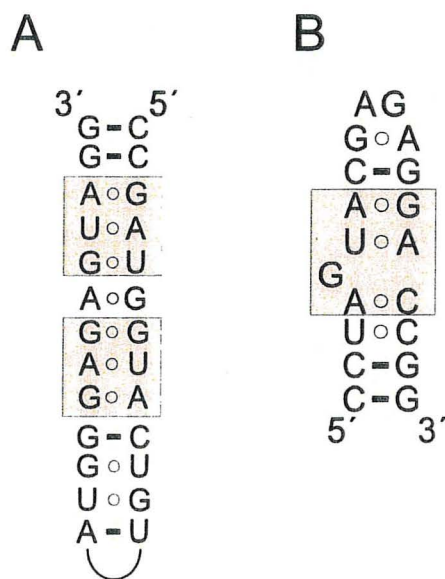
	Rôle de virulence chez certaines bactéries comme <i>Salmonella typhimurium</i> (Garcia Vescovi et al., 1996).	1996). ARNnc : Mg sensor : Chez <i>Salmonella typhimurium</i> se retrouve en amont du gène <i>mgtA</i> (Cromie et al., 2006). ykoK : Chez <i>Bacillus subtilis</i> se retrouve en amont du gène <i>mgtE</i> (Dann et al., 2007).	
Mn²⁺	Le rôle de la protection pendant le stress oxydatif (Turner et al., 2015). Cofacteur de plusieurs enzymes de métabolisme de carbone (Kehres et al., 2003). La dégradation de diguanylate cyclique (c-di-GMP) dépend du magnésium (Papp-Wallace et al., 2006). Fixation d'azote chez certaine bactéries (Hood et al., 2017). Rôle dans la virulence chez certaines bactéries (Juttukonda et al., 2015).	Protéines : MntR : En faible concentration du manganèse, MntR active la transcription de l'opéron <i>mntABCD</i> , qui est un ABC transporteur (Que et al., 2000). tandis que en condition d'excès, MntR réprime l'expression de <i>mntABCD</i> et <i>mntH</i> deux transporteurs de manganèse ; MntR active aussi la transcription de deux pompes d'efflux de manganèse, MneP et MneS en concentration élevée (Huang et al., 2017). Chez <i>E. coli</i> , MntR régule l'expression de la pompe d'efflux, MntP pour le manganèse (Waters et al., 2011). MntX est un exportateur de manganèse chez <i>N. meningitidis</i> (Veyrier et al., 2011). ARNnc : En faible concentration un arrêt de transcription empêche la transcription du gène <i>mntH</i> , tandis qu'en concentration élevée, le gène <i>mntH</i> se transcrit (Shi et al., 2014). En faible concentration chez <i>Lactococcus lactis</i> le <i>riboswitch yybP-ykoY</i> empêche la transcription du gène <i>yoaB</i> (Price et al., 2015). Chez <i>E. coli</i> en concentration élevée le <i>riboswitch yybP-ykoY</i> permet la traduction du gène <i>mntP</i> , qui est un exportateur de manganèse (Dambach et al., 2015).	10-100 µM (Finney et al., 2003). De µM à 0.5 mM (Helmann, 2014). La concentration pourrait atteindre à l'ordre de mM (Papp-Wallace et al., 2006).
Ni²⁺/Co²⁺	nickel : L'activité de l'uréase dépend du nickel chez <i>S. aureus</i> (Hiron et al., 2010) et <i>H. pylori</i> (Benoit et al., 2013). Repliement de l'uréase (Hu et al., 2017). cobalt : Cofacteur d'enzymes telle que méthylmalonyl-CoA, la méthionine – aminopeptidase et la glucose	Protéines : Transporteurs : CorA (Smith et al., 1993; Niegowski et al., 2007). ATP Binding Cassette (Rodionov et al., 2006). NikR : En concentration élevée empêche la transcription de <i>nikABCDE</i> (Chivers et al., 2000), qui code pour ATP Binding Cassette chez <i>E. coli</i> (Navarro et al., 1993). Nika : Faire rentrer en se liant au	Faible abondance (Finney et al., 2003).

	<p>isomérase (Kobayashi et al., 1999), monoxyde de carbone déshydrogénase (Ragsdale et al., 1996), l'uréase et l'hydrolase chez <i>H. pylori</i> (Olson et al., 2001).</p>	<p>nickel <i>via</i> deux canaux transmembranaires, NikB et NikC (de Pina et al., 1995; Mulrooney et al., 2003).</p> <p>NiCoT : Transporteur de nickel et cobalt (Eitinger et al., 2005).</p> <p>RcnR : régule la pompe d'efflux, RcnA en présence de nickel et/ou cobalt (Iwig et al., 2008; Blaha et al., 2011). Le cobalt peut être exporté par la pompe d'efflux CzcDABC (Nies, 1995).</p> <p>ARNnc :</p> <p>Le <i>riboswitch NiCo</i> se retrouve en amont du gène de czcD chez plusieurs espèces (Furukawa et al., 2015).</p>	
Ca ²⁺	<p>La sporulation (Herbaud et al., 1998), division cellulaire (Norris et al., 1996), virulence (Guragain et al., 2013; Broder et al., 2016).</p> <p>Chez certaines bactéries action de flagelles dépend de calcium (Imazawa et al., 2016). La formation de biofilm chez certains pathogènes (Garrison-Schilling et al., 2011). Régulateur pour certaines enzymes comme la gyrase GyrA (Karkare et al., 2012).</p>	<p>Protéines :</p> <p>CaBP (<i>Calcium Binding Proteins</i>) : Des protéines régulatrices ressemblant à la calmoduline (Norris et al., 1996) contenant un ou deux motifs de EF-hand protéine, hélice-boucle-hélice. Il existe d'autres CaBPs moins répandus.</p> <p>L'efflux :</p> <p>Système à deux composants :</p> <p>PhoP/PhoQ : Chez <i>Salmonella typhimurium</i> (Garcia Vescovi et al., 1996).</p> <p>CarSR calcium-regulated Sensor-Regulator Chez <i>Vibrio cholerae</i> (Bilecen et al., 2009). Chez <i>P. aeruginosa</i> PAO1 CarSR considéré comme régulateur de calcium (Guragain et al., 2016). Exportateur comme CaxP chez <i>Streptococcus pneumoniae</i> (Rosch et al., 2008), LmrP chez <i>Lactococcus lactis</i> (Schaedler et al., 2012).</p> <p>L'afflux :</p> <p>Le calcium peut rentrer par polyhydroxybutyrate-polyphosphate (PHB-PP) (Das et al., 2001), des canaux ioniques voltage-dépendant (Bruni et al., 2017).</p>	<p>~0.1mM (Finney et al., 2003). Chez <i>E. coli</i> 90 ± 10 nM (Gangola et al., 1987) ; peut atteindre à 1-2 µM (Naseem et al., 2007). Chez <i>Pseudomonas aeruginosa</i> 100-200 nM (Guragain et al., 2013). La concentration libre :100-300 nM (Dominguez, 2004).</p>

1.5 Le motif boucle E (*loop E*) :

1.5.1 Structure de la boucle E :

Certains ARN sont constitués de certaines boucles internes (*internal loop*), qui séparent la double hélice en deux parties de manière symétrique ou asymétrique (Hendrix et al., 2005). Ces boucles internes peuvent avoir des paires de bases non Watson-Crick (Hendrix et al., 2005). Une boucle qui s'appelle, la boucle E a été rapportée pour la première fois chez l'ARN ribosomal 5S de *E. coli* (Figure 1.8 A) (Fox et al., 1975). Les nucléotides de boucle E forment des paires de bases non Watson-Crick et la conservation absolue n'est pas essentielle pour le repliement tertiaire de l'ARN ribosomal 5S, mais fournit un caractère intrinsèque afin de maintenir la structure des boucles internes *via* des paires de bases non canoniques (Brunel et al., 1991). Ce motif, également nommé Sarcin-Ricin (Figure 1.8 B) ou bulged-G, a la séquence conservée RAGUA, (Endo et al., 1993). La boucle E (*loop E*), a été identifiée comme un centre de liaison de métaux (Correll et al., 1997; Feig et al., 1999).



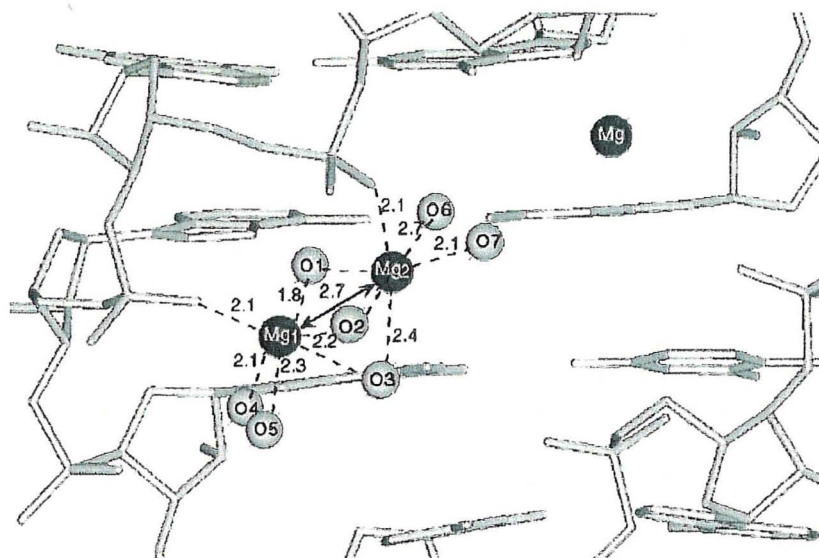
8. Figure 1.8 Le consensus de la boucle E retrouvé chez l'ARN ribosomal 5S et 23S.

Le consensus de la boucle E retrouvé chez l'ARN ribosomal 5S. B) Le motif AGUA est présenté dans la boucle Sarcin/Ricin de l'ARN 23S chez *E. coli*. Adapté de (Leontis et al., 1986; Leontis et al., 1998).

1.5.2 Le rôle de la boucle E et les métaux :

Il a été montré que le repliement de l'ARN ribosomal 5S a besoin d'une certaine concentration du magnésium et il est aussi possible en présence de calcium (Leontis et al., 1986) et plus tard la boucle E a été identifiée comme un centre de liaison de métaux (Correll et al., 1997). En fait, quatre ions de magnésium stabilisent les paires de bases non Watson-Crick (Figure 1.9) (Correll et al., 1997). Deux atomes sur quatre de ces atomes de magnésium sont présentés (Figure 1.9) et deux de ces atomes pourraient être remplacés par d'autres ions dépendamment des propriétés des ions (Feig et al., 1999).

D'ailleurs, une boucle similaire à la boucle E ayant l'affinité pour le zinc a été identifiée (Ciesiolka et al., 1995). Il a aussi été montré que le potassium comme un ion monovalent peut interagir avec les paires de bases non Watson-Crick, mais de manière moins spécifique que le magnésium (Auffinger et al., 2004).



9. Figure 1.9 L'interaction de magnésium avec la boucle E de l'ARNr 5S.

Quatre ions de magnésium stabilisent les paires de bases non Watson-Crick de la boucle E de l'ARN ribosomal 5S, déterminées à 1.5 Å. Deux atomes parmi les quatre observés dans le cristal sont montrés. Tiré de (Feig et al., 1999).

1.6 La prédiction d'ARN non codants :

La prédiction de structures secondaires d'ARN permet de trouver des nouveaux ARNnc avant d'explorer leur fonction (Seetin et al., 2012a) quoique cela requiert beaucoup de temps et d'énergies au point de vue de calcul. La génomique comparative est un outil qui permet de trouver des ARNnc dans n'importe quelle espèce, mais en particulier chez des procaryotes (Rivas et al., 2001a,b) ou des champignons (Li et al., 2017). Plusieurs candidats de *riboswitchs* ont été découverts en utilisant la méthode de génomique comparative (Weinberg et al., 2010; Weinberg et al., 2017). Un algorithme permettant de détecter des covariations au sein d'alignements de séquences a été développé (Rivas et al., 2001a,b), ce qui a permis de trouver des ARNnc chez *E. coli* (Rivas et al., 2001b). Contrairement aux protéines, les séquences primaires d'ARNnc sont souvent peu conservées et par conséquent, l'utilisation de *Basic Local Alignment Search Tool nucleotide* (BLASTn) (Altschul et al., 1990; Camacho et al., 2009) pour trouver l'homologie n'est pas toujours efficace (Nawrocki et al., 2013). En effet, l'alignement de séquences d'ARNnc homologues montre souvent une identité de 60 % ou moins (Gardner et al., 2005).

Les *riboswitchs* se retrouvent dans la région 5'-UTR de bactéries (Breaker, 2011) et par conséquent chercher ces ARNnc dans cette région paraît logique (Barrick et al., 2004). En utilisant BLASTn sur les régions intergéniques chez *B. subtilis* et en comparant la conservation, plusieurs *riboswitchs* ont été découverts (Barrick et al., 2004). Une autre méthode prend tous les gènes homologues et prédit une structure secondaire pour retrouver des nouveaux *riboswitchs* (Yao et al., 2007). Une méthode qui retrouve des séquences homologues et ensuite prédit la structure secondaire a été utilisée pour découvrir des nouveaux *riboswitchs* (Tseng et al., 2009). La méthode de génomique comparative utilise des méthodes de prédiction de la structure secondaire d'ARN.

Plusieurs méthodes de prédiction de la structure secondaire d'ARN seront expliquées ci-dessous. Chaque méthode a ses avantages et ses inconvénients au point de vue de la vitesse de prédiction et de précision de la prédiction. On peut distinguer deux approches différentes pour la prédiction de la structure d'ARN, certaines méthodes sont basées sur l'alignement de séquences tandis que d'autres ne sont pas basées sur l'alignement de séquences (Hamada, 2015).

La prédiction de la structure secondaire d'ARN à partir d'une seule séquence est basée sur le repliement d'ARN ayant l'énergie libre minimum, *MFE* (*Minimum Free energy*) (Mathews et al., 2010). Dans cette méthode la structure d'ARN (ayant des paires de bases)

est en équilibre avec les parties non structurées (Seetin et al., 2012b). Les nucléotides qui font des paires de bases diminuent l'énergie de la structure au point de vue thermodynamique, ΔG ; et les nucléotides qui se retrouvent dans la partie simple brin, boucle ou bulge, augmentent l'énergie de la structure (Lorenz et al., 2016). Le premier algorithme a été développé pour maximiser le nombre de paires de bases dans une séquence (Nussinov et al., 1980). Une version avancée pour la prédiction de structure secondaire d'ARN a été implémentée par *mfold* (Zuker et al., 1981; Zuker, 2003) et *Vienna package* (Hofacker, 2003; Lorenz et al., 2016) Ces deux méthodes, *Vienna package* et *mfold*, utilisent l'énergie libre minimum *MFE*, afin de prédire la structure secondaire (Mathews et al., 2010). Afin de surmonter des problèmes de *MFE* (l'ARN a une seule conformation d'équilibre et des paires de bases voisines diminuent toujours l'énergie minimale) (Seetin et al., 2012b), la fonction de partition (*partition function*) a été utilisé. C'est une sommation de toutes les constantes d'équilibre pour toutes les structures possibles (McCaskill, 1990). La méthode *Vienna package* prend en compte la probabilité de la structure prédite avec *MFE* parmi d'autre structures, (*partition function*) et la probabilité de former des paires de bases dans une structure donnée parmi d'autres structures prédites (Seetin et al., 2012b). La version avancée de *mfold* (Zuker, 2003) fonctionne avec une approche pour chercher des structures « sous optimales ». Dans cette approche plusieurs structures ont été considérées ayant l'énergie libre minimum, *MFE* (Seetin et al., 2012b). Cette approche est utile pour la prédiction de structures d'ARN ayant des structures alternatives comme des *riboswitchs* (Li et al., 2014).

Des méthodes de prédiction de structures basées sur l'apprentissage de machine (ML) ont été aussi développées (Hamada, 2014). Parmi ces méthodes, celles des statistiques comme, *sfold* (Ding et al., 2005), *CONTRAFold* (Do et al., 2006) et *Stochastic context free grammar* (SCFG) (Dowell et al., 2004) ont été considérées dans cette catégorie.

La prédiction de la structure secondaire d'ARN en utilisant une seule séquence ne peut pas être précise. La limitation vient du manque de connaissance sur le modèle d'énergie d'ARN et de la simplification de beaucoup de paramètres de simples brins (Backofen et al., 2014). Afin de surmonter ces problèmes, la prédiction de la structure secondaire basée sur l'alignement de séquences a été proposée (Backofen et al., 2014). La structure secondaire et la structure tertiaire subissent moins de changement au sein d'évolution par rapport à la séquence. Des substitutions de séquences seront compensées afin de préserver la structure et évidemment la fonction de l'ARN (Aigner et al., 2012). Il existe trois approches différentes pour la prédiction de structure secondaire basée sur l'alignement

de séquences. **I)** Aligner des séquences d'abord et ensuite les replier. **II)** Replier et aligner des séquences en même temps. **III)** Replier d'abord et ensuite aligner (Seetin et al., 2012b). Les exemples ci-dessous expliquerons mieux ces trois approches.

I) *RNAalifold* (Hofacker et al., 2002; Bernhart et al., 2008), prend en compte le minimum d'énergie des séquences alignées au lieu d'une seule séquence et ensuite prédit la structure secondaire de cet alignement (Backofen et al., 2014).

II) Parmi des méthodes qui alignent et replient en même temps, l'algorithme de Sankoff est bien connu (Sankoff, 1985). Dans cette nouvelle méthode, des séquences seront repliées comme une seule séquence, pourvu que toutes les contraintes de prédiction de paires de bases, de la *MFE*, soient considérées pour toutes les séquences dans l'alignement (Havgaard et al., 2014). *FOLDalign* (Gorodkin et al., 1997; Havgaard et al., 2007) est un exemple qui maximise des probabilités de paires de bases au sein des séquences, plutôt que l'énergie libre minimale (Seetin et al., 2012b). *LocARNA* (Will et al., 2007) implémente aussi l'algorithme de Sankoff et permet de détecter la structure conservée de l'alignement des séquences. *LocARNA* maximise le nombre de paires de bases possibles entre deux séquences (Seetin et al., 2012b). Dans la même optique que *LocARNA*, *PART* (Harmanci et al., 2008) utilise des méthodes probabilistes comme la fonction de partition, (*partition function*), pour maximiser des nombres de paires de bases (Seetin et al., 2012b).

III) *RNAshapes* (Steffen et al., 2006) produit d'abord une énumération d'une représentation abstraite de l'ARN (*Shapes*) et calcule la probabilité de chaque structure. Ensuite, un calcul thermodynamique permet de choisir la structure la plus probable qui ont cette forme abstraite en commun (Seetin et al., 2012b).

Une approche qui détecte l'ARN fonctionnel a été proposée par la méthode *RNAz* (Washietl et al., 2005), ce qui prend en compte la structure secondaire de l'ARN conservée selon le consensus d'alignement de séquences et puis calcule la stabilité de cette structure selon le point de vue thermodynamique (Washietl et al., 2005). Une autre méthode de prédiction a été implémentée par *CMfinder*, qui prédit la structure d'ARN selon la covariance de séquences (Yao et al., 2006). *CMfinder* permet de détecter des motifs d'ARN avec une faible conservation de séquences (Yao et al., 2007). *CMfinder* pourrait être considéré comme un outil puissant pour retrouver des nouveaux ARNnc, y compris les *riboswitchs* (Ruzzo et al., 2014). Plusieurs méthodes ont été développées basées sur le repliement et l'alignement de la structure (Will et al., 2007; Seetin et al., 2012b; Havgaard et al., 2014).

La plupart de ces méthodes utilisent un nombre de séquences limitées, car en augmentant le nombre de séquences et leurs longueurs, le temps de calcul augmente de

manière exponentielle. Une méthode qui combine plusieurs méthodes de prédiction nommées ci-dessus a été utilisée par *GraphClust* (Heyne et al., 2012). *GraphClust* permet de prendre plusieurs séquences de différentes longueurs et ensuite de les mettre dans les groupes (*clusters*). Contrairement aux autres méthodes nommées auparavant, le temps de calcul est linéaire et donc le nombre et la longueur des séquences ne sont pas limitantes pour la prédiction de la structure secondaire d'ARN.

1.7 Problématiques et hypothèse et objectifs

1.7.1 Problématique :

La découverte de RhyB et d'autres *riboswitchs* pour le Mg^{2+} , Mn^{2+} et Ni^{2+}/Co^{2+} montre qu'il existe aussi un système de régulation des métaux au niveau de l'ARN. Après la découverte de deux *riboswitchs* spécifiques aux ions, Mg^{2+} , nous avons estimé qu'il devrait probablement exister d'autres ARN non codant ou *riboswitchs* spécifiques aux métaux. Actuellement, il existe deux classes pour les *riboswitchs* magnésium, deux classes pour les *riboswitchs* manganèse et aussi une classe de *riboswitch* qui reconnaît Ni^{2+}/Co^{2+} . Ces données suggèrent qu'il y aurait d'autres ARN non codant régulateurs des métaux. Plusieurs questions pourraient se poser ainsi : Comment d'autres métaux pourraient se réguler ? Est-ce qu'il existe d'autres *riboswitchs* pour les ions de transition comme le Cd^{2+} ou des métaux alcalins comme le Ca^{2+} ? Est-ce qu'il y aurait un ARN régulateur pour chaque ion métallique ou existe-t-il un même ARN non codant pour plusieurs différents métaux ? Est-ce qu'il y aurait d'autres classes de *riboswitchs* pour des ions tel que Mg^{2+} , Mn^{2+} et Ni^{2+}/Co^{2+} ?

1.7.2 Hypothèse :

Nous émettons l'hypothèse qu'il existe d'autres *riboswitchs* spécifiques aux métaux qui pourraient moduler certains gènes chez les procaryotes et notre objectif général est de les trouver et éprouver leur authenticité.

1.7.3 Objectifs :

Afin de réaliser ce projet deux objectifs principaux et plusieurs sous-objectifs ont été établis :

- A. Trouver un ou plusieurs *riboswitch(s)* en utilisant des méthodes bioinformatiques de génomique comparative tenant compte des prédictions de structures secondaires de l'ARN.
 - I. Mettre en place une base de données pour faciliter l'extraction des séquences intergéniques en amont de gènes avec des fonctions associées aux métaux.

- II. Utiliser cette base de données en combinant avec d'autres bases de données publiques pour faire de nouvelles recherches sur une fonction particulière, par exemple le rôle de l'ARNnc dans la virulence.
 - III. Établir un « *pipeline* » pour la découverte de nouvelles structures conservées.
- B. Éprouver l'authenticité des ARNnc candidats pour la régulation des métaux, incluant les *riboswitchs* candidats trouvés suite à l'objectif « A », par des méthodes de biologie moléculaire comme « *in-line probing* » et des essais fonctionnels en bactérie.
- I. Chercher et caractériser des ribozymes *hammerhead* agissant potentiellement comme senseurs d'ions métalliques.
 - II. Caractériser des *riboswitchs* ressemblant à *yybP-ykoY* qui serait potentiellement des *riboswitchs* pour le calcium.
 - III. Caractériser de potentiels *riboswitchs* calcium trouvés chez *Sinorhizobium* et *Yersinia/Serratia*.

2. Single mutation in hammerhead ribozyme favors cleavage activity with manganese over magnesium.

La traduction en français :

Une seule mutation dans le cœur catalytique du ribozyme *hammerhead* favorise le manganèse plutôt que le magnésium. .

Auteurs :

Mohammad Reza Naghdi¹, Émilie Boutet¹, Clarisse Mucha¹, Jonathan Ouellet², Jonathan Perreault¹.

Affiliations professionnelles :

¹ Institut National de la Recherche Scientifique (INRS), Centre Armand Frappier Santé Biotechnologie, 531 boul. des Prairies, Laval, Québec, H7V 1B7, Canada ;
jonathan.perreault@inrs.iaf.ca; Tel.: (1-450-687-5010, ext. 4411);

² Department of Chemistry and Physics, Monmouth University, 400 Cedar Avenue, West Long Branch, NJ, 07764, USA

Titre de la revue ou de l'ouvrage :

non-coding RNA

Publié le 20 mars 2020

DOI : <https://doi.org/10.3390/ncrna6010014>

Référence pour l'article :

<https://www.mdpi.com/2311-553X/6/1/14>

Contribution des auteurs :

Le concept de l'article a été imaginé par J. Perreault. Des tests initiaux sur le ribozyme *hammerhead Bcep176* ont été faits par MR. Naghdi, de même que pour les autres ribozymes décrits (avec l'assistance de C. Mucha et J. Ouellet) pendant le premier stage en 2012. De plus, les constructions, la validation, la cinétique en présence des divers ions sur le ribozyme *hammerhead* bactériophage *Bcep176* et toutes les analyses pour déterminer le k_{obs} ont été planifiées et réalisées par MR. Naghdi. Les mutants de *Bcep176* ont été construits par MR. Naghdi et ensuite testés par E. Boutet. De plus E. Boutet a construit le ribozyme

hammerhead type II (variant sauvage, A6 et mutant, A6C) et l'a testé en présence de Mg^{2+} et Mn^{2+} . Toutes les analyses pour déterminer le k_{obs} pour le ribozyme *hammerhead* type II ont été faites par MR Naghdi. La section du ribozyme *hammerhead* type II dans la partie matériels et méthodes a été écrite par E. Boutet, toutes les autres parties de l'article ont été écrites par MR. Naghdi et retravaillées avec J. Perreault. Des modifications et la révision ont été faites par MR. Naghdi, E. Boutet, J. Ouellet et J. Perreault.

Résumé :

Le ribozyme *hammerhead* est un des ARN les plus étudiés au point de vue biochimique et structural. L'activité d'auto-coupeure du ribozyme *hammerhead* dépend des cations, en particulier des cations divalents. Le magnésium étant le cation divalent le plus abondant dans les cellules, il joue un rôle important dans l'activité du ribozyme *hammerhead*, mais d'autres cations tel que le manganèse permettent aussi son activité. Le cœur catalytique du ribozyme *hammerhead* est très conservé. L'importance des nucléotides du cœur catalytique du ribozyme *hammerhead* pour son activité d'auto-coupeure a été largement étudiée et la majorité de ces nucléotides sont considérés invariables. Dans cet article on montre l'effet d'une variation naturelle dans le cœur catalytique du ribozyme *hammerhead Bcep176* qui favorise l'activité d'auto-coupeure en présence de manganèse plutôt que d'autres cations y compris le magnésium.

Telle que présentée dans cet article, la déviation du consensus appelé désormais la « mutation » A6C, qui favorise un ligand (le manganèse) par rapport à d'autres ligands (tel le magnésium), nous permet de jeter un nouveau regard sur l'évolution de motifs d'ARN complexes. En effet, nos résultats pourraient contribuer à mieux comprendre l'évolution de certains ARNnc (tel le ribozyme *hammerhead* et des *riboswitchs*) vis-à-vis les métaux en démontrant qu'une simple mutation ponctuelle peut grandement affecter la préférence pour de tels ligands, ce point sera davantage élaboré dans la partie discussion (chapitre 6) de cette thèse. Ces travaux permettent aussi de reconsidérer l'effet de changements dans le cœur catalytique de l'important modèle d'étude de l'ARN qu'est le ribozyme *hammerhead* dans le contexte de ribozymes complets, c'est-à-dire avec l'interaction entre les tiges 1 et 2, par opposition au ribozyme minimal qui a été plus étudié. Ainsi, non seulement le ribozyme *hammerhead* apparaît-il comme un exemple d'un ARNnc ayant le potentiel de fonctionner comme un senseur d'ions métalliques, mais aussi comme un exemple d'une possible variation de spécificité de ligand métallique tel que décrit dans le chapitre de résultat (chapitre 5) sur le *riboswitch yybP-ykoY* et ses variants.

Abstract :

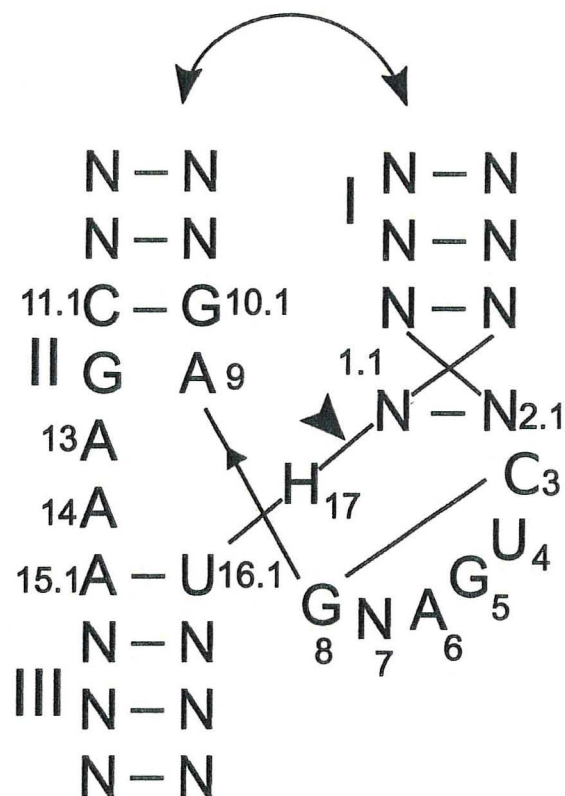
Hammerhead ribozymes are one of the most studied classes of ribozymes so far, from both the structural and biochemical point of views. The activity of most hammerhead ribozymes is cation-dependent. Mg^{2+} is one of the most abundant divalent cations in the cell and therefore plays a major role in cleavage activity for most hammerhead ribozymes. Besides Mg^{2+} , cleavage can also occur in the presence of other cations such as Mn^{2+} . The catalytic core of hammerhead ribozymes is highly conserved, which could contribute to a preference of hammerhead ribozymes toward certain cations. Here, we show a naturally occurring variation in the catalytic core of hammerhead ribozymes, A6C, that can favor one metallic ion, Mn^{2+} , over several other cations.

Keywords : manganese; magnesium; cation; RNA variants; C6 hammerhead variants; A6.

2.1 Introduction

Independent discoveries by the laboratories of Thomas Cech and Sidney Altman, leading to a shared Nobel prize in chemistry in 1989, demonstrated that RNA could catalyze chemical reactions [1–5]. Naturally occurring ribozymes [3–6] include group I [1] and II [7] introns, RNase P RNA [2], spliceosomal RNA [8] and ribosomal RNA [6], as well as small hammerhead [9,10], Varkud satellite (VS) [11], hairpin [12], hepatitis delta virus (HDV) [13] (and HDV-like [14]), twister [15] (and twister sister), pistol, hatchet [16] and *glmS* ribozymes [17]. Hammerhead ribozymes (HHRz) were observed for the first time in the late eighties in tobacco plants [9,10]. The HHRz were the most studied ribozymes for self-cleavage activity, becoming models for research on RNA structure and function [18]. Since then, it has been shown that HHRz are widespread and could be found in all domains of life [19–22].

At physiological pH level, the activity of HHRz depends on metal ions, especially Mg^{2+} [23], which supports cleavage in vitro for a minimal, but sub-optimal, HHRz sequence at 10 mM [24]. Other ions can also activate the self-cleavage of HHRz [25]: cations like ammonium (NH_4^+) can support the activity of HHRz [26] and large tetraalkylammonium ions significantly increase the rate of HHRz in addition to Mg^{2+} [27]. The cleavage rate of HHRz was tested with transition metals and depending on the conditions and ribozymes tested, cleavage with Mn^{2+} showed three times [28] and up to seventy-six times [29] better cleavage than Mg^{2+} . In fact, metal ions like Mn^{2+} bind to specific nucleotides of the catalytic core, such as the phosphate of A9, the nitrogen from G10.1 and the oxygen of G12 [30–33] (Figure 2.1). Nevertheless, the finding that Mn^{2+} bound to hammerhead ribozymes and bound more strongly than Mg^{2+} or K^+ [34] is not surprising given that Mn^{2+} also binds RNA better, in general [35].



10. Figure 2.1 The structure and sequence consensus of the hammerhead ribozyme (HHRz).

The structure and sequence consensus of the hammerhead ribozyme (HHRz). The standard numbering of positions in the catalytic core of HHRz is shown. The cleavage site is indicated by an arrow. H: stands for all the nucleotides except G. The curved arrows illustrate the tertiary interaction between the stems I and II.

The minimal catalytic core of HHRz is made of the core consensus C3U4G5A6NG8A9-G12A13A14 with the A15-U16 base pair and H17 cleavage site surrounded by three helical stems [25] (Figure 2.1), which are necessary for cleavage activity. Nevertheless, some rare variations at certain core positions decrease the cleavage rate in a few natural HHRz, but the ribozymes presumably remain functionally active *in vivo* [20]. Two examples of variants, U(2a)G(2b)U(3)U4G5A6C7G8A9 and G(2a)C(2b)C(3)U4G5A6C7G8A9 from halophilic organisms, were suggested to modulate gene expression according to divalent cation concentrations [20].

We hypothesized that some other HHRz would also be likely to have varying ion specificity. We set our goal to determine first whether a previously identified core variant

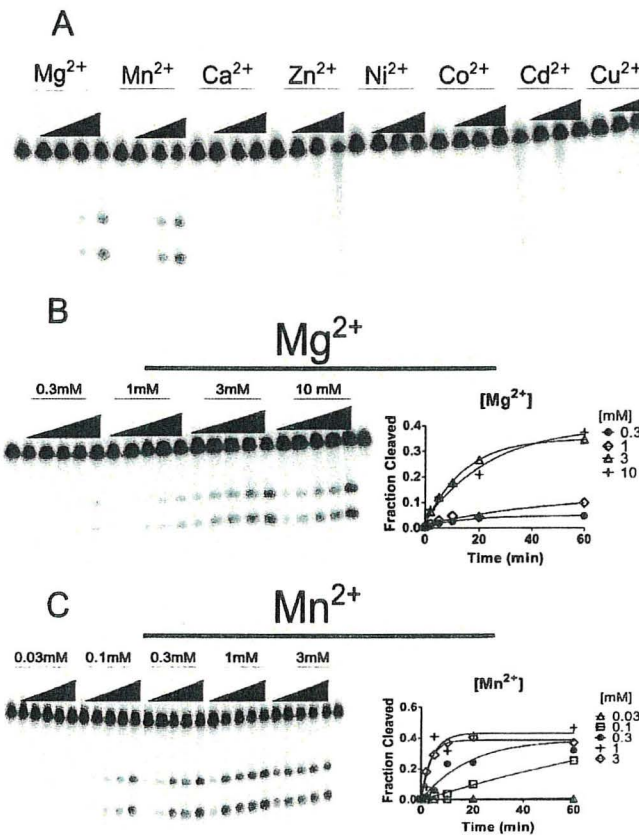
(A6C) from bacteriophage *Bcep176* could have altered cation preferences, and second if this single A6C substitution within the core could alter ion preference for other HHRz. To keep the naming convention clear, the natural variant *Bcep176* will be denoted as *Bcep176* (C6). In this paper, we show that this naturally occurring variation from the typical catalytic core is deleterious for cleavage activity with Mg^{2+} (and other divalent cations), but still allows good cleavage activity with Mn^{2+} .

2.2 Results

2.2.1 Varying metal ion preference of a HHRz variant :

We assayed over a dozen putative ribozymes (selected from [20]) that either had a variant core or gene context suggestive of cation regulation (Tables S1 and S2, l'annexe III). Five were active in our assay conditions, including the *Bcep176* (C6) variant which barely cleaved during transcription, but was active in the presence of Mn^{2+} after purification (Table S1, l'annexe III). We determined how this natural variation (C6) could affect the cleavage of *Bcep176* (C6) in the presence of various ions and was found marked differences between activity in Mg^{2+} and Mn^{2+} . To verify the specificity of *Bcep176* (C6) for metal ions, Mg^{2+} , Mn^{2+} and other metals such as Ca^{2+} , Zn^{2+} , Ni^{2+} , Co^{2+} , Cd^{2+} and Cu^{2+} were tested at 0.01, 0.1 and 1 mM, with the exception of Cu^{2+} , which was tested at 0.01 and 0.1 mM (Figure 2.2 A). Cleavage occurred solely in the presence of either Mg^{2+} or Mn^{2+} . To determine the cleavage activity of RNA *Bcep176* (C6), assays were performed for up to 60 min in the presence of Mg^{2+} at 0.3, 1, 3 and 10 mM; and for Mn^{2+} at 0.01, 0.03, 0.1, 0.3, 1 and 3 mM (Figure 2.2 B, C). The cleavage activity of *Bcep176* (C6) at 0.1 mM Mn^{2+} was observed, whereas no cleavage activity was observable at 0.1 mM Mg^{2+} (Figure 2.2 B, C). The k_{obs} values were calculated as 0.31 min^{-1} and 0.29 min^{-1} at 1 and 3 mM Mn^{2+} , respectively (Figure 2.2 C), whereas k_{obs} values at equivalent Mg^{2+} concentration were calculated as 0.0041 min^{-1} and 0.051 min^{-1} , respectively¹ (Figure 2.2 B).

¹ All data can be found in complementary annex (l'annexe III-C).

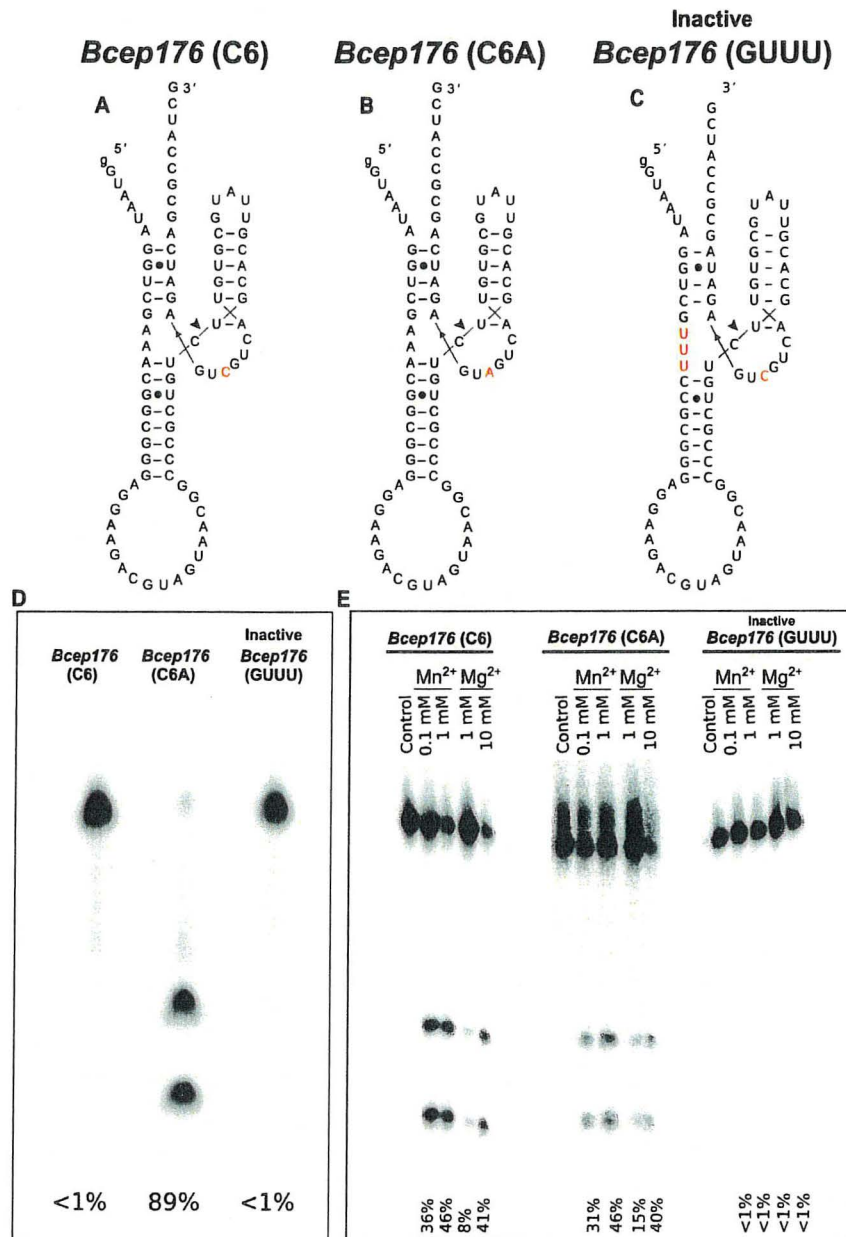


11. Figure 2.2 Cleavage assay for the hammerhead ribozyme *Bcep176* (C6).

A) Cleavage assay of *Bcep176* (C6) in the presence of other metals. For Mg^{2+} , concentrations of 0, 0.01, 0.1, 1 and 10 mM were used. For the other metal ions tested, the concentrations were 0, 0.01, 0.1 and 1 mM. For assays with Cu^{2+} , the concentrations were 0, 0.01 and 0.1 mM. B, C) Cleavage assay of HHRz *Bcep176* (C6) in the presence of indicated concentrations of Mg^{2+} and Mn^{2+} . Incubation times were 2, 5, 10, 20 and 60 min at 37 °C. The first and last wells are negative controls (no divalent cations).

To better understand the role of *Bcep176* (C6) regarding higher activity with Mn^{2+} compared to Mg^{2+} , we did the inverse mutation C6A, leading to *Bcep176* (C6A), reverting to consensus A6 and a negative control mutation of GAAA → GUUU (sequences B and C of Figure 2.3, respectively). Cleavage did not take place in the inactive mutant GAAA → GUUU, as expected. In contrast, for the inverse mutant *Bcep176* (C6A), more efficient cleavage took place, both in the presence of 1 mM Mn^{2+} and during transcription (25 mM Mg^{2+}) (Figure 2.3 D), indicating that the “consensus-like” *Bcep176* (C6A) mutant did not

discriminate between Mg^{2+} and Mn^{2+} . The higher cleavage (36% and 46%) observed in the presence of 0.1 and 1 mM Mn^{2+} , respectively, for native *Bcep176* (C6) suggests that the nucleotide C6 in this WT hammerhead causes a preference for Mn^{2+} over Mg^{2+} (Figure 2.3 E). The fact that *Bcep176* (C6) barely cleaved during transcription (25 mM Mg^{2+}) (Figure 2.3 D), but cleaved to ~40% with 10 mM Mg^{2+} (Figures 2.2 and 2.3 E), may be due to differences of folding during *in vitro* transcription compared to folding after purification and snap cooling.



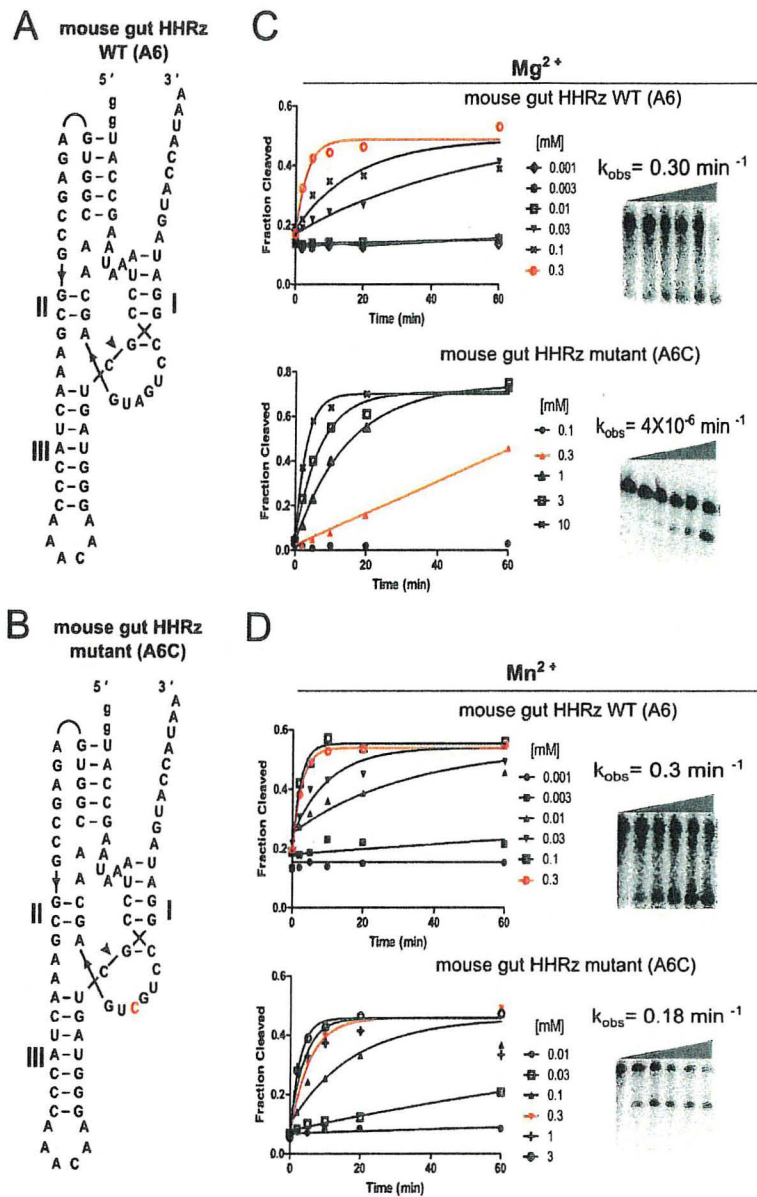
12. Figure 2.3 Importance of C6 in the bacteriophage hammerhead ribozyme *Bcep176*.

A) Native *Bcep176* (C6) bears a natural variation from the consensus catalytic core, C6, instead of A6, shown in red. B) *Bcep176* (C6A) reversed to "standard consensus" (C6A), in red. C) Introducing a mutation to change GAAA to GUUU in order to inactivate the ribozyme *Bcep176*. D) Cleavage during in vitro transcription (in the presence of 25 mM Mg²⁺) for native *Bcep176* (C6), mutant *Bcep176* (C6A) and inactive *Bcep176* (GUUU). E) Cleavage assay of *Bcep176* (C6), mutant *Bcep176* (C6A) and inactive *Bcep176* (GUUU) in the presence of 0.1 mM and 1 mM Mn²⁺ compared to 1 mM and 10 mM Mg²⁺. The incubation time was 60 min for all the assays.

2.2.2 Effect of A6C mutation on another HHRz :

Furthermore, we used a different HHRz with high self-cleavage activity in the presence of Mg^{2+} to explore if an A6C mutation within a consensus HHRz catalytic core would lead to the same phenotype, i.e., cleavage in the presence of Mn^{2+} would be favored over cleavage with Mg^{2+} . The mutated CUGCUGA version of a pseudoknotted type II HHRz derived from mouse gut (hereinafter referred to as mouse gut HHRz) (from [20]) (Figure 2.4A) cleaved better with Mn^{2+} compared to Mg^{2+} , similar to that observed with *Bcep176* (C6). This native mouse gut HHRz (A6) showed high self-cleavage activity during transcription (25 mM Mg^{2+}). The native mouse gut HHRz had a similar cleavage efficiency ($k_{obs} = 0.3 \text{ min}^{-1}$) with Mn^{2+} and Mg^{2+} at 300 μM (Figure 2.4B, C, Table 2.1; Supplementary Figure S1, l'annexe III). However, for the mouse gut HHRz (A6C) mutant, there was a greater than 10,000-fold rate difference between Mn^{2+} and Mg^{2+} at the same ion concentration of 300 μM ($k_{obs} = 0.18 \text{ min}^{-1}$ vs $k_{obs} = 3.55 \times 10^{-6} \text{ min}^{-1}$, respectively)² (Figure 2.4B, C, Table 2.1; Supplementary Figure S1, l'annexe III). It should be noted that even if the C6 mutation favors better cleavage with Mn^{2+} over Mg^{2+} , our data does not necessarily indicate better binding for Mn^{2+} over Mg^{2+} .

² All data can be found in complementary annex (l'annexe III-C).



13. Figure 2.4 Effect of A6C mutation on self-cleavage activity for a pseudoknotted type II consensus core HHRz (A6) derived from the mouse gut metagenome (mouse gut HHRz).

A, B) Sequences and secondary structures of native mouse gut HHRz (A6) versus mutated mouse gut HHRz mutant (A6C), shown in red. C, D) The gels are showing the self-cleavage activity of native mouse gut HHRz (A6) versus mutant HHRz (A6C) in the presence of 300 μM of Mg^{2+} or Mn^{2+} . The graphs correspond to the fraction cleaved for all the concentrations indicated, red curves correspond to 300 μM . The incubation times were 0, 2, 5, 10, 20 and 60 min.

4. Table 2.1 Cleavage rates of *Bcep176* (C6) variant and pseudoknotted type II mouse gut HHRz (mouse gut HHRz) with comparative Mn²⁺ and Mg²⁺ concentrations.

[mM]	<i>Bcep176</i>		mouse gut HHRz			
	WT(C6)		WT(A6)		Mutant(A6C)	
	Mg ²⁺	Mn ²⁺	Mg ²⁺	Mn ²⁺	Mg ²⁺	Mn ²⁺
0.001	ND	ND	ND	<10 ⁻⁶	ND	ND
0.003	ND	ND	1.2x10 ⁻⁵	0.0025	ND	ND
0.01	ND	ND	0.00070	0.031	ND	0.00093
0.03	ND	0.00012	0.019	0.096	ND	0.0070
0.1	ND	0.017	0.056	0.45	<10 ⁻⁶	0.056
0.3	0.0018	0.057	0.30	0.39	3.6x10⁻⁶	0.18
1	0.0041	0.31	ND	ND	0.067	0.24
3	0.051	0.29	ND	ND	0.14	0.37
10	0.041	ND	ND	ND	0.33	ND

Table 2.1. The different k_{obs} obtained are presented for Mg²⁺ and Mn²⁺ (min⁻¹). The red highlights show that at 0.3 mM concentration, Mn²⁺ is favored over Mg²⁺ for *Bcep176* (C6) and pseudoknotted type II HHRz derived from mouse gut (mouse gut HHRz) mutant (A6C), whereas green highlights no significant difference with Mn²⁺ vs Mg²⁺ for standard mouse gut HHRz (A6). ND stands for "not determined" for the indicated concentration³.

³ The standard deviation error and Confidence Interval can be found in complementary annex (l'annexe III-C).

purified by denaturing 6 % PAGE for 2 h. RNA is then eluted and dissolved in 21 μL RNase-free water. 1 μL of this sample is used to determine the concentration of RNA by using a Nanodrop spectrophotometer.

3.4.3 Dephosphorylation and Labeling:

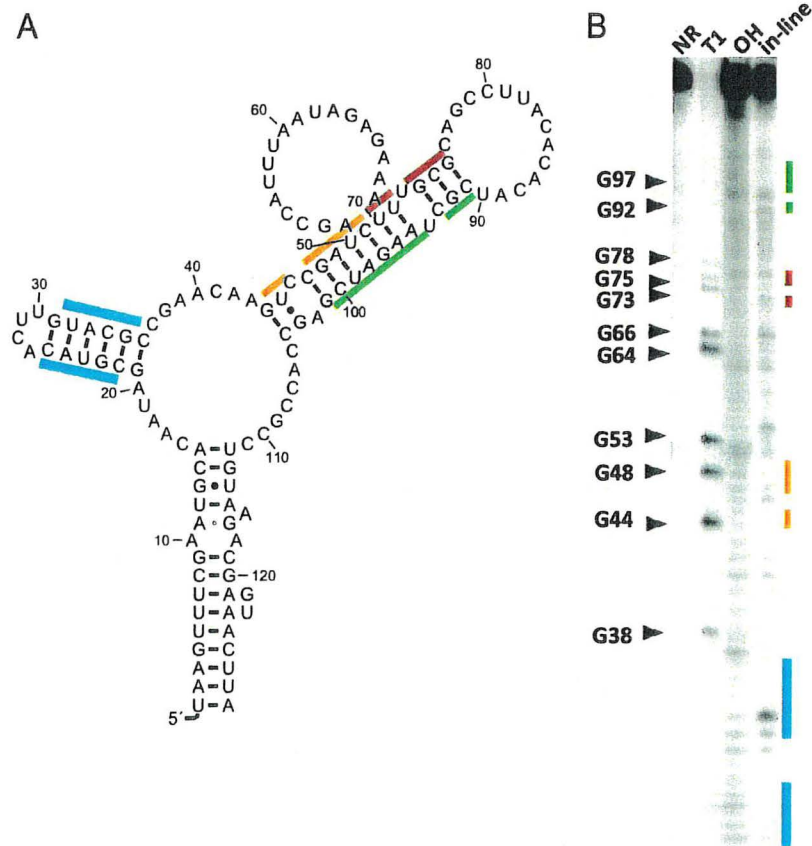
Dephosphorylation of RNA is performed by following the manufacturer's instructions for Antarctic phosphatase (NEB). To label RNA, 2 μL radioactive ATP ($\gamma\text{-}^{32}\text{P}$), 3 to 10 pmoles of dephosphorylated RNA, 1 μL of 10 U/ μL polynucleotide T4 kinase and PNK buffer (NEB) in 20 μL is incubated at 37 °C for 1 h. The labeled product is purified on denaturing 6 % PAGE.

3.4.4 Determination of candidate RNA structure and potential modulation by in-line probing:

To determine the activity of an RNA suspected to be a riboswitch, RNA is incubated in conditions favoring a structure-dependent degradation pattern with in-line probing. In these conditions, different concentrations of ligand can be assayed to test ligand binding and get data necessary for K_D calculation. In-line reactions are carried out for 40 h at room temperature. Standard in-line reactions are 50mM Tris pH 8.3, 100 mM KCl and 20 mM MgCl_2 , but in the case of metal ion ligands, the in-line reaction can be carried with varying concentrations of Mg^{2+} . To be able to determine RNA structure, two types of ladder are prepared. Up to three times as much of labeled RNA can be digested by T1 enzyme and alkaline digestion. T1 reactions are carried out with the radiolabeled RNA and 1.5 μL of T1 RNase 1 U/ μL in T1 solution incubated at 56 °C for 5 minutes. Alkaline digestion is conducted with the RNA being incubated in 20 μL of 1X alkaline solution at 90° C for 1 minute and 20 seconds. The samples are then run on 10 % denaturing PAGE for approximately 3 hours at 70 W, exposed with phosphorimaging screens and scanned by a Typhoon FL9500. The technique has been described in more detail by Regulski and Breaker [15].

In-line probing was carried out on a construct of the IGR upstream of *mnmC* in *E. coli str. JM109* to confirm the structure of the RNA. We prepared a control in-line probing reaction containing a spontaneously digested RNA without metabolite, a no-reaction sample of undigested RNA, RNA subjected to partial digestion by RNase T1, and a partial alkaline

digestion (Na_2CO_3). The labeled RNA was incubated for 35 hours and then the pattern of RNA degradation was examined by denaturing 10 % PAGE (Fig. 3.6 A and B). The small highly conserved stem is not supported by the in-line probing data in this construct. This could be due to inappropriate structure prediction, especially considering that the stem is not supported by covariation. Alternatively, since riboswitches and many types of regulatory RNA elements have at least two mutually exclusive structures, in-line probing (Fig. 3.6 B) may not represent the conserved RNA motif candidate RNA-methyl-28, but rather an alternative structure. An additional example for the Mg^{2+} -II riboswitch is provided (Supplementary material section S4.1, l'annexe IV).



19. Figure 3.6 In-line probing of the RNA-methyl-28 candidate.

A) Secondary structure of the sequence used for the in-line probing experiment. B) In-line probing of the RNA illustrated in B. RNA was incubated for 40 h. The colored lines along the gel correspond to regions indicated with the same colors on the secondary structure representation.

3.5 Discussion

Targeted comparative genomics, where sequences upstream of homologous proteins are aligned to look for conserved RNA structures, have been fruitful in the past [57]. However, these studies were undertaken with data extracted from databases such as “NCBI Nucleotide” using scripts or homemade programs and could only be performed by someone with programming skills. Here we demonstrate the accessibility and usefulness of RiboGap in extracting and exploring intergenic sequences for ncRNA. Beyond its ease of use, RiboGap extends the types of sequence ensembles the user can make by allowing function-based queries, rather than protein domain-based queries, providing additional data useful for downstream analysis. RiboGap can be used on a regular basis by most genomics researchers interested in obtaining results from simple or complex queries. Since RiboGap centralizes data from many different databases, it permits users to optimize their research by querying the combined data from various original databases. This cuts down on the laborious compilation and parsing of multiple sets of data required prior to analysis, associated with drawing information from regular databases.

The extraction of intergenic sequences is an important part of the pipeline, which can lead to *de novo* predictions of ncRNAs with GraphClust. Even if the latter is relatively efficient to compare sequences on a large scale, comparing all IGRs of all sequenced bacteria and archaea would require considerable computing resources, as opposed to targeting limited datasets provided by RiboGap. This targeted strategy focuses on sets of genes, which have a higher likelihood of harboring regulating ncRNAs. Alternatively, the users can choose from multiple sets of functions to explore less obvious associations, which may link more subtle regulatory mechanisms related to the ncRNA candidate structures. In both cases, using RiboGap can greatly reduce computing time as well as the number of candidate ncRNAs to evaluate, which is even more time consuming. Perhaps even more critical, choosing which candidate ncRNA to study from the large number of putative ncRNAs requires the analysis of countless alignments and structures, many of which might be interesting, but most of which would not be.

While RiboGap is a powerful tool for extracting intergenic sequences associated with chosen gene functions, it is limited by gene annotations. Poorly annotated genes can either prevent the user from getting a set of sequences corresponding to the chosen function, or lead to the prediction of ncRNAs associated with another unrelated function. Here, we show an example of the latter. Intergenic sequences associated with genes annotated as “urea carboxylases” were searched for the presence of known ncRNAs. As

expected, this led us to find *ykkC* and mini-*ykkC* guanidine riboswitches [31-33]. In this case, the annotation of *uca* is likely wrong due to a lack of knowledge regarding guanidine biology, resulting in missannotation of *uca* as encoding urea decarboxylase enzyme, rather than a guanidine decarboxylase.

Our initial screen of magnesium-related genes only identified members of one of the two known Mg^{2+} riboswitches. We thus adjusted parameters of BLASTClust from 90 % to 98 % to find the Mg^{2+} -I (*ykoK*) riboswitch as we had previously done with our RNA-methyl-28 search. Analysis of the GraphClust results should be performed with precaution as none of the currently available software can comprehensively predict the existing RNA secondary structures. Yet, because GraphClust uses both CMfinder and LocARNA, it benefits from different covariation model predictions and alignments of secondary structure instead of merely sequence. Finally, one should not forget that even though some very powerful tools are available for ncRNA prediction, the inspection of the candidate motifs is necessary to appropriately evaluate them and decide which ones to prioritize for experimental validation.

3.6 Acknowledgements

J.P. thanks support from Natural Sciences and Engineering Council of Canada (NSERC) [418240 to J.P.]. J.P. is a junior 1 FRQS research scholar. R.R.B. is supported by the NIH (GM022778) and by the Howard Hughes Medical Institute. The authors wish to thank J. Lajoie, V. Korniakova, R. Walsh and E. Boutet for helpful discussions. Computations and data extraction were made on the supercomputer Mammouth parallèle 2, managed by Calcul Québec and Compute Canada (funded by CFI, NanoQuébec, RMGA and FRQ-NT).

3.7 References

- [1] R. Backofen, F. Amman, F. Costa, S. Findeiss, A.S. Richter, P.F. Stadler, Bioinformatics of prokaryotic RNAs, *RNA biology* 11(5) (2014).
- [2] R.R. Breaker, Prospects for riboswitch discovery and analysis, *Molecular cell* 43(6) (2011) 867-79.
- [3] M. Gossringer, R.K. Hartmann, 3'-UTRs as a source of regulatory RNAs in bacteria, *The EMBO journal* 31(20) (2012) 3958-60.
- [4] A. Serganov, E. Nudler, A decade of riboswitches, *Cell* 152(1-2) (2013) 17-24.
- [5] E. Masse, N. Majdalani, S. Gottesman, Regulatory roles for small RNAs in bacteria, *Current opinion in microbiology* 6(2) (2003) 120-4.
- [6] T. Mizuno, Regulation of gene expression by a small RNA transcript (micRNA), *Tanpakushitsu kakusan koso. Protein, nucleic acid, enzyme* 29(11) (1984) 908-13.
- [7] M.M. Meyer, *The role of mRNA structure in bacterial translational regulation*, Wiley interdisciplinary reviews. RNA (2016).

- [8] M.T. Morita, Y. Tanaka, T.S. Kodama, Y. Kyogoku, H. Yanagi, T. Yura, Translational induction of heat shock transcription factor sigma32 : evidence for a built-in RNA thermosensor, *Genes & development* 13(6) (1999) 655-65.
- [9] A. Nahvi, N. Sudarsan, M.S. Ebert, X. Zou, K.L. Brown, R.R. Breaker, Genetic control by a metabolite binding mRNA, *Chemistry & biology* 9(9) (2002) 1043.
- [10] W.C. Winkler, A. Nahvi, N. Sudarsan, J.E. Barrick, R.R. Breaker, An mRNA structure that controls gene expression by binding S-adenosylmethionine, *Nature structural biology* 10(9) (2003) 701-7.
- [11] D.A. Benson, K. Clark, I. Karsch-Mizrachi, D.J. Lipman, J. Ostell, E.W. Sayers, GenBank, *Nucleic acids research* 43(Database issue) (2015) D30-5.
- [12] F. Cunningham, M.R. Amode, D. Barrell, K. Beal, K. Billis, S. Brent, D. Carvalho-Silva, P. Clapham, G. Coates, S. Fitzgerald, L. Gil, C.G. Giron, L. Gordon, T. Hourlier, S.E. Hunt, S.H. Janacek, N. Johnson, T. Juettemann, A.K. Kahari, S. Keenan, F.J. Martin, T. Maurel, W. McLaren, D.N. Murphy, R. Nag, B. Overduin, A. Parker, M. Patricio, E. Perry, M. Pignatelli, H.S. Riat, D. Sheppard, K. Taylor, A. Thormann, A. Vullo, S.P. Wilder, A. Zadissa, B.L. Aken, E. Birney, J. Harrow, R. Kinsella, M. Muffato, M. Ruffier, S.M. Searle, G. Spudich, S.J. Trevanion, A. Yates, D.R. Zerbino, P. Flicek, *Ensembl 2015, Nucleic acids research* 43(Database issue) (2015) D662-9.
- [13] M.G. Seetin, D.H. Mathews, RNA structure prediction : an overview of methods, *Methods Mol Biol* 905 (2012) 99-122.
- [14] S. Heyne, F. Costa, D. Rose, R. Backofen, GraphClust : alignment-free structural clustering of local RNA secondary structures, *Bioinformatics* 28(12) (2012) i224-32.
- [15] E.E. Regulski, R.R. Breaker, In-line probing analysis of riboswitches, *Methods Mol Biol* 419 (2008) 53-67.
- [16] P.J. Farnham, T. Platt, Rho-independent termination : dyad symmetry in DNA causes RNA polymerase to pause during transcription in vitro, *Nucleic acids research* 9(3) (1981) 563-77.
- [17] K.S. Wilson, P.H. von Hippel, Transcription termination at intrinsic terminators : the role of the RNA hairpin, *Proceedings of the National Academy of Sciences of the United States of America* 92(19) (1995) 8793-7.
- [18] S. Okuda, A.C. Yoshizawa, ODB : a database for operon organizations, 2011 update, *Nucleic acids research* 39(Database issue) (2011) D552-5.
- [19] T. Tatusova, S. Ciufu, B. Fedorov, K. O'Neill, I. Tolstoy, RefSeq microbial genomes database : new representation and annotation strategy, *Nucleic acids research* 43(7) (2015) 3872.
- [20] E.P. Nawrocki, S.R. Eddy, Infernal 1.1 : 100-fold faster RNA homology searches, *Bioinformatics* 29(22) (2013) 2933-5.
- [21] S. Will, K. Reiche, I.L. Hofacker, P.F. Stadler, R. Backofen, Inferring noncoding RNA families and classes by means of genome-scale structure-based clustering, *PLoS computational biology* 3(4) (2007) e65.
- [22] R. Lorenz, S.H. Bernhart, C. Honer Zu Siederdisen, H. Tafer, C. Flamm, P.F. Stadler, I.L. Hofacker, ViennaRNA Package 2.0, *Algorithms Mol Biol* 6 (2011) 26.
- [23] A.R. Gruber, S. Findeiss, S. Washietl, I.L. Hofacker, P.F. Stadler, RNAz 2.0 : improved noncoding RNA detection, *Pac Symp Biocomput* (2010) 69-79.
- [24] E.P. Nawrocki, D.L. Kolbe, S.R. Eddy, Infernal 1.0 : inference of RNA alignments, *Bioinformatics* 25(10) (2009) 1335-7.
- [25] Z. Yao, Z. Weinberg, W.L. Ruzzo, CMfinder--a covariance model based RNA motif finding algorithm, *Bioinformatics* 22(4) (2006) 445-52.
- [26] S. Janssen, R. Giegerich, The RNA shapes studio, *Bioinformatics* 31(3) (2015) 423-5.
- [27] C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, T.L. Madden, BLAST+ : architecture and applications, *BMC bioinformatics* 10 (2009) 421.
- [28] S. Griffiths-Jones, RALEE--RNA ALignment editor in Emacs, *Bioinformatics* 21(2) (2005) 257-9.

- [29] Z. Weinberg, R.R. Breaker, R2R--software to speed the depiction of aesthetic consensus RNA secondary structures, *BMC bioinformatics* 12 (2011) 3.
- [30] T.L. Bailey, M. Boden, F.A. Buske, M. Frith, C.E. Grant, L. Clementi, J. Ren, W.W. Li, W.S. Noble, MEME SUITE : tools for motif discovery and searching, *Nucleic acids research* 37(Web Server issue) (2009) W202-8.
- [31] M.E. Sherlock, R.R. Breaker, Biochemical Validation of a Third Guanidine Riboswitch Class in Bacteria, *Biochemistry* (2017).
- [32] R.R. Breaker, R.M. Atilho, S.N. Malkowski, J.W. Nelson, M.E. Sherlock, The Biology of Free Guanidine As Revealed by Riboswitches, *Biochemistry* (2017).
- [33] M.E. Sherlock, S.N. Malkowski, R.R. Breaker, Biochemical Validation of a Second Guanidine Riboswitch Class in Bacteria, *Biochemistry* (2017).
- [34] R.K. Montange, R.T. Batey, Structure of the S-adenosylmethionine riboswitch regulatory mRNA element, *Nature* 441(7097) (2006) 1172-5.
- [35] J.X. Wang, R.R. Breaker, Riboswitches that sense S-adenosylmethionine and S-adenosylhomocysteine, *Biochem Cell Biol* 86(2) (2008) 157-68.
- [36] P. Gollnick, P. Babitzke, A. Antson, C. Yanofsky, Complexity in regulation of tryptophan biosynthesis in *Bacillus subtilis*, *Annual review of genetics* 39 (2005) 47-68.
- [37] C. Valverde, M. Lindell, E.G. Wagner, D. Haas, A repeated GGA motif is critical for the activity and stability of the riboregulator RsmY of *Pseudomonas fluorescens*, *J Biol Chem* 279(24) (2004) 25066-74.
- [38] K. Lapouge, M. Schubert, F.H. Allain, D. Haas, Gac/Rsm signal transduction pathway of gamma-proteobacteria : from RNA recognition to regulation of social behaviour, *Mol Microbiol* 67(2) (2008) 241-53.
- [39] F. Jean-Pierre, J. Perreault, E. Deziel, Complex autoregulation of the post-transcriptional regulator RsmA in *Pseudomonas aeruginosa*, *Microbiology* 161(9) (2015) 1889-96.
- [40] T. Romeo, C.A. Vakulskas, P. Babitzke, Post-transcriptional regulation on a global scale : form and function of Csr/Rsm systems, *Environmental microbiology* 15(2) (2013) 313-24.
- [41] P.B. Lodato, P.K. Hsieh, J.G. Belasco, J.B. Kaper, The ribosome binding site of a mini-ORF protects a T3SS mRNA from degradation by RNase E, *Molecular microbiology* 86(5) (2012) 1167-82.
- [42] N. Mathy, L. Benard, O. Pellegrini, R. Daou, T. Wen, C. Condon, 5'-to-3' exonuclease activity in bacteria : role of RNase J1 in rRNA maturation and 5' stability of mRNA, *Cell* 129(4) (2007) 681-92.
- [43] H. Qiao, N. Lu, E. Du, L. Yao, H. Xiao, S. Lu, Y. Qi, Rare codons in uORFs of baculovirus p13 gene modulates downstream gene expression, *Virus research* 155(1) (2011) 249-53.
- [44] T.L. Bailey, C. Elkan, Fitting a mixture model by expectation maximization to discover motifs in biopolymers, *Proc Int Conf Intell Syst Mol Biol* 2 (1994) 28-36.
- [45] V. de Lorenzo, F. Giovannini, M. Herrero, J.B. Neilands, Metal ion regulation of gene expression. Fur repressor-operator interaction at the promoter region of the aerobactin system of pColV-K30, *Journal of molecular biology* 203(4) (1988) 875-84.
- [46] M.M. Meyer, T.D. Ames, D.P. Smith, Z. Weinberg, M.S. Schwalbach, S.J. Giovannoni, R.R. Breaker, Identification of candidate structured RNAs in the marine organism 'Candidatus Pelagibacter ubique', *BMC genomics* 10 (2009) 268.
- [47] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, *Journal of molecular biology* 215(3) (1990) 403-10.
- [48] E.P. Nawrocki, S.W. Burge, A. Bateman, J. Daub, R.Y. Eberhardt, S.R. Eddy, E.W. Floden, P.P. Gardner, T.A. Jones, J. Tate, R.D. Finn, Rfam 12.0 : updates to the RNA families database, *Nucleic acids research* 43(Database issue) (2015) D130-7.
- [49] Z.Z. Yao, Z. Weinberg, W.L. Ruzzo, CMfinder – a covariance model based RNA motif finding algorithm, *Bioinformatics* 22(4) (2006) 445-452.
- [50] G.E. Fox, C.R. Woese, 5S RNA secondary structure, *Nature* 256(5517) (1975) 505-7.

- [51] S. Will, T. Joshi, I.L. Hofacker, P.F. Stadler, R. Backofen, LocARNA-P : accurate boundary prediction and improved detection of structural RNAs, *RNA* 18(5) (2012) 900-14.
- [52] K.A. Curry, C.S. Tomich, Effect of ribosome binding site on gene expression in *Escherichia coli*, *DNA* 7(3) (1988) 173-9.
- [53] H. Chen, M. Bjerknes, R. Kumar, E. Jay, Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs, *Nucleic acids research* 22(23) (1994) 4953-7.
- [54] A. El Korbi, J. Ouellet, M.R. Naghdi, J. Perreault, Finding instances of riboswitches and ribozymes by homology search of structured RNA with Infernal, *Methods Mol Biol* 1103 (2014) 113-26.
- [55] J.N. Kim, A. Roth, R.R. Breaker, Guanine riboswitch variants from *Mesoplasma florum* selectively recognize 2'-deoxyguanosine, *Proceedings of the National Academy of Sciences of the United States of America* 104(41) (2007) 16092-7.
- [56] S. Tian, J.D. Yesselman, P. Cordero, R. Das, Primerize : automated primer assembly for transcribing non-coding RNA domains, *Nucleic acids research* 43(W1) (2015) W522-6.
- [57] Z. Weinberg, J.E. Barrick, Z. Yao, A. Roth, J.N. Kim, J. Gore, J.X. Wang, E.R. Lee, K.F. Block, N. Sudarsan, S. Neph, M. Tompa, W.L. Ruzzo, R.R. Breaker, Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline, *Nucleic acids research* 35(14) (2007) 4809-19.
- [58] G.E. Crooks, G. Hon, J.M. Chandonia, S.E. Brenner, WebLogo : a sequence logo generator, *Genome research* 14(6) (2004) 1188-90.
- [59] S.H. Bernhart, I.L. Hofacker, S. Will, A.R. Gruber, P.F. Stadler, RNAalifold : improved consensus structure prediction for RNA alignments, *BMC bioinformatics* 9 (2008) 474.

4. A survey of *cis* regulatory non-coding RNAs involved in bacterial virulence.

La traduction en français :

Le rôle des ARN non codants (ARNnc) régulateurs en *cis* dans la virulence de bactéries.

Auteurs :

Mohammad Reza Naghdi¹, Katia Smail¹, Samia Djerroud¹ and Jonathan Perreault¹.

Affiliations professionnelles :

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

Titre de la revue ou de l'ouvrage :

Frontiers

Received : date ; Accepted : date ; Published : date :

Contribution des auteurs :

La partie bioinformatique y compris les scripts pour chercher des gènes ou locus_tag correspondant à ceux de la base de données PATRIC et d'autres bases de données, les scripts pour le tri et pour chercher des résultats de BLAST, des terminateurs « Rho indépendant » ont été faits par MR. Naghdi. Au-delà de travaux mentionnés, beaucoup de révisions et de vérifications par rapport aux résultats obtenus ont été faites par MR. Naghdi. Les Northern blots ont été faits par MR. Naghdi. La majorité des analyses et des tables ont été faites et préparées par MR. Naghdi. Les ARNt ont été cherchés par K. Smail. La rédaction de l'article a été faite par MR. Naghdi et K. Smail. La mise-à-jour de RiboGap a été faite par S. Djerroud. J. Perreault a participé à la conceptualisation du projet et la révision de l'article.

Résumé :

L'étude de la pathogenèse chez les bactéries est importante pour trouver de nouvelles cibles de médicaments pour traiter les infections bactériennes. Les bactéries pathogènes, y compris les opportunistes, expriment de nombreux gènes dits de virulence pour échapper aux défenses naturelles de l'hôte et au système immunitaire. L'hôte en revanche a développé plusieurs méthodes afin d'éliminer les pathogènes ; par exemple, restreindre l'accessibilité des pathogènes aux métaux comme le magnésium. De ce fait les bactéries doivent répondre rapidement aux changements exercés par l'hôte et les gènes de virulences ont besoin d'être régulés de façon précise et étroite.

Telle régulation peut être obtenue par des ARN cis-régulateurs, comme les *riboswitchs* ou les thermorégulateurs. En dépit des centaines de familles d'ARN annotées comme régulatrices en *cis*, il existe chez les bactéries relativement peu d'exemples d'ARNnc dans les Régions 5' – non-traduites (*UnTranslated*, UTR) décrits pour réguler les gènes de virulence en aval. Un exemple bien étudié est le contrôle de l'expression de *PrfA* par au moins deux ARNnc différents, un thermorégulateur et un *riboswitch* SAM (S-adénosylméthionine). Il n'existe toutefois pas de liste exhaustive pour les gènes de virulence régulés par des ARNnc. Pour réévaluer les rôles potentiels de tels éléments régulateurs dans la pathogenèse bactérienne, nous avons collecté et présenté les gènes de virulence à partir de différentes bases de données et évalué la présence d'ARNnc dans leurs UTR.

Afin de retrouver des ARNnc en aval des gènes de virulence, les données de tous les gènes ayant un ARNnc dans leurs séquences intergéniques (en provenance de RiboGap) et les annotations de gènes de virulence (en provenance de la base de donnée PATRIC) ont été combinées.

À date, cette liste est la plus exhaustive pour tenter d'établir des liens entre des ARNnc et la régulation de gènes de virulence, soit directement soit en fournissant des indices sur les facteurs susceptibles d'activer ou réprimer ces gènes. Puisque des atténuateurs et certains *riboswitchs* modulent l'expression du gène en aval avec l'arrêt de transcription « Rho indépendant », ainsi retrouver ce motif dans une région intergénique permet d'évaluer préalablement l'existence de *riboswitchs* ou d'atténuateurs.

RiboGap nous a permis aussi d'évaluer l'existence de terminateurs « Rho indépendant » en amont de nombreux gènes de virulence, ce qui suggère l'existence de *riboswitchs* ou d'atténuateurs. Ces résultats obtenus à partir de 86300 gènes associés à la virulence et les

ARNnc *en cis* (des motifs qui régulent le transcrit du gène en aval) ont été mises au point dans le texte de l'article.

Bien qu'à notre connaissance cette liste d'ARNnc associés à des gènes de virulence soit la plus complète disponible, elle n'est pas absolument exhaustive. Notamment, les hits obtenus vis-à-vis des régulateurs ARN associés aux métaux sont restreints. Par exemple nous n'avons retrouvé aucun hit pour le *riboswitch yybP-ykoY*, régulateur du transport de manganèse, alors que le rôle du manganèse dans la virulence a été bien démontré. Nous avons néanmoins retrouvé plusieurs hits pour les *riboswitchs* magnésium et Ni/Co, pour le motif ARN TerC et un petit ARN lié à la régulation du zinc, ces métaux sont impliqués dans la virulence.

Dans un premier temps, cet article a pour but de montrer un exemple un exemple de l'utilité de RiboGap pour combiner des informations de différentes provenances, dans ce cas-ci pour établir des liens entre des ARNnc régulateurs et la virulence. Les résultats obtenus fournissent des renseignements importants pour les ARNnc impliqués dans la virulence et pourraient, Dans un deuxième temps ces résultats permettent aux chercheurs intéressés par certains de ces gènes de virulence d'explorer de nouvelles avenues en lien avec le type de régulation suggéré par les ARNnc en amont.

Abstract :

Study of pathogenesis in bacteria is important to find new drug targets to treat bacterial infections. Pathogenic bacteria, including opportunists, express numerous so-called virulence genes to escape the host natural defenses and immune system. Regulation of virulence genes is often required for bacteria to infect their host. Such regulation can be achieved by *cis-regulatory* RNAs, like the metabolite-binding riboswitches or thermoregulators. In spite of the hundreds of RNA families annotated as *cis-regulatory*, there are relatively few examples of ncRNAs in 5'-UnTranslated Regions (UTRs) of bacteria described to regulate downstream virulence genes. To reassess the potential roles of such regulatory elements in bacterial pathogenesis, we collected and present genes important for virulence from different databases and evaluated the presence of ncRNAs in their UTRs.

4.1 Introduction

Virulence factors (VFs) are in large part responsible of the relative pathogenicity of bacteria and often counter the host immune system (Casadevall and Pirofski, 2003). The expression of VFs is controlled in accordance to environmental cues and signaling changes from the host (DiRita et al., 2000), or changes like temperature increase while moving from water to a mammalian host, or signals of microbiota living in the host, to give only a few examples (Bäumler and Sperandio, 2016). The communication between the host and microorganisms is known as inter-kingdom signaling and is crucial for activation of VFs which activates cascades of signaling pathways (Hughes and Sperandio, 2008). For example, it has been shown that sensing a specific metabolite such as fucose in intestines is crucial for effective colonization (Pacheco et al., 2012). The importance of two component regulatory systems has also been shown, like PhoPQ in Mg^{2+} sensing (García Véscovi et al., 1996, Véscovi et al., 1997) or QseC, a quorum-sensing regulator which detects eukaryote hormones for VF regulation (Clarke et al., 2006, Rasko et al., 2008).

Tight regulation of VFs improves the ability of pathogens to infect their host (Caldelari et al., 2013). Quick regulation is a key for successful pathogenicity, which is expected in a highly changing environment like hosts, especially when the latter react to the invasion (Fris and Murphy, 2016). The regulation with RNA has been shown to be more effective than regulation with proteins in some contexts (Gripenland et al., 2010). Noncoding RNAs (ncRNAs) are a heterogeneous group of RNA that do not code for proteins but instead directly enact a function, often related to gene control. Regulation carried out by ncRNAs can impact one or several genes during transcription or translation (Eddy, 2001). The ncRNAs can be divided into two major groups as *cis*-regulatory or non *cis*-regulatory (*trans*) RNA. *Cis*-regulatory RNAs are located mostly in the 5'-UnTranslated Regions (UTRs) of genes and have a direct effect on expression of the downstream gene. Regulation can occur at the transcriptional or translational level (Abduljalil, 2018). Good examples of *cis*-regulatory elements are riboswitches (Nahvi et al., 2002), RNA thermosensors (Morita et al., 1999) and T-boxes (Grundy and Henkin, 1993). For example, It has been shown that several VFs in *Listeria monocytogenes* are controlled by *cis*-regulatory RNAs such as riboswitches and thermoregulator RNAs (Lebreton and Cossart, 2017). The latter regulate the *prfA* gene (Johansson et al., 2002), a major regulator of virulence in *L. monocytogenes* (Leimeister-Wächterchter et al., 1990, Mengaud et al., 1991). *L. monocytogenes* also uses different flavors of a SAM riboswitch to regulate its virulence

(Loh et al., 2009). Here we evaluated *cis*-regulatory RNAs involved in regulation of virulence genes (VFs).

Increased use of RNAseq and improved bioinformatics prediction pipelines revealed numerous ncRNAs (Weinberg et al., 2010, Weinberg et al., 2017, Leonard et al., 2019, Stav et al., 2019) many of which we find may have an impact on virulence and may help expose more regulatory roles of ncRNAs *in vivo* (Hör et al., 2018). We used an exhaustive list of ncRNAs to explore their association to known virulence genes by combining data from different relational databases to provide a collection of *cis*-regulatory ncRNAs regulating VFs to assist research on virulence regulation.

4.2 Materials and methods

4.2.1 Bioinformatics :

In order to determine the existence of all *cis*-regulatory elements in front of virulence genes, an exhaustive list of VFs was established by using PATRIC database (Wattam et al., 2014) and two other databases : VFDB (Chen et al., 2016) and Victors (Sayers et al., 2019). The virulence is defined as a capacity of a bacteria to infect a host by using the VFs which help the bacteria colonize (Sharma et al., 2017) and escape the host immune system which results in infection and disease (Mao et al., 2015). The VFs are the genes, which do not affect the viability of bacteria by losing their activities (Brown et al., 2012).

Next, we looked for all *cis*-regulatory RNAs upstream of these genes by using the RiboGap database (Naghdi et al., 2017) and BLASTp (Altschul et al., 1990, Camacho et al., 2009). All the bacterial intergenic 5'-UTR sequences having ncRNAs (484136 sequences) were extracted with their corresponding cds using RiboGap, i.e. all prokaryotic ncRNAs as annotated in Rfam, as well as a few more RNAs (Supplementary Data, l'annexe V). BLASTp was then used to determine homology between genes downstream of ncRNAs and the list of VFs (9319 genes). PERL scripts (supplementary data, l'annexe V) were used to analyze the results obtained by BLASTp. To avoid getting genes with common domains, but that are non-orthologous, the BLASTp condition was set to 98 % coverage for High Scoring Pair (HSP). The BLASTp result was then sorted to keep hits with at least 40 % identity. Only *cis*-regulatory RNAs on the same strand as the downstream gene were taken into account, except for tRNAs. Since the main role of tRNAs in cells is not gene regulation but rather protein synthesis, we have considered them as separate entities and we wonder to figure

out their role in VFs regulation and more stringent criteria were used to consider them for their potential role in VF regulation.

4.2.2 tRNA searches :

The tRNAs were searched for separately. RNA distance from start codon of VF was also taken in consideration. The search was carried out for the same virulence genes as described above. Because all genomes harbor many tRNAs, numerous genes are expected to have tRNAs upstream of their coding sequences just by chance, so samples of genes (three replicas of 100 randomly chosen genes) were also used to put results in context. To evaluate the presence of pseudo-tRNAs and also obtain information on tRNA identity, tRNAscan-SE (Chan and Lowe, 2019) was used instead of RiboGap annotations, but RiboGap was used to fetch all the UTRs.

4.2.3 Northern of co-transcribed tRNAs :

Three tRNAs were identified upstream of "Elongation Factor Tu" (*Ef-Tu*) gene in *Neisseria*. To determine whether this gene is transcribed alone or co-transcribed with the tRNAs upstream, we selected *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Neisseria sicca* and *Neisseria elongata*. Oligonucleotides complementary to each of the three tRNAs upstream of the *EF-Tu* cds and to the cds itself were ordered from IDT to probe the membrane (Table S1, l'annexe V). Similarly, an oligonucleotide complementary to the tmRNA for the four *Neisseria* species was ordered from IDT and used as control.

Total RNA of *Neisseria gonorrhoeae*, *Neisseria meningitidis* MC58_NMB0124, *Neisseria sicca* and *Neisseria elongata* was migrated on a 6 % polyacrylamide gel and then transferred onto nitrocellulose membrane (Amersham Hybond™ N⁺ from GE healthcare). The oligonucleotides were labeled in 5' by using 5 pmoles of oligonucleotide, 2 μL ATP (γ-³²-P), 1 μL of 10 U/μL polynucleotide T4 kinase and PNK buffer (NEB) in 20 μL, then incubated at 37 °C for 1 h. The labeled products were then purified on denaturing 6 % polyacrylamide gel. The labeled oligonucleotides were incubated for 24 hours at 42° C with the membrane by using the hybridization buffer SCC 5X prepared form SCC 20 X (175,3 g NaCl, 88, 2 g sodium citrate in 800 ml, pH 7,0 with NaOH volume to 1 L, autoclaved) at the rotating oven and the day after washed twice with (SCC 2X, 1 % SDS) and SCC (0.2 X, 0.1 % SDS) respectively (Perreault et al., 2011). Membranes were then exposed overnight on a phosphorimaging plate. The plate was scanned with a Typhoon FLA9500.

4.3 Result

4.3.1 *Cis*-regulatory RNA distribution upstream of virulence factors :

We decided to not limit our search to the genes listed as VFs in the PATRIC, Victors and VFBD databases. The focus of these databases is on experimentally validated VF genes, but can omit their orthologs in other pathogens. We even extended our survey of *cis*-regulatory ncRNAs to non-pathogens because regulation of a gene in such species can still be informative for their VF orthologs in pathogenic counterparts, or between different pathogens. For example, a gene encoding a magnesium-translocating P-type ATPase is regulated by the *ykoK* Mg²⁺ riboswitch in the non-pathogen *Lactococcus lactis* (Dann et al., 2007), by the Mg-sensor riboswitch in the pathogen *Salmonella enterica* (Cromie et al., 2006) or by the MgtC leader in *Klebsiella aerogenes* (Table 1 and S2, S3, l'annexe V) as well as the PhoP/PhoQ two component system (García Véscovi et al., 1996, Cromie et al., 2006), all indicative of a common regulatory signal, in spite of different mechanisms. Therefore, extending searches to VF orthologs may provide hints on the regulation of these genes in pathogens.

We found 86300 genes associated with virulence (and orthologs) downstream of ncRNAs (supplementary data S3, l'annexe V). From these, *cis*-regulatory RNAs (as annotated "type" in Rfam) were selected out of all ncRNAs based on the criteria described in materials and methods. This list includes 18 riboswitches for metabolites, 16 thermoregulatory RNAs and 6 cation-associated regulators (Table 1), as well as many additional ncRNAs such as the histidine leader or T-boxes (4162 hits in the latter case). The purine riboswitch was found to be the most common riboswitch among *cis*-regulatory RNAs (5790 instance), whereas AdoCbl was found just one time and many riboswitches, such as THF, guanidine (I, II and III) and fluoride riboswitches, were not observed with any genes associated with virulence. Among the cation associated ncRNAs, the most common RNA family is associated to a zinc metalloproteinase, *Listeria snRNA rli51*, followed by the Mg²⁺ riboswitch and *ykoK*. Thermoregulators are clearly important ncRNAs regarding pathogenesis as 16 families of such ncRNAs are found upstream of a total of 7021 instances of VFs (or VF homologs). The most abundant thermoregulator identified is *cspA* followed by TrxA-5'UTR and *prfA*-UTR.

6. Table 4.1 Virulence factors (VFs) with *cis-regulatory* RNAs

RNA	Hits	Genes	Most abundant virulence gene
Riboswitches (metabolites)			
ydaO/yuaA leader (c-di-AMP)	163	9	resuscitation-promoting factor RpfA
Cyclic di-GMP-I	13	4	peptide chain release factor 3
Cyclic di-GMP-II	18	2	type II secretion system protein E
M. florum (deoxyguanosine)	193	2	glutamine-hydrolyzing GMP synthase
FMN	222	3	riboflavin synthase
TPP	567	7	energy-coupled thiamine transporter ThiT
ZMP/ZTP	693	8	bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase PurH
Lysine	1185	8	amino acid permease
Glycine	2241	9	glycine dehydrogenase (aminomethyl-transferring)
Purine	5790	8	glutamine-hydrolyzing GMP synthase
SAM riboswitch	62	2	homoserine O-succinyltransferase
SAM-I/IV variant riboswitch	2	1	homoserine O-succinyltransferase
SAM (S box)	1506	~20	methionine import ATP-binding protein MetN
SAH	36	2	methionine synthase
Cobalamin	1819	28	cobyrinate a, c-diamide synthase
AdoCbl	1	1	anaerobic ribonucleoside-triphosphate reductase
Moco (molybdenum cofactor)	243	4	GTP 3',8-cyclase MoaA
Cations			
Mg ²⁺ Sensor	87	1	Mg ²⁺ -translocating P-type ATPase
ykoK	261	5	Mg ²⁺ -translocating P-type ATPase
MgtC leader	207	2	Mg ²⁺ -translocating P-type ATPase
NiCo	4	2	heavy metal translocating P-type ATPase
terC	2	2	Hypothetical protein
Listeria snRNA rli51	566	1	zinc metalloproteinase PrtA
Thermoregulators			
PrfA UTR	959	1	listeriolysin transcriptional regulator PrfA
ToxT 5' UTR	2	1	pilus/toxin transcriptional regulator ToxT
CnfY 5' UTR	6	1	DUF4765 domain-containing protein
RhIA 5' UTR ROSE like	7	2	rhamnosyltransferase 1 subunit A
KatA 5' UTR	30	1	Catalase
SodC 5' UTR	36	1	superoxide dismutase
LasI 5' UTR ROSE like	87	1	acyl-homoserine-lactone synthase
OppA 5' UTR	88	3	oligopeptide ABC transporter substrate-binding protein OppA
shuA/chuA 5' UTR	145	3	TonB-dependent hemoglobin/transferrin/lactoferrin family receptor
PepN 5' UTR	88	1	aminopeptidase N
HtrA 5' UTR	524	2	periplasmic serine endoprotease DegP
TrxA 5' UTR	1035	2	thioredoxin TrxA
ROSE element	841	2	heat-shock protein IbpA
LcrF	318	1	transcriptional regulator
cspA	2257	16	cold-shock protein CspC
AiiA 5' UTR	598	1	attachment protein

4.3.2 tRNAs upstream of virulence factors :

We observed tRNAs upstream of hundreds of genes. Interestingly, several VFs have pseudo-tRNAs in their UTR, such as *clpP* encoding a protease ; and numerous genes have tRNA sequences on the opposite polarity in their 5'-UTR, like the *mr* encoded ribonuclease for many *Betaproteobacteria* species (Table S4, l'annexe V). We tried to determine whether *Ef-Tu* was co-transcribed with the three tRNAs observed upstream of its cds in several *Neisseria* strains. Indeed, the tRNA closest to *Ef-Tu* is at only 46 bases from its start codon. The northern blot result reveals the three tRNAs (Tyr, Gly, Thr) are co-transcribed (supplementary Figure S1, l'annexe V), but co-transcription with *Ef-Tu* was not apparent.

4.3.3 Rho-independent transcription terminators (RiTt) :

We evaluated existing "Rho-independent" terminators (RiTt) for VFs as we did for other ncRNAs. These results can recapitulate several instances of RiTt deemed responsible of riboregulation as determined by Term-seq (Dar et al., 2016) (supplementary data Table S5, l'annexe V). For example, *rli51* is a *cis*-regulatory RNA in *L. monocytogenes* that we highlight as associated to VFs and where we predict a RiTt, which is corroborated by the Term-seq results of Dar et al. 2016. This illustrates the usefulness of RiboGap to rapidly gather information on intergenic sequences and infer hypotheses that can be evaluated.

4.4 Discussion

To our knowledge, the results obtained from RiboGap and virulence databases are the most exhaustive results to date for known VFs related to *cis*-regulatory RNAs. Several hits are already well known as ncRNAs that regulate virulence factors (eg. thermoregulator RNA and *prfA* gene) (Johansson et al., 2002). The importance of thermoregulator RNAs in virulence of *Yersinia* sp (Nuss et al., 2017) and other species like *Shigella* sp. and pathogenic *E. coli* (Heroven et al., 2017) has been discussed by others. The prevalence of thermoregulator RNAs is high, as this is an excellent way for bacteria that can live in soil or water to determine they have moved to a hot-blooded animal.

Similarly, it is known that metals play important roles in virulence (Papp-Wallace and Maguire, 2006, Broder et al., 2016, Guragain et al., 2016, Imazawa et al., 2016, Palmer and Skaar, 2016, Wedekind et al., 2017). It is thus also not surprising that the list includes more than one thousand ncRNAs, which appear connected to metal cations, in one way or another. The role of Mg^{2+} in virulence was previously connected to a Mg^{2+} riboswitch (Cromie et al., 2006, Dann et al., 2007) and plays a major role in the pathogenicity of *Salmonella enterica* serovar Typhimurium (Groisman et al., 2006, Ramesh and Winkler, 2010, Groisman et al., 2013). The role of virulence for other cations such as nickel (Benoit et al., 2013), cobalt (Kersey et al., 2012, Remy et al., 2013), calcium (Sarkisova et al., 2005, Guragain et al., 2013, Sarkisova et al., 2014, Dar et al., 2016, Guragain et al., 2016, Hay et al., 2017), manganese (Boyer et al., 2002, Papp-Wallace and Maguire, 2006, Shi et al., 2014, Juttukonda and Skaar, 2015) or zinc (Dintilhac et al., 1997, Corbett et al., 2012, Mastropasqua et al., 2018, Velasco et al., 2018) has been demonstrated as well. We have not found the manganese riboswitch (*yybP-ykoY*) (Barrick et al., 2004, Dambach et al., 2015, Price et al., 2015), in spite of several known links between Mn^{2+} and virulence. Conversely, in addition to the few Ni-Co riboswitches (Furukawa et al., 2015) from the list, we have found some genes involved with nickel and/or cobalt transport associated to the riboswitch cobalamin and TPP, respectively (Supplementary Data, l'annexe V). Likewise, ZTP-sensing has been previously associated to Zn homeostasis (Nies, 2019). Moreover, there is one hit for an Mg^{2+} ATPase C transporter found in association with the CspA

thermoregulator (Supplementary Data, l'annexe V). These examples are indicative that genes are not always regulated by the most obvious signals, further emphasizing the importance of this compendium of VF-associated ncRNAs.

Second messengers are often involved in regulation of virulence of bacteria (Hall et al. 2018) and several second messengers have been shown to be sensed by riboswitches in the last decade including : cyclic-di-GMP (Sudarsan et al., 2008), cyclic-di-AMP (Nelson et al., 2013), cyclic-GAMP (Nelson et al., 2015) and ppGpp (Sherlock et al., 2018). The cyclic-di-GMP (I and II) riboswitches are known to regulate several VFs (Tamayo, 2019) which have been found in our searches, as well as other genes not recognized as VFs in PATRIC such as : *gbpA*, a characterized colonization factor from *Vibrio cholerae* (Sudarsan et al., 2008, Kariisa et al., 2016) ; the collagen adhesion protein from the well-known insect-killing bacteria, *Bacillus thuringiensis* (Tang et al., 2016) ; or several putative virulence genes from *Clostridiodes difficile* (Abt et al., 2016). Also, cyclic-di-GMP is known to influence bacterial behavior with regards to motility or formation of biofilm, which can impact virulence (Ha and O'Toole, 2015, Valentini and Filloux, 2016), but many of the genes involved in these processes are not necessarily VFs. Thus, while we tried to be as thorough as possible, clearly the list of thousands of instances of VFs and orthologs putatively regulated by ncRNAs should not be considered as absolutely exhaustive. Other possibilities not yet annotated may also exist, e.g. while no guanidine riboswitches have been found in our search, we could presume that for bacteria which cause infection in the urinary track, guanidine riboswitches would be a good way of determining they have reached this site, and thus express relevant VFs, since guanidine is present at much higher concentration in urine (Wishart et al., 2007, Wishart et al., 2009, Durantou et al., 2012, Wishart et al., 2013, Wishart et al., 2018). Several RNA motifs known to be involved VF regulation were not included in the present study because their annotation is deficient. Perhaps the best example for this is the RNA motif bound by the CsrA/RsmA proteins, which have a major impact on virulence (Vakulskas et al., 2015). This motif is composed of a stem-loop with a single stranded "GGA" in the loop and it is usually found in tandem where one of the two loops overlaps the ribosome binding site (RBS) (Valverde et al., 2004, Lapouge et al., 2008) (Curry and Tomich, 1988, Chen et al., 1994). While our list includes the Two-AYGGAY (RF01731) family, which most likely corresponds to a subset of the 5'-UTRs targeted by CsrA/RsmA, hundreds of targets are known for these

proteins (Kulkarni et al., 2014) and the binding motif consensus appears relatively relaxed, making it more difficult to annotate with a high degree of confidence. Other examples of RNA-binding protein affecting VFs exist, such as the TRAP complex which binds ~10 repeats of (U/G) AG within one UTR (Gollnick et al., 2005), regulating genes such as *trpE* (encoding an anthranilate synthase, already shown to be regulated by TRAP (Gollnick et al., 2005), as well as *cna3* (encoding a collagen adhesin) in *Streptococcus gallolyticus* but which we merely predicted by pattern matching (Naghdi et al., 2017) and was not confirmed experimentally as a TRAP target. We thus avoided this type of motifs for our compilation to avoid spurious annotations as much as possible.

One of the ncRNAs that was searched independently was tRNA. Many VFs on the list exhibit presence of tRNAs very close to their coding sequence (less than 30 nt). While we could not show by Northern blot that *Ef-Tu* is indeed co-transcribed with these tRNAs, they are still likely to be, given the short distance of only 46 bases separating them from the AUG. The rate of processing of the tRNAs might be too fast to permit detection of a transcript including the tRNAs together with *Ef-Tu*. In fact, co-transcription was previously observed in *E. coli* (Miyajima et al., 1981) and the proximity of tRNAs to *Ef-Tu* was already noticed in several species (Cousineau et al., 1992), which we find is generalized to numerous bacteria, whether they are *Proteobacteria*, *Bacteroidetes* or *Firmicutes*. Presumed co-transcription of *Ef-Tu* with these tRNAs could suggest potential regulation by tRNA or merely co-regulation due to the use of the same promoter. This is further supported by the absence of predicted promoters, between the tRNA closest to *Ef-Tu* and the start codon, as well as by the presence of a few promoters upstream of the three tRNA sequences, promoters which would thus also be responsible of *Ef-Tu* expression. Several roles beyond the transfer of amino acids have been demonstrated for tRNAs or fragments of tRNAs (Ryckelynck et al., 2005, Raina and Ibba, 2014, Fricker et al., 2019). The tRNA sequences found on the opposite strand, if co-transcribed with the gene downstream, could potentially be targeted by tRNAs (or tRNA fragments) in a way analogous to many sRNAs. Also, some of the machinery involved in tRNA processing and modification is known to act on mRNAs and affect their expression, like the NSun2 tRNA methylase (Zhang et al., 2012) which may imply that tRNA sequences (and pseudo-tRNAs) found in UTRs could act as substrates for such modification and processing. Furthermore, several viruses use

tRNA-like motifs, either for their replication or to initiate translation in eucaryotes (Skuzeski et al., 1996, Hacker and Kaper, 2000, Zeenko et al., 2002). Finally, many bacteriophages use tRNA sequences to integrate in bacterial chromosomes, making it likely to find tRNAs in proximity to pathogenicity islands and related mobile elements (Hacker and Kaper, 2000) and implying that these tRNAs may have a critical role in horizontal gene transfer and the evolution of virulence. Yet, the role of tRNAs upstream of VFs, if any, remains to be elucidated in most cases.

Bacteria respond to signals coming from the host and its immune system. Such signal can be simple and yet present an acute change in the bacterial environment, like the change in temperature when entering a host, to which bacteria need to respond very quickly. Regulation by ncRNAs is very fast and less energetically demanding compared to regulation by protein. Discovering more ncRNAs involved in VF regulation helps better understand the means of bacteria to escape the host immune system, as well as potential targets to overcome bacteria pathogenicity as a promising way for treatment (Mulhbacher et al., 2010).

4.5 Acknowledgement

We acknowledge Dr. Frédéric Verrier (INRS-IAF) for the kind gift of *Neisseria* RNA extracts and Emilie Boutet for her help with the Northern blot technique.

4.6 Funding

Natural Sciences and Engineering Council of Canada (NSERC) [418240 to J.P.]. J.P. is a junior FRQS scholar. KS received fellowships from the Fondation Armand-Frappier and NSERC.

4.7 References

- Abduljalil, J. (2018). Bacterial riboswitches and RNA thermometers: Nature and contributions to pathogenesis. *Noncoding RNA Res* 3, 54-63.
- Abt, M., McKenney, P. and Pamer, E. (2016). *Clostridium difficile* colitis: pathogenesis and host defence. *Nat. Rev. Microbiol.* 14, 609-620.
- Altschul, S., Gish, W., Miller, W., Myers, E. and Lipman, D. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410.
- Barrick, J., Corbino, K., Winkler, W., Nahvi, A., Mandal, M., Collins, J. et al. (2004). New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6421-6426.
- Bäumler, A. and Sperandio, V. (2016). Interactions between the microbiota and pathogenic bacteria in the gut. *Nature* 535, 85-93. doi: 10.1038/nature18849.
- Benoit, S., Miller, E. and Maier, R. (2013). *Helicobacter pylori* stores nickel to aid its host colonization. *Infect. Immun.* 81, 580-584.
- Boyer, E., Bergevin, I., Malo, D., Gros, P. and Cellier, M. (2002). Acquisition of Mn(II) in addition to Fe(II) is required for full virulence of *Salmonella enterica* serovar Typhimurium. *Infect. Immun.* 70, 6032-6042.
- Broder, U., Jaeger, T. and Jenal, U. (2016). LadS is a calcium-responsive kinase that induces acute-to-chronic virulence switch in *Pseudomonas aeruginosa*. *Nat Microbiol* 2, 16184.
- Brown, S., Cornforth, D. and Mideo, N. (2012). Evolution of virulence in opportunistic pathogens: generalism, plasticity, and control. *Trends Microbiol* 20, 336-342. doi: 10.1016/j.tim.2012.04.005.
- Caldelari, I., Chao, Y., Romby, P. and Vogel, J. (2013). RNA-mediated regulation in pathogenic bacteria. *Cold Spring Harb Perspect Med* 3, a010298. doi: 10.1101/cshperspect.a010298.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10, 421.
- Casadevall, A. and Pirofski, L. (2003). The damage-response framework of microbial pathogenesis. *Nat Rev Microbiol* 1, 17-24. doi: 10.1038/nrmicro732.
- Chan, P. and Lowe, T. (2019). tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences. *Methods Mol. Biol.* 1962, 1-14.
- Chen, H., Bjerknes, M., Kumar, R. and Jay, E. (1994). Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs. *Nucleic Acids Res.* 22, 4953-4957.
- Chen, L., Zheng, D., Liu, B., Yang, J. and Jin, Q. (2016). VFDB 2016: hierarchical and refined dataset for big data analysis--10 years on. *Nucleic Acids Res* 44, D694-D697. doi: 10.1093/nar/gkv1239.
- Clarke, M., Hughes, D., Zhu, C., Boedeker, E. and Sperandio, V. (2006). The QseC sensor kinase: a bacterial adrenergic receptor. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10420-10425.
- Corbett, D., Wang, J., Schuler, S., Lopez-Castejon, G., Glenn, S., Brough, D. et al. (2012). Two zinc uptake systems contribute to the full virulence of *Listeria monocytogenes* during growth in vitro and in vivo. *Infect. Immun.* 80, 14-21.
- Cousineau, B., Cerpa, C., Lefebvre, J. and Cedergren, R. (1992). The sequence of the gene encoding elongation factor Tu from *Chlamydia trachomatis* compared with those of

other organisms. *Gene* 120, 33-41.

Cromie, M., Shi, Y., Latifi, T. and Groisman, E. (2006). An RNA sensor for intracellular Mg²⁺. *Cell* 125, 71-84.

Curry, K. and Tomich, C. (1988). Effect of ribosome binding site on gene expression in *Escherichia coli*. *DNA* 7, 173-179.

Dambach, M., Sandoval, M., Updegrove, T., Anantharaman, V., Aravind, L., Waters, L. et al. (2015). The ubiquitous yybP-ykoY riboswitch is a manganese-responsive regulatory element. *Mol. Cell* 57, 1099-1109.

Dann, C., Wakeman, C., Sieling, C., Baker, S., Irnov, I. and Winkler, W. (2007). Structure and mechanism of a metal-sensing regulatory RNA. *Cell* 130, 878-892.

Dar, D., Shamir, M., Mellin, J., Koutero, M., Stern-Ginossar, N., Cossart, P. et al. (2016). Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria. *Science* 352, aad9822.

Dintilhac, A., Alloing, G., Granadel, C. and Claverys, J. (1997). Competence and virulence of *Streptococcus pneumoniae*: Adc and PsaA mutants exhibit a requirement for Zn and Mn resulting from inactivation of putative ABC metal permeases. *Mol. Microbiol.* 25, 727-739.

DiRita, V., Engleberg, N., Heath, A., Miller, A., Crawford, J. and Yu, R. (2000). Virulence gene regulation inside and outside.. *Philos Trans R Soc Lond B Biol Sci* 355, 657-665. doi: 10.1098/rstb.2000.0606.

Durantou, F., Cohen, G., De Smet, R., Rodriguez, M., Jankowski, J., Vanholder, R. et al. (2012). Normal and pathologic concentrations of uremic toxins. *J. Am. Soc. Nephrol.* 23, 1258-1270.

Eddy, S. (2001). Non-coding RNA genes and the modern RNA world.. *Nat Rev Genet* 2, 919-929. doi: 10.1038/35103511.

Fricke, R., Brogli, R., Luidalepp, H., Wyss, L., Fasnacht, M., Joss, O. et al. (2019). A tRNA half modulates translation as stress response in *Trypanosoma brucei*. *Nat Commun* 10, 118.

Fris, M. and Murphy, E. (2016). Riboregulators: Fine-Tuning Virulence in *Shigella*.. *Front Cell Infect Microbiol* 6, 2. doi: 10.3389/fcimb.2016.00002.

Furukawa, K., Ramesh, A., Zhou, Z., Weinberg, Z., Vallery, T., Winkler, W. et al. (2015). Bacterial riboswitches cooperatively bind Ni²⁺ or Co²⁺ ions and control expression of heavy metal transporters. *Mol. Cell* 57, 1088-1098.

García Vescovi, E., Soncini, F. and Groisman, E. (1996). Mg²⁺ as an extracellular signal: environmental regulation of *Salmonella* virulence.. *Cell* 84, 165-174.

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

Gripenland, J., Netterling, S., Loh, E., Tiensuu, T., Toledo-Arana, A. and Johansson, J. (2010). RNAs: regulators of bacterial virulence.. *Nat Rev Microbiol* 8, 857-866. doi: 10.1038/nrmicro2457.

Groisman, E., Cromie, M., Shi, Y. and Latifi, T. (2006). A Mg²⁺-responding RNA that controls the expression of a Mg²⁺ transporter.. *Cold Spring Harb Symp Quant Biol* 71, 251-258. doi: 10.1101/sqb.2006.71.005.

Groisman, E., Hollands, K., Kriner, M., Lee, E., Park, S. and Pontes, M. (2013). Bacterial Mg²⁺ homeostasis, transport, and virulence.. *Annu Rev Genet* 47, 625-646. doi: 10.1146/annurev-genet-051313-051025.

Grundy, F. and Henkin, T. (1993). tRNA as a positive regulator of transcription antitermination in *B. subtilis*.. *Cell* 74, 475-482.

Guragain, M., King, M., Williamson, K., Perez-Osorio, A., Akiyama, T., Khanam, S. et al.

(2016). The *Pseudomonas aeruginosa* PAO1 Two-Component Regulator CarSR Regulates Calcium Homeostasis and Calcium-Induced Virulence Factor Production through Its Regulatory Targets CarO and CarP. *J. Bacteriol.* 198, 951-963.

Guragain, M., Lenaburg, D., Moore, F., Reutlinger, I. and Patrauchan, M. (2013). Calcium homeostasis in *Pseudomonas aeruginosa* requires multiple transporters and modulates swarming motility. *Cell Calcium* 54, 350-361.

Ha, D. and O'Toole, G. (2015). c-di-GMP and its Effects on Biofilm Formation and Dispersion: a *Pseudomonas Aeruginosa* Review. *Microbiol Spectr* 3, 0003-2014.

Hacker, J. and Kaper, J. (2000). Pathogenicity islands and the evolution of microbes. *Annu. Rev. Microbiol.* 54, 641-679.

Hay, A., Yang, M., Xia, X., Liu, Z., Hammons, J., Fenical, W. et al. (2017). Calcium Enhances Bile Salt-Dependent Virulence Activation in *Vibrio cholerae*. *Infect. Immun.* 85, Heroven, A., Nuss, A. and Dersch, P. (2017). RNA-based mechanisms of virulence control in Enterobacteriaceae. *RNA Biol* 14, 471-487.

Hör, J., Gorski, S. and Vogel, J. (2018). Bacterial RNA Biology on a Genome Scale.. *Mol Cell* 70, 785-799. doi: 10.1016/j.molcel.2017.12.023.

Hughes, D. and Sperandio, V. (2008). Inter-kingdom signalling: communication between bacteria and their hosts.. *Nat Rev Microbiol* 6, 111-120. doi: 10.1038/nrmicro1836.

Imazawa, R., Takahashi, Y., Aoki, W., Sano, M. and Ito, M. (2016). A novel type bacterial flagellar motor that can use divalent cations as a coupling ion.. *Sci Rep* 6, 19773. doi: 10.1038/srep19773.

Johansson, J., Mandin, P., Renzoni, A., Chiaruttini, C., Springer, M. and Cossart, P. (2002). An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*.. *Cell* 110, 551-561.

Juttukonda, L. and Skaar, E. (2015). Manganese homeostasis and utilization in pathogenic bacteria. *Mol. Microbiol.* 97, 216-228.

Kariisa, A., Weeks, K. and Tamayo, R. (2016). The RNA Domain Vc1 Regulates Downstream Gene Expression in Response to Cyclic Diguanylate in *Vibrio cholerae*. *PLoS ONE* 11, e0148478.

Kersey, C., Agyemang, P. and Dumenyó, C. (2012). CorA, the magnesium/nickel/cobalt transporter, affects virulence and extracellular enzyme production in the soft rot pathogen *Pectobacterium carotovorum*. *Mol. Plant Pathol.* 13, 58-71.

Kulkarni, P., Jia, T., Kuehne, S., Kerkering, T., Morris, E., Searle, M. et al. (2014). A sequence-based approach for prediction of CsrA/RsmA targets in bacteria with experimental validation in *Pseudomonas aeruginosa*. *Nucleic Acids Res.* 42, 6811-6825.

Lapouge, K., Schubert, M., Allain, F. and Haas, D. (2008). Gac/Rsm signal transduction pathway of gamma-proteobacteria: from RNA recognition to regulation of social behaviour. *Mol. Microbiol.* 67, 241-253.

Lebreton, A. and Cossart, P. (2017). RNA- and protein-mediated control of *Listeria monocytogenes* virulence gene expression. *RNA Biol* 14, 460-470.

Leimeister-Wächter, M., Haffner, C., Domann, E., Goebel, W. and Chakraborty, T. (1990). Identification of a gene that positively regulates expression of listeriolysin, the major virulence factor of *Listeria monocytogenes*.. *Proc Natl Acad Sci U S A* 87, 8336-8340.

Leonard, S., Meyer, S., Lacour, S., Nasser, W., Hommais, F. and Reverchon, S. (2019). APERO: a genome-wide approach for identifying bacterial small RNAs from RNA-Seq data. *Nucleic Acids Res.*

Loh, E., Dussurget, O., Gripenland, J., Vaitkevicius, K., Tiensuu, T., Mandin, P. et al. (2009). A trans-acting riboswitch controls expression of the virulence regulator PrfA in

Listeria monocytogenes. Cell 139, 770-779. doi: 10.1016/j.cell.2009.08.046.

Mao, C., Abraham, D., Wattam, A., Wilson, M., Shukla, M., Yoo, H. et al. (2015). Curation, integration and visualization of bacterial virulence factors in PATRIC. *Bioinformatics* 31, 252-258.

Mastropasqua, M., Lamont, I., Martin, L., Reid, D., D'Orazio, M. and Battistoni, A. (2018). Efficient zinc uptake is critical for the ability of *Pseudomonas aeruginosa* to express virulence traits and colonize the human lung. *J Trace Elem Med Biol* 48, 74-80.

Mengaud, J., Dramsi, S., Gouin, E., Vazquez-Boland, J., Milon, G. and Cossart, P. (1991). Pleiotropic control of *Listeria monocytogenes* virulence factors by a gene that is autoregulated. *Mol Microbiol* 5, 2273-2283.

Miyajima, A., Shibuya, M., Kuchino, Y. and Kaziro, Y. (1981). Transcription of the *E. coli* *tufB* gene: Cotranscription with four tRNA genes and inhibition by guanosine-5'-diphosphate-3'-diphosphate. *Molecular and General Genetics MGG* 183, 13-19. doi: 10.1007/BF00270131.

Morita, M., Tanaka, Y., Kodama, T., Kyogoku, Y., Yanagi, H. and Yura, T. (1999). Translational induction of heat shock transcription factor sigma32: evidence for a built-in RNA thermosensor. *Genes Dev* 13, 655-665.

Mulhbachter, J., Brouillette, E., Allard, M., Fortier, L., Malouin, F. and Lafontaine, D. (2010). Novel riboswitch ligand analogs as selective inhibitors of guanine-related metabolic pathways. *PLoS Pathog.* 6, e1000865.

Naghdi, M., Smail, K., Wang, J., Wade, F., Breaker, R. and Perreault, J. (2017). Search for 5'-leader regulatory RNA structures based on gene annotation aided by the RiboGap database. *Methods* 117, 3-13. doi: 10.1016/j.ymeth.2017.02.009.

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

Nelson, J., Sudarsan, N., Furukawa, K., Weinberg, Z., Wang, J. and Breaker, R. (2013). Riboswitches in eubacteria sense the second messenger c-di-AMP. *Nat. Chem. Biol.* 9, 834-839.

Nelson, J., Sudarsan, N., Phillips, G., Stav, S., Lunse, C., McCown, P. et al. (2015). Control of bacterial exoelectrogenesis by c-AMP-GMP. *Proc. Natl. Acad. Sci. U.S.A.* 112, 5389-5394.

Nies, D. (2019). The ancient alarmone ZTP and zinc homeostasis in *Bacillus subtilis*. *Mol. Microbiol.*

Nuss, A., Heroven, A. and Dersch, P. (2017). RNA Regulators: Formidable Modulators of *Yersinia* Virulence. *Trends Microbiol.* 25, 19-34.

Pacheco, A., Curtis, M., Ritchie, J., Munera, D., Waldor, M., Moreira, C. et al. (2012). Fucose sensing regulates bacterial intestinal colonization. *Nature* 492, 113-117. doi: 10.1038/nature11623.

Palmer, L. and Skaar, E. (2016). Transition Metals and Virulence in Bacteria. *Annu Rev Genet* 50, 67-91. doi: 10.1146/annurev-genet-120215-035146.

Papp-Wallace, K. and Maguire, M. (2006). Manganese transport and the role of manganese in virulence. *Annu. Rev. Microbiol.* 60, 187-209.

Perreault, J., Weinberg, Z., Roth, A., Popescu, O., Chartrand, P., Ferbeyre, G. et al. (2011). Identification of hammerhead ribozymes in all domains of life reveals novel structural variations. *PLoS Comput. Biol.* 7, e1002031.

Price, I., Gaballa, A., Ding, F., Helmann, J. and Ke, A. (2015). Mn(2+)-sensing mechanisms of *yybP-ykoY* orphan riboswitches. *Mol. Cell* 57, 1110-1123.

Raina, M. and Ibbá, M. (2014). tRNAs as regulators of biological processes. *Front Genet* 5, 171.

- Ramesh, A. and Winkler, W. (2010). Magnesium-sensing riboswitches in bacteria.. *RNA Biol* 7, 77-83.
- Rasko, D., Moreira, C., Li, D., Reading, N., Ritchie, J., Waldor, M. et al. (2008). Targeting QseC signaling and virulence for antibiotic development.. *Science* 321, 1078-1080. doi: 10.1126/science.1160354.
- Remy, L., Carriere, M., Derre-Bobillot, A., Martini, C., Sanguinetti, M. and Borezee-Durant, E. (2013). The *Staphylococcus aureus* Opp1 ABC transporter imports nickel and cobalt in zinc-depleted conditions and contributes to virulence. *Mol. Microbiol.* 87, 730-743.
- Ryckelynck, M., Giege, R. and Frugier, M. (2005). tRNAs and tRNA mimics as cornerstones of aminoacyl-tRNA synthetase regulations. *Biochimie* 87, 835-845.
- Sarkisova, S., Lotlikar, S., Guragain, M., Kubat, R., Cloud, J., Franklin, M. et al. (2014). A *Pseudomonas aeruginosa* EF-hand protein, EfhP (PA4107), modulates stress responses and virulence at high calcium concentration. *PLoS ONE* 9, e98985.
- Sarkisova, S., Patrauchan, M., Berglund, D., Nivens, D. and Franklin, M. (2005). Calcium-induced virulence factors associated with the extracellular matrix of mucoid *Pseudomonas aeruginosa* biofilms. *J. Bacteriol.* 187, 4327-4337.
- Sayers, S., Li, L., Ong, E., Deng, S., Fu, G., Lin, Y. et al. (2019). Victors: a web-based knowledge base of virulence factors in human and animal pathogens. *Nucleic Acids Res.* 47, D693-D700.
- Sharma, A., Dhasmana, N., Dubey, N., Kumar, N., Gangwal, A., Gupta, M. et al. (2017). Bacterial Virulence Factors: Secreted for Survival. *Indian J. Microbiol.* 57, 1-10.
- Sherlock, M., Sudarsan, N. and Breaker, R. (2018). Riboswitches for the alarmone ppGpp expand the collection of RNA-based signaling systems. *Proc. Natl. Acad. Sci. U.S.A.* 115, 6052-6057.
- Shi, Y., Zhao, G. and Kong, W. (2014). Genetic analysis of riboswitch-mediated transcriptional regulation responding to Mn²⁺ in *Salmonella*. *J. Biol. Chem.* 289, 11353-11366.
- Skuzeski, J., Bozarth, C. and Dreher, T. (1996). The turnip yellow mosaic virus tRNA-like structure cannot be replaced by generic tRNA-like elements or by heterologous 3' untranslated regions known to enhance mRNA expression and stability. *J. Virol.* 70, 2107-2115.
- Stav, S., Atilho, R., Mirihana Arachchilage, G., Nguyen, G., Higgs, G. and Breaker, R. (2019). Genome-wide discovery of structured noncoding RNAs in bacteria. *BMC Microbiol.* 19, 66.
- Sudarsan, N., Lee, E., Weinberg, Z., Moy, R., Kim, J., Link, K. et al. (2008). Riboswitches in eubacteria sense the second messenger cyclic di-GMP. *Science* 321, 411-413.
- Tamayo, R. (2019). Cyclic diguanylate riboswitches control bacterial pathogenesis mechanisms. *PLoS Pathog.* 15, e1007529.
- Tang, Q., Yin, K., Qian, H., Zhao, Y., Wang, W., Chou, S. et al. (2016). Cyclic di-GMP contributes to adaption and virulence of *Bacillus thuringiensis* through a riboswitch-regulated collagen adhesion protein. *Sci Rep* 6, 28807.
- Vakulskas, C., Potts, A., Babitzke, P., Ahmer, B. and Romeo, T. (2015). Regulation of bacterial virulence by Csr (Rsm) systems. *Microbiol. Mol. Biol. Rev.* 79, 193-224.
- Valentini, M. and Filloux, A. (2016). Biofilms and Cyclic di-GMP (c-di-GMP) Signaling: Lessons from *Pseudomonas aeruginosa* and Other Bacteria. *J. Biol. Chem.* 291, 12547-12555.
- Valverde, C., Lindell, M., Wagner, E. and Haas, D. (2004). A repeated GGA motif is

critical for the activity and stability of the riboregulator RsmY of *Pseudomonas fluorescens*. *J. Biol. Chem.* 279, 25066-25074.

Velasco, E., Wang, S., Sanet, M., Fernandez-Vazquez, J., Jove, D., Glaria, E. et al. (2018). A new role for Zinc limitation in bacterial pathogenicity: modulation of α -hemolysin from uropathogenic *Escherichia coli*. *Sci Rep* 8, 6535.

Véscovi, E., Ayala, Y., Di Cera, E. and Groisman, E. (1997). Characterization of the bacterial sensor protein PhoQ. Evidence for distinct binding sites for Mg^{2+} and Ca^{2+} . *J Biol Chem* 272, 1440-1443.

Wattam, A., Abraham, D., Dalay, O., Disz, T., Driscoll, T., Gabbard, J. et al. (2014). PATRIC, the bacterial bioinformatics database and analysis resource.. *Nucleic Acids Res* 42, D581-D591. doi: 10.1093/nar/gkt1099.

Wedekind, J., Dutta, D., Belashov, I. and Jenkins, J. (2017). Metalloriboswitches: RNA-based inorganic ion sensors that regulate genes. *J. Biol. Chem.* 292, 9441-9450.

Weinberg, Z., Lunse, C., Corbino, K., Ames, T., Nelson, J., Roth, A. et al. (2017). Detection of 224 candidate structured RNAs by comparative analysis of specific subsets of intergenic regions. *Nucleic Acids Res.* 45, 10811-10823.

Weinberg, Z., Wang, J., Bogue, J., Yang, J., Corbino, K., Moy, R. et al. (2010). Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol.* 11, R31.

Wishart, D., Feunang, Y., Marcu, A., Guo, A., Liang, K., Vazquez-Fresno, R. et al. (2018). HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* 46, D608-D617.

Wishart, D., Jewison, T., Guo, A., Wilson, M., Knox, C., Liu, Y. et al. (2013). HMDB 3.0--The Human Metabolome Database in 2013. *Nucleic Acids Res.* 41, D801-807.

Wishart, D., Knox, C., Guo, A., Eisner, R., Young, N., Gautam, B. et al. (2009). HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Res.* 37, D603-610.

Wishart, D., Tzur, D., Knox, C., Eisner, R., Guo, A., Young, N. et al. (2007). HMDB: the Human Metabolome Database. *Nucleic Acids Res.* 35, D521-526.

Zeenko, V., Ryabova, L., Spirin, A., Rothnie, H., Hess, D., Browning, K. et al. (2002). Eukaryotic elongation factor 1A interacts with the upstream pseudoknot domain in the 3' untranslated region of tobacco mosaic virus RNA. *J. Virol.* 76, 5678-5691.

Zhang, X., Liu, Z., Yi, J., Tang, H., Xing, J., Yu, M. et al. (2012). The tRNA methyltransferase NSun2 stabilizes p16INK⁴ mRNA by methylating the 3'-untranslated region of p16. *Nat Commun* 3, 712.

5. ARNnc candidats potentiels de *riboswitchs* calcium

Ce chapitre présente des résultats qui seront probablement publiés suite à plusieurs autres expériences.

5.1 Mise en contexte :

Plusieurs cations sont très importants pour les activités vitales de tous les domaines de la vie. D'une part certains métaux sont essentiels pour les bactéries et d'autres part ils sont toxiques, voire létiaux, lorsqu'en concentration trop élevée. De ce fait, les micro-organismes ont développé différents systèmes de régulation.

Les *riboswitchs* sont capables de reconnaître une molécule particulière, des métabolites ou des ions sans avoir besoin de protéines (Nahvi et al., 2002) et par conséquent régule l'expression du gène en aval. Il y a des *riboswitchs* spécifiques pour les métaux, tel que les *riboswitch* magnésium (Cromie et al., 2006; Dann et al., 2007) le *riboswitch* manganèse (Dambach et al., 2015; Price et al., 2015) de même que le *riboswitch* nickel et cobalt (Furukawa et al., 2015). Notre motivation est de retrouver des nouveaux *riboswitchs* pour d'autres métaux.

5.2 Matériels et Méthodes :

5.2.1 Bioinformatique :

Afin de retrouver les séquences intergéniques ayant un potentiel élevé de contenir des *riboswitchs* dont le ligand serait un cation, la base de données de RiboGap (Naghdi et al., 2017) a été utilisée. RiboGap permet d'utiliser le « *pattern matching* » ce qui est très puissant pour chercher des mots clés particuliers. Le motif de « *regular expression* » (REGEX), (`[[:<:]Ca[:>:]]Calcium`) a été formulé pour chercher toutes les séquences intergéniques qui ont le mot calcium ou le symbole chimique Ca^{2+} dans la description du gène annoté (Naghdi et al., 2017). Après avoir extrait toutes les séquences intergéniques pour le mot clé utilisé la méthode déjà décrite dans le chapitre 3 a été utilisée (Naghdi et al., 2017). Cette méthode utilise plusieurs logiciels comme

GraphClust (Heyne et al., 2012), *CMfinder* (Yao et al., 2006) et *Infernal* (Nawrocki et al., 2013) pour trouver les modèles de structures secondaires conservées.

5.2.2 Validation fonctionnelle :

5.2.2.1 Détermination de la structure et de l'affinité *in vitro* par *in line probing* :

Parmi les *riboswitchs* potentiels trouvés, les séquences intergéniques de *Desulfovibrio magneticus* RS-1, avec le numéro d'accèsion NC_012796 : 5019180 – 5019383 et celle de *Desulfovibrio desulfuricans* subsp. *desulfuricans* str. ATCC 27774, (numéro d'accèsion NC_011883 : 720681 – 720371 et locus_tag : Ddes_0594) ayant le gène « échangeur du Sodium/Calcium » ont été choisies pour être testées (voir l'annexe I). Différentes constructions de *riboswitchs* ont été réalisées par la méthode PCR chevauchante en utilisant l'outil en ligne « Primerize » (Tian et al., 2015).

La méthode *in-line probing* (Regulski et al., 2008b) a été utilisée pour déterminer la modulation de l'ARN en présence des ions manganèse et calcium tels que décrite dans le chapitre 3 (Naghdi et al., 2017). La mesure de la fraction d'ARN qui module se calcule ainsi :

$$\frac{L' \text{ intensité normalisée de chaque bande} - l' \text{ intensité normalisée des bandes}}{L' \text{ intensités normalisée maximale des bandes} - l' \text{ intensité normalisée minimale}}$$

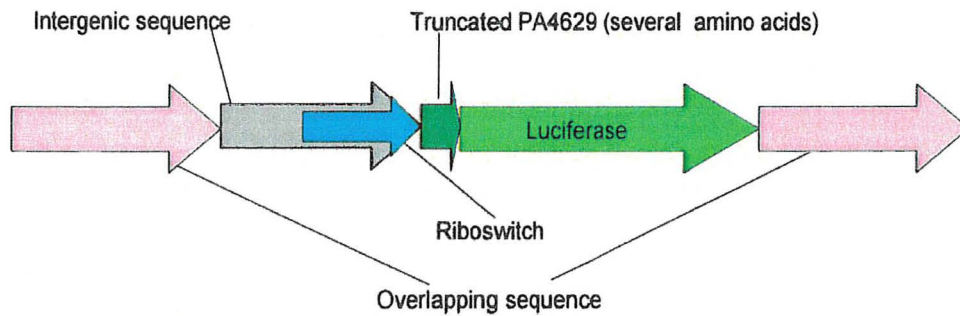
Il faut ensuite tracer la fraction d'ARN clivée en fonction logarithmique de la concentration du ligand sachant que le K_d est la concentration du ligand dans laquelle la moitié de l'ARN est clivée.

5.2.2.2 Clonage :

La séquence du *riboswitch* *yybP-ykoY* avec 15 nucléotides de séquence codante de *Pseudomonas aeruginosa* PA14, locus-tag : PA1461260 a été clonée dans le vecteur de mini-CTX-Lux (Becher et al., 2000) en aval du gène *luxC*, en utilisant la méthode Gibson de NEB (décrit ci-dessous) (Gibson et al., 2009), (Figure 5.1 et Figure 5.2).

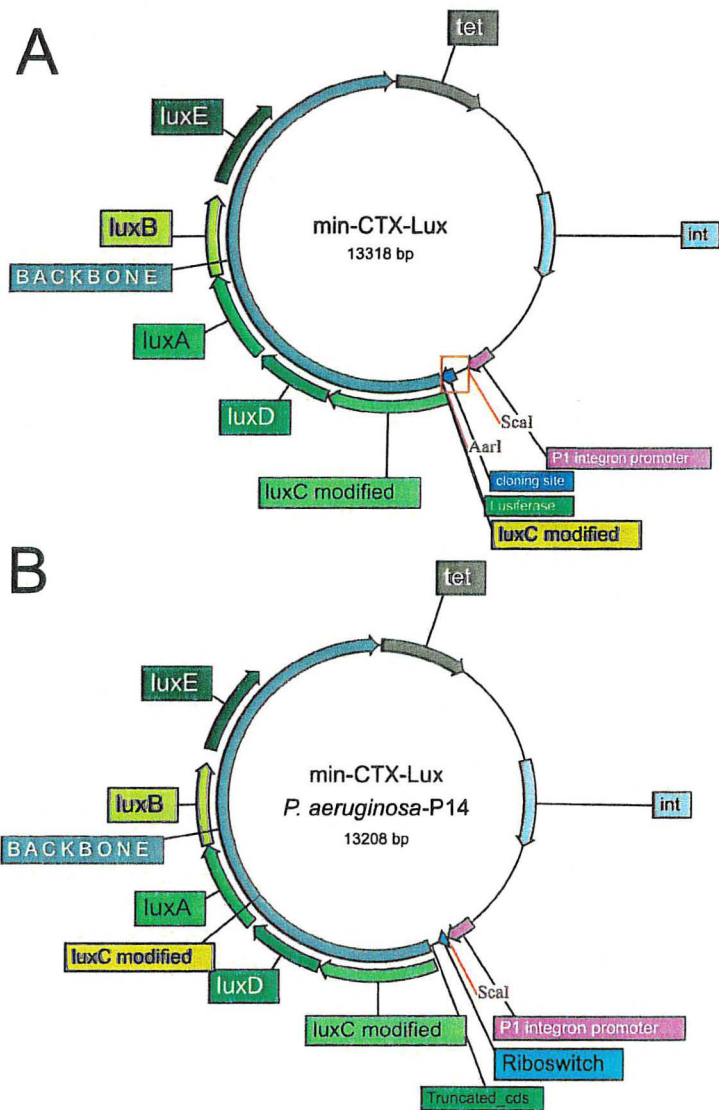
Très brièvement la séquence à cloner a été construite par la PCR classique à partir d'ADN génomique de *Pseudomonas aeruginosa* PA14. Des amorces (Fw et Rev) se retrouvent dans l'annexe I. Cette séquence comprend la séquence du *riboswitch* *yybP-ykopY* prédit par Rfam et 15 nucléotides de séquence codante du locus_tag : PA4629 encadrée par deux parties communes (séquences) du vecteur qui permet d'intégrer la séquence cible dans le vecteur selon la méthode Gibson (Figure 5.1). Le site du clonage du vecteur, juste en amont du gène *LuxC* a été coupé par deux enzymes de restriction *AarI* et *Scal* (Figure 5.2 A, B), en incubant pendant 16 h selon le protocole fourni par Thermo Fisher.

Le vecteur cloné (contenant le *riboswitch*) a été transformé dans les bactéries compétentes *E. coli* DH5 α . Les bactéries ont été étalées sur le milieu LB Agar avec tétracycline (125 μ g/mL). La séquence clonée a été vérifiée par la méthode séquençage de Sanger. Après avoir confirmé le clonage, le vecteur cloné a été transformé dans les bactéries compétentes, *E. coli* SM10 et ensuite à partir de cette dernière dans la bactérie *Pseudomonas aeruginosa* PA14 par la conjugaison.



20. Figure 5.1 La présentation schématique de la séquence à cloner dans le vecteur mini-CTX-Lux.

La construction de la séquence à cloner dans le vecteur mini-CTX-Lux. La séquence a été construite par la méthode PCR sur l'ADN génomique de la bactérie *Pseudomonas aeruginosa* PA14.



21. Figure 5.2 Le vecteur mini-CTX-Lux et le *riboswitch yybP-ykoY* cloné.

A) Le vecteur mini-CTX-Lux avec le site du clonage encerclé. Ce site contient deux enzymes de restriction *AarI* et *ScaI*. D'autres enzymes de restriction n'ont pas été montrées. B) Le vecteur mini-CTX-Lux cloné avec le *riboswitch*. Le site du clonage a été remplacé par le *riboswitch* et 15 nucléotides de la séquence codante du locus_tag : PA14-61260.

5.2.2.3 Mesure d'expression par luminescence :

Afin de réaliser la mesure d'expression par luminescence, les bactéries doivent pousser dans des conditions où le *riboswitch* peut réguler l'expression du gène rapporteur, et par conséquent la variation de l'intensité de luminescence de ce dernier est directement liée à la régulation par le *riboswitch*.

Le milieu M9 sans calcium a été préparé à partir des sels M9 : (64 g $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$, 15 g KH_2PO_4 , 2.5 g NaCl , 5 g NH_4Cl dans 1 Litre d'eau). Ensuite 200 mL de sel M9 préparé avec 2 mL de 1M MgSO_4 , et 20 mL de glucose 20 % est dissous dans 1 L d'eau. Sachant que les concentrations finales du magnésium et du glucose seront 2 mM et 0,4 % respectivement.

Le test d'expression a été effectué dans le milieu M9 et LB avec différentes concentrations d'ions divalents, de 100 μM à 3 mM calcium et de 100 μM à 300 μM manganèse en présence et en absence de 1 mM (concentration finale) d'agents chélateurs Acide Ethylene Diamine Tetraacetic (EDTA) et Acide Éthylène glycol-bis (β -aminoethyl ether)-N, N, N', N'-tetraacetic (EGTA). Sachant que le manganèse est toxique dans des concentrations plus élevées.

Les bactéries *Pseudomonas aeruginosa* PA14 ont été mises-en-culture pendant la nuit dans le milieu minimal M9 sans calcium et / ou LB pour les tests d'expression (Chan et al., 2016). Les deux milieux contiennent 75 $\mu\text{g}/\text{mL}$ de tétracycline. La Densité Optique (DO) à 600 nm a été mesurée après ~ 18 h et la quantité nécessaire des bactéries a été calculée pour que la mesure de l'activité de la régulation (luminescence) s'effectue pendant la phase exponentielle de la croissance et se prolonge dans la phase stationnaire. Par exemple la culture bactérienne se dilue pour arriver à la Densité Optique (DO) 0,01. La mesure de luminescence s'est effectuée en utilisant l'appareil de « Cytation 3; BioTek Instruments, Inc., Winooski, VT » avec les plaques 96 puits. La croissance de bactéries a été mesurée à 600 nm. Les résultats se normalisent automatiquement avec les blancs (LB ou M9) et ensuite ils ont été utilisés pour calculer la moyenne et l'écart type. Des valeurs obtenus ($\text{Lum}/\text{DO}_{600}$) ont été présentées en fonction de différentes conditions dans des histogrammes suivants (Figure 5.7 A, B et 5.8). Des scripts maisons ont été réalisés pour faciliter cette analyse.

5.3 Résultats :

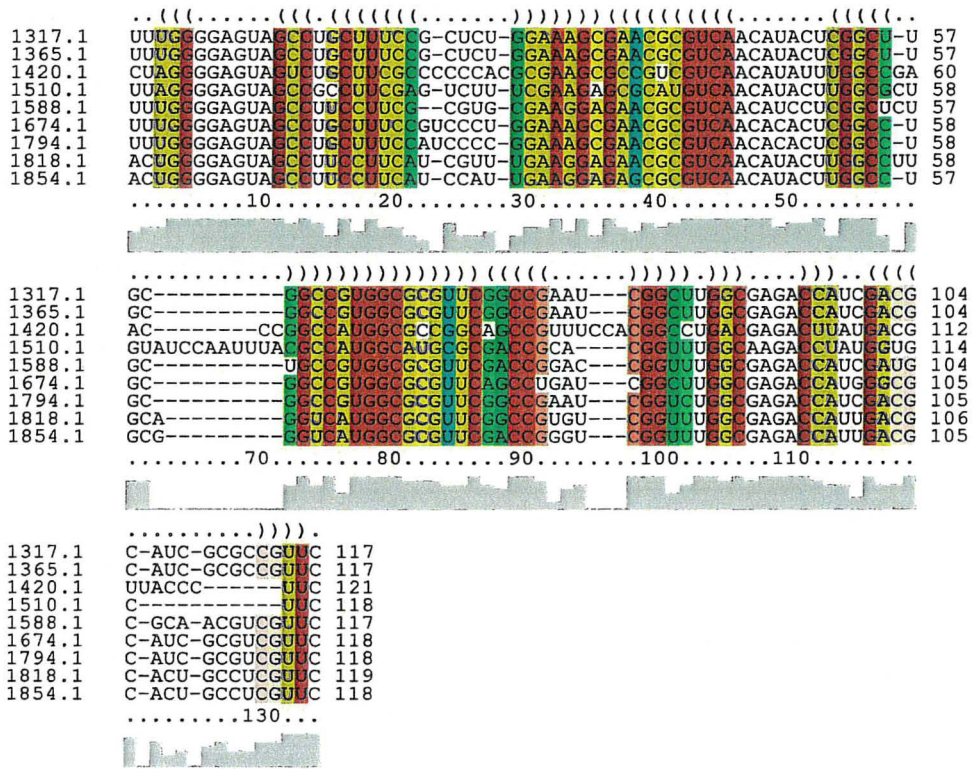
5.3.1 Prédiction de nouvelles familles d'ARNnc :

Après avoir utilisé la méthode décrite auparavant (Naghdi et al., 2017) plusieurs clusters intéressants ont été identifiés pour le mot clé « calcium ». Chaque cluster contient la prédiction de structure secondaire avec des alignements de séquences de plusieurs programmes y compris *CMfinder* (Yao et al., 2006) et *LocARNA* (Will et al., 2007). *Graphclust* prédit par défaut 10 clusters probables de différentes structures secondaires probables. Enfin deux clusters ont été considérés comme suffisamment prometteurs pour les valider expérimentalement.

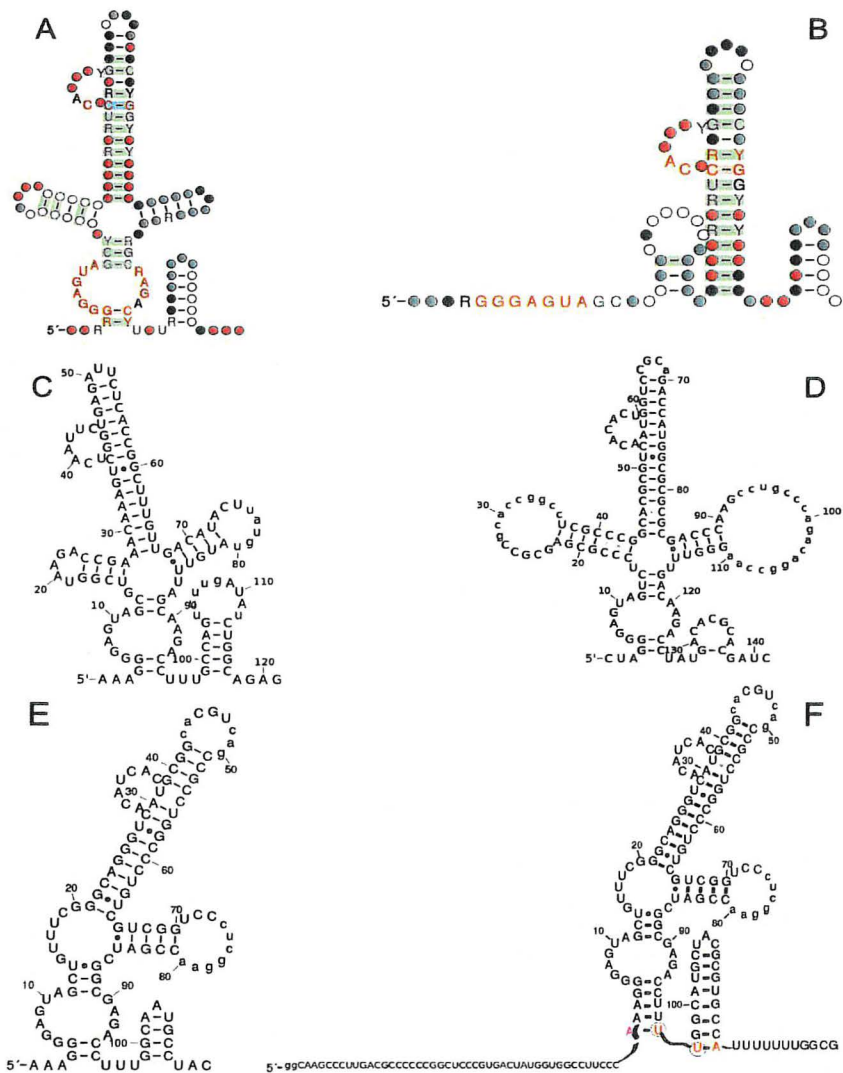
5.3.1.1 Structure ressemblant à *yybP-ykoY* :

Le cluster #1.1 trouvé en amont de gènes associés au calcium était particulièrement intéressant, car il avait une structure élaborée avec certaines régions simples brins bien conservées et beaucoup de covariation dans toutes ces tiges (Figure 5.3). De plus, une évaluation rapide de plusieurs des instances de ce motif semblait indiquer qu'il ne correspondait à aucune famille d'ARN de Rfam. Cependant, un examen plus approfondi a révélé que ce motif d'ARN correspondait au motif déjà décrit *yybP-ykoY* (Barrick et al., 2004; Dambach et al., 2015; Price et al., 2015).

Bien que notre recherche ait révélé plusieurs instances qui n'étaient pas annotées, probablement dû à de petites différences dans les consensus des deux motifs (Figure 5.3). Peu de temps après, cette structure d'ARN a été publiée comme étant sensible au manganèse (Dambach et al., 2015; Price et al., 2015). Néanmoins, la présence de ces motifs en amont de gènes associés au calcium (et non au manganèse) était pour le moins intrigante et quelques exemples ont donc été choisis pour être étudiés.



22. Figure 5.3 L'alignement obtenu à partir des séquences retrouvées en aval des gènes ayant un rapport au calcium.



23. Figure 5.4 La comparaison des modèles du *riboswitch* du manganèse et le candidat du calcium.

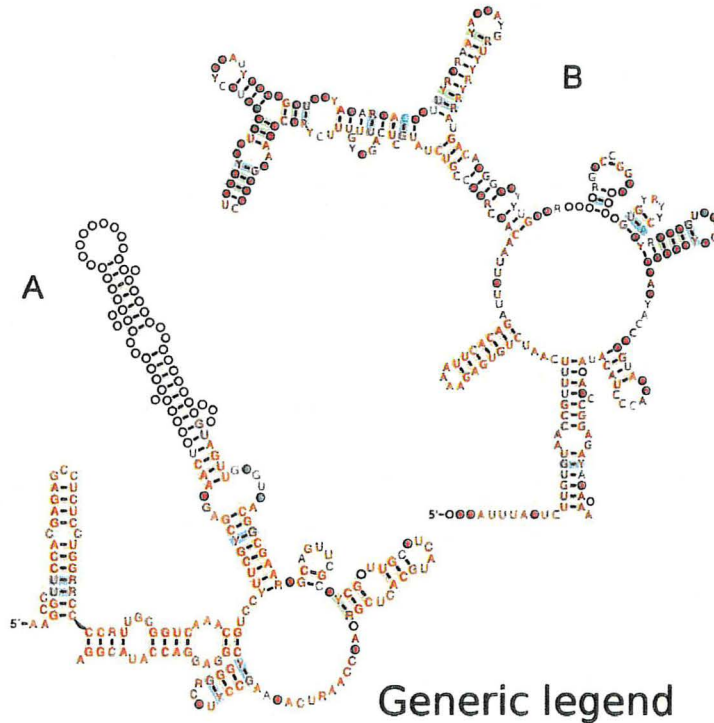
A) Le consensus du motif potentiel pour le *riboswitch* calcium. B) Le consensus fait à partir de l'alignement du *riboswitch* manganèse de Rfam. C) La structure correspond au motif *yybP-ykoY* chez *Lactococcus lactis* (Price et al., 2015). D) La structures *Pseudomonas aeruginosa* PA14 faite à partir du consensus du modèle retrouvé par *GraphClust*. E, F) Le candidat potentiel de *Desulfovibrio magneticus* RS-1 dans le modèle ON (E), contrairement à l'image, présumément la boucle E ne se formerait pas sous la forme « ON ». F) Le modèle potentiel pour la forme « OFF », avec un terminateur de transcription ; Les deux nucléotides entourés en rouge (U) représentent une même position avec deux possibilités de former les paires de bases avec A (mauve et rouge), tantôt en amont, tantôt en aval du U97. La numérotation des positions en (F) a été conservée comme en (E) pour fin de clarté.

Un motif intéressant a été retrouvé chez le genre *Desulfovibrio*, tel que l'exemple de *Desulfovibrio magneticus* RS-1, (NC_012796) (Figure 5.4 A (consensus), E (modèle ON) et F (modèle OFF)) et *Desulfovibrio desulfuricans* subsp. *desulfuricans* str. ATCC 27 774, (NC_011883). En comparant le motif *yybP-ykoY* avec nos séquences obtenues par la prédiction (*GraphClust* et *Infernal*), nous constatons la formation d'une grande tige avec un bulge (*E-value* égale à $8.977e^{-05}$ et $5.44e^{-06}$) pour *Desulfovibrio magneticus* RS-1 et *Desulfovibrio desulfuricans* subsp. *desulfuricans* str. ATCC 27774 respectivement ; tandis que la modèle prédite de *GraphClust* donnera le *E-value* plus intéressante égale à $9,74e^{-11}$ et $8,066e^{-11}$ pour *Desulfovibrio magneticus* RS-1 et *Desulfovibrio desulfuricans* subsp. *desulfuricans* str. ATCC 27774 respectivement (voir l'annexe II, B.1, B.2). En conclusion le modèle que l'on a prédit, ce n'est pas un consensus typique mais pas trop éloigné du motif *yybP-ykoY*.

Le motif retrouvé chez *Desulfovibrio magneticus* RS-1 a été étudié plus en profondeur par *in-line probing* (Figure 5.6). Un autre motif de candidat a été retrouvé dans la séquence intergénique de *Pseudomonas aeruginosa* PA14 (Figure 5.4 D), locus-tag : PA4629 (hypothetical protein) et a été choisi pour être étudié par des essais fonctionnels avec des gènes rapporteurs.

5.3.1.2 Nouvelles structures de riboswitchs candidats pour le calcium :

Deux autres candidats pour le calcium ont été retrouvés (Figure 5.5) par bio-informatique chez genre de *Sinorhizobium*, *Yersinia* et *Serratia*. Ces deux candidats montrent des structures tout à fait différentes de toutes structures connues (du moins dans Rfam). Plus de tests, surtout des tests *in-vivo*, sont nécessaires pour ces deux candidats afin de pouvoir déterminer l'effet du calcium sur l'ARN et sur l'expression des gènes en aval. Le clonage en fusion traductionnelle a été entrepris pour ces deux ARNnc potentiels dans un vecteur pRS414 pour tester la structure de *Serratia* / *Yersinia* chez *E. coli* (également une entérobactérie) avec des essais Miller. Il a aussi été amorcé avec un vecteur spécifique (pTO1) à *Sinorhizobium* que nous avons dû modifier pour permettre une fusion traductionnelle avec GFP.



Generic legend

base pair annotations
 - covarying mutations
 - compatible mutations
 - no mutations observed

nucleotide present		nucleotide identity	
● 97%	● 75%	N	97%
● 90%	● 50%	N	90%
		N	75%

24. Figure 5.5 Deux candidats obtenus prédits, à partir de gènes liés au calcium.

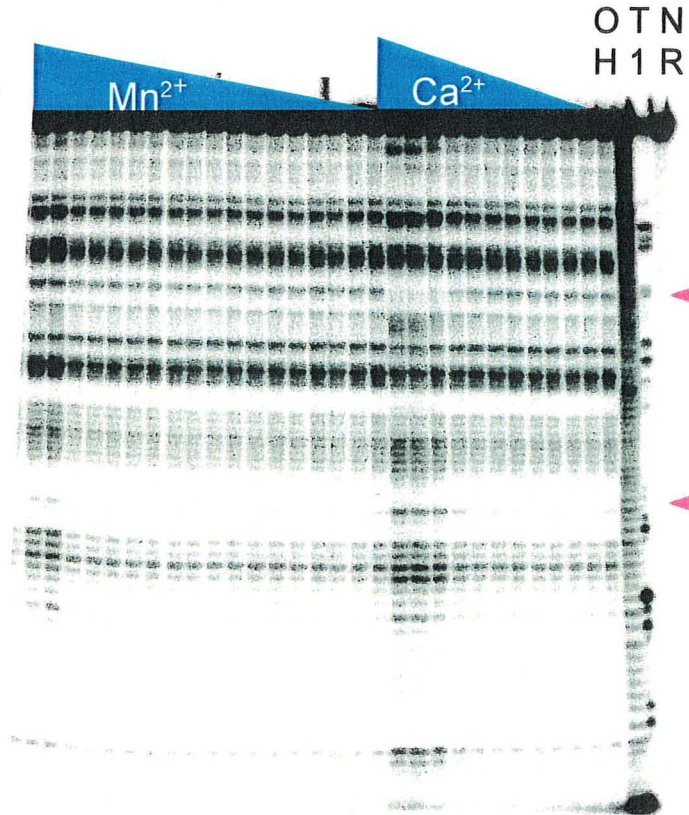
A) Le candidat pour l'ARNnc retrouvé chez *Sinorhizobium meliloti*. **B)** Le candidat pour l'ARNnc retrouvé chez *Serratia et Yersinia*.

5.3.2 Caractérisation des ARNnc ressemblants à *yybP-ykoY* :

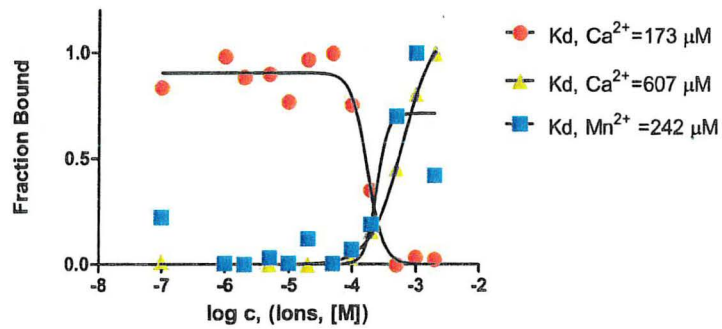
5.3.2.1 *in line-probing* :

Très brièvement, l'ARN marqué avec [³²P] a été incubé dans le tampon alcalin (pH ≈ 8.3) pendant ≈ 40 h à la température pièce avec des concentrations différentes de calcium et de manganèse. Le principe de la méthode est basé sur la dynamique moléculaire de l'ARN qui se dégrade naturellement. Si l'ARN lie à un ligand particulier, le patron de dégradation changera comparativement à l'ARN sans ligand. Donc la modulation du patron de dégradation sur le gel, indique le changement de la structure de l'ARN à cause de la liaison au ligand. À partir de ces données, il est possible de déterminer la constante de dissociation, ou K_d , ce qui permet d'évaluer l'affinité de l'ARN au ligand. De plus, le coefficient de Hill donne un indice sur le nombre de site spécifique pour le ligand et nous renseigne sur la coopérativité de la liaison du ligand à l'ARN si le n (le nombre de site) $n > 1$ tandis que $n=1$ indique aucune coopérativité. La coopérativité pour certains *riboswitchs* liant des ions a été déjà montrée comme dans le cas du *riboswitch* nickel/cobalt (Furukawa et al., 2015). Suite au *in-line probing* le K_d a été calculé pour le calcium et le manganèse ($K_d=173 \mu\text{M}$ et $K_d= 242 \mu\text{M}$ respectivement) (Figure 5.6 B). Les courbes qui montent et celles qui descendent correspondent à des bandes dont l'intensité diminue et augmente respectivement (les flèches indiquées dans la Figure 5.6 A). Nos données indiquent une apparente coopérativité avec un coefficient de Hill calculé de ~ 1.7 , ce qui suggère qu'il existe au moins deux sites pour le ligand.

A



B



25. Figure 5.6 *in-line probing* obtenu à partir du motif retrouvé chez *Desulfovibrio magneticus* RS-1 en présence de manganèse et calcium.

A) La concentration du calcium (Zero, 1, 2, 5, 10, 20, 50, 100, 200, 500 μM ; 1, 2 mM) et du manganèse (Zero, 2, 5, 10, 20, 50, 100, 200, 500 nM ; 1, 2, 5, 10, 20, 50, 100, 200, 500 μM) ; En présence de 2 mM $[\text{Mg}^{2+}]$. B) Le graphe présente le K_d calculé pour les bandes de calcium rond rouge et triangle ocre) et le manganèse (carré bleu).

5.3.2.2 Mesurer l'expression du gène rapporteur ayant le motif *yybP-ykoY* chez *Pseudomonas aeruginosa*.

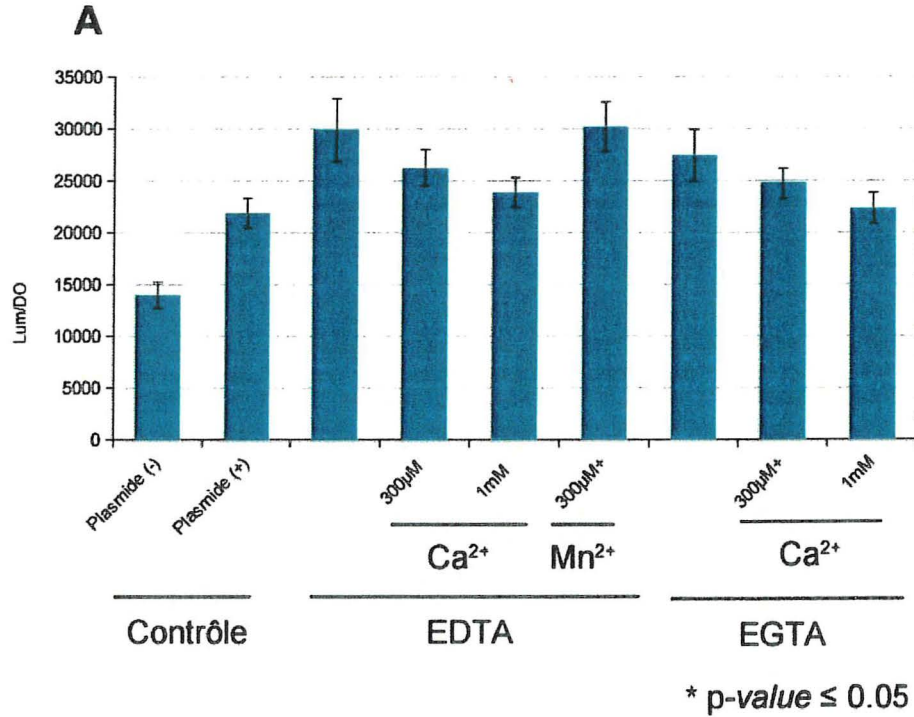
Notre recherche par *cmsearch* a révélé que le *riboswitch yybP-ykoY* se retrouve dans la région intergénique du locus tag PA14-61260 chez *Pseudomonas aeruginosa*, cet exemple de l'ARN *yybP-ykoY* avait d'ailleurs déjà été annoté dans Rfam, et donc aussi dans RiboGap. Ce *riboswitch* a donc été cloné et plusieurs essais chez *P. aeruginosa* (sauvage et mutant PA14-61260) ayant le gène rapporteur *Lux* ont été effectués afin de déterminer l'effet des ions sur la régulation de ce gène (locus-tag). De l'EDTA et de l'EGTA à 1 mM (concentration finale) ont été utilisés pour chélater tous les ions, y compris le calcium (et plus spécifiquement le calcium dans le cas de l'EGTA). Les résultats de luminescence obtenus ont été normalisés en divisant la luminescence par la Densité Optique (DO). De manière globale, les résultats obtenus suggèrent que l'expression de la luminescence augmente en présence d'agents chélateurs et est réprimée en présence d'excès de cations divalents (Figure 5.7 A). Parmi les résultats obtenus, il existe des cas où l'expression du gène rapporteur ne change pas vis-à-vis de contrôles en présence des agents chélateurs, mais changeait lorsqu'exposé à un excès d'ions⁴ Ca²⁺, ou encore l'inverse. Néanmoins, en résumé, de façon générale, une faible concentration d'ions semblait favoriser l'expression, alors qu'une concentration élevée semblait favoriser la répression (Figure 5.7 B). À noter que le résultat de la (Figure 5.7 A) est reproductible avec le milieu LB. En conclusion le locus tag PA14-61260 semble affectée par les ions divalents. Les essais avec la bactérie mutante, pour le locus tag PA14-61260 montrent le même résultat que pour la souche sauvage. Des tests statistiques (t-test) ont été effectués pour déterminer si le changement d'expression du gène rapporteur est significatif (Figure 5.7 A et B).

Un *riboswitch* SAM a été utilisé comme contrôle négatif. Contrairement au *riboswitch yybP-ykoY*, il montre une diminution d'expression en présence d'agents chélateurs (Figure 5.8). Ainsi, l'agent chélateur aurait un effet général négatif sur l'expression des gènes et sur la croissance des bactéries, la DO en présence d'EDTA indique d'ailleurs une croissance quasi nulle. Ceci supporterait le fait que la faible augmentation d'expression du rapporteur contrôlé par *yybP-ykoY* ressemblant (*yybP-*

⁴ À noter que certains ions se précipitent dans le milieu M9 en concentration élevée par exemple le calcium pourrait se précipiter à 3 mM.

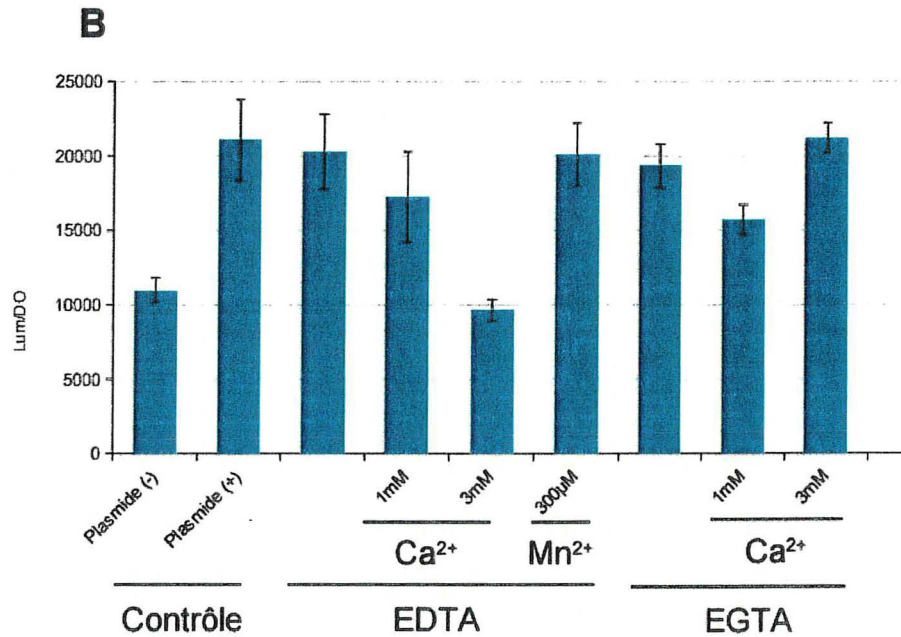
ykoY-r) serait due à une carence en ion qui met le *riboswitch* en mode « ON » et par conséquent active l'expression. Le test statistique ONE way ANOVA montre la valeur-p (*p-value*) ≤ 0.05 (l'annexe VIII).

Les résultats obtenus ont été aussi normalisés par rapport aux ceux du *riboswitch* SAM (Figure 5.9). Les résultats d'expression du *riboswitch* SAM ont été considérés comme contrôle négatif et permettent d'être utilisé comme blanc. Parmi trois expériences, il y a deux cas qui montrent l'expression du gène rapporteur augmente en présence d'agent chélateur, l'EDTA, tandis qu'en compensant avec le calcium l'expression diminue.



	t-test	
	P (T<=t) one-tail	P (T<=t) two-tail
Entre PI (-) et PI (+)	7.2×10^{-14}	1.4×10^{-13}
Entre PI (+) et EDTA	1.4×10^{-6}	2
Entre l'EDTA et l'EDTA+300 µM Ca ²⁺	5.0×10^{-10}	1×10^{-9}
Entre l'EDTA et l'EDTA+1 mM Ca ²⁺	1.0×10^{-6}	2.1×10^{-6}
Entre l'EDTA et l'EDTA+300 µM Mn ²⁺	4.8×10^{-6}	9.5×10^{-6}
Entre PI (+) et l'EGTA	5.9×10^{-7}	1.2×10^{-6}
Entre l'EGTA et l'EGTA+300 µM Ca ²⁺	1.0×10^{-9}	2.1×10^{-9}

26. Figure 5.7 A Déterminer l'effet des agents chélateurs l'EDTA et l'EGTA dans le milieu M9 pour la bactérie *Pseudomonas aeruginosa*.



* p-value ≤ 0.05

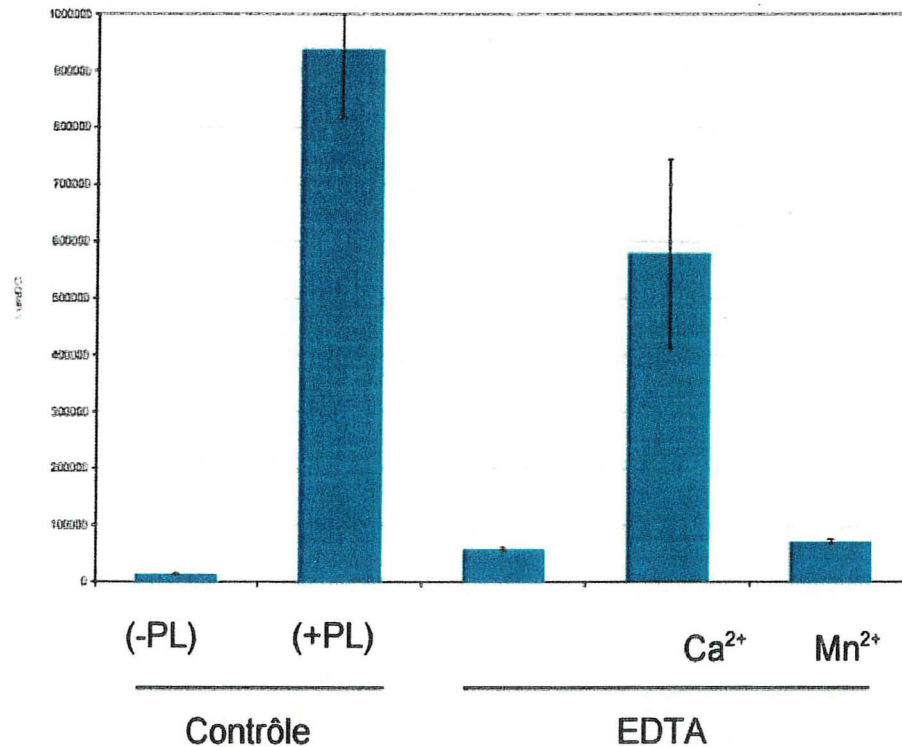
	t-test	
	P (T<=t) one-tail	P (T<=t) two-tail
Entre PI (-) et PI (+)	8.8 x 10 ⁻¹⁰	1.8 x 10 ⁻⁹
Entre PI (+) et l'EDTA	5.8 x 10 ⁻³	1.2 x 10 ⁻²
Entre l'EDTA et l'EDTA+1 mM Ca ²⁺	1.3 x 10 ⁻⁸	2.7 x 10 ⁻⁸
Entre l'EDTA et l'EDTA+3 mM Ca ²⁺	3.9 x 10 ⁻¹⁹	7.7 x 10 ⁻¹⁹
Entre l'EDTA et l'EDTA+300 µM Mn ²⁺	2.6 x 10 ⁻⁸	5.3 x 10 ⁻⁸
Entre PI (+) et l'EGTA	4.9 x 10 ⁻¹	2
Entre l'EGTA et l'EGTA+1 mM Ca ²⁺	9.5 x 10 ⁻⁶	1.9 x 10 ⁻⁵
Entre l'EGTA et l'EGTA+3 mM Ca ²⁺	1.0 x 10 ⁻⁹	2.0 x 10 ⁻⁹

27. Figure 5.7 B Déterminer l'effet des agents chélateurs l'EDTA et l'EGTA dans le milieu M9 pour la bactérie *Pseudomonas aeruginosa*.

L'effet des agents chélateurs a été compensé en rajoutant du calcium et du manganèse dans le milieu M9.

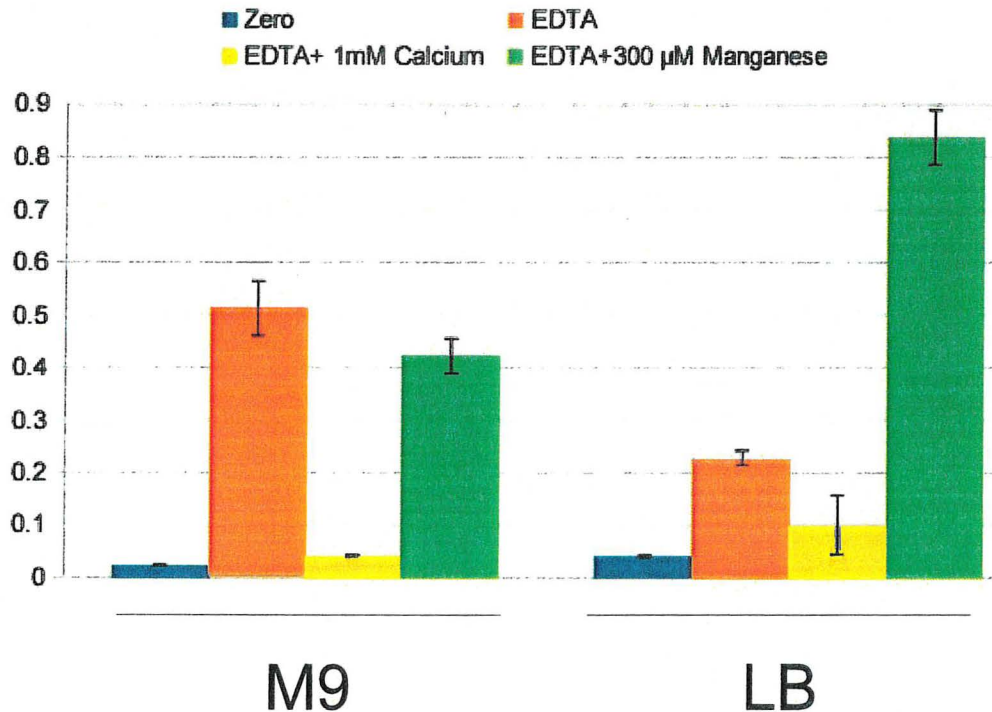
A) Colonne 1, C1 : *P. aeruginosa* sans plasmide (aucune luminescence). C2 : *P. aeruginosa* avec le plasmide. C3 : EDTA 1 mM chélate tous les ions divalents. C4, C5, C6 : EDTA 1 mM a été compensé en rajoutant 300 μ M, de calcium, 1 mM de calcium ou 300 μ M de manganèse, respectivement. C7 : EGTA 1 mM chélate le calcium. C8, C9 : EGTA 1 mM a été compensé en rajoutant de 300 μ M et 1mM de calcium, respectivement.

B) Colonne 1, C1 : *P. aeruginosa* sans plasmide (aucune luminescence). C2 : *P. aeruginosa* avec le plasmide. C3 : EDTA 1 mM chélate tous les ions. C4, C5, C6 : EDTA 1 mM a été compensé en rajoutant 1 mM ou 3 mM de calcium ou 300 μ M de manganèse respectivement. C7 : EGTA 1 mM chélate le calcium. C8, C9 : EGTA 1 mM a été compensé en rajoutant 1mM et 3 mM de calcium respectivement. Chaque condition a été mesurée trois fois et la moyenne et l'écart-type ont été calculés. Chaque colonne de l'histogramme présente la somme des moyennes et d'écart-type de tous les points. Le test statistique a été fait à partir de 30 points dans le temps (sur 16 heures) pour des triplicatas pour chaque condition afin de calculer la valeur-p (*p-value*). À noter, que comme la représentation simplifiée des histogrammes ne comprend que la sommation de tous les points dans le temps, les valeurs-p (*p-value*) calculées sur l'ensemble des points sont plus appropriées.



28. Figure 5.8 L'effet de l'agent chélateur sur le contrôle négatif (un *riboswitch* SAM) dans le milieu M9.

En rajoutant l'agent chélateur l'expression sous le contrôle du *riboswitch* SAM diminue contrairement au cas de *yypP-ykoY*. Colonne 1, C1 : *P. aeruginosa* sans plasmide (aucune luminescence). C2 : *P. aeruginosa* avec le plasmide contient le *riboswitch* SAM sans calcium. C3 : EDTA 1 mM chélate tous les ions dans le milieu M9. C4, C5 : EDTA 1 mM a été compensé en rajoutant 1 mM de calcium et 300 μ M de manganèse respectivement. Le test statistique ONE way ANOVA a été effectué et la valeur p (*p-value*) ≤ 0.05 a été obtenue.



29. Figure 5.9 La normalisation des résultats obtenus par rapport à ceux du *riboswitch* SAM chez *Pseudomonas aeruginosa*.

La normalisation des résultats du gène rapporteur chez *Pseudomonas aeruginosa*. Deux milieux M9 et LB ont été utilisés. Ces résultats montrent une augmentation de l'expression en présence d'EDTA et l'expression diminue en présence du calcium. Le niveau d'expression du manganèse ne change pas ou augmente par rapport à celui d'EDTA.

Les chapitres suivants présentent des perspectives et une discussion sur différents éléments présentés dans cette thèse, incluant ceux concernant les résultats obtenus dans ce chapitre.

6. Discussion

Ce chapitre est la discussion de cette thèse basée sur les résultats obtenus dans les chapitres précédents (le chapitre 5 et les articles). Ce chapitre de discussion générale revoit certains des résultats présentés dans les chapitres précédents avec plus de recul et à la lumière d'autres découvertes tirées de la littérature. Nous ramenons un contexte plus global et nous y réinterprétons certains résultats et émettons de nouvelles hypothèses qui sont discutées.

6.1 L'ARN, les métaux et les origines :

La planète terre est composée de plusieurs éléments y compris les métaux ; par exemple le fer est le plus abondant métal sur la terre (Kadner, 2005), tandis que le molybdène est le plus abondant (100 nM) sur la surface de la mer (Collier, 1984). Deux pourcents des métaux de la terre sont aussi constitués de magnésium (Eisenhauer, 2011). Le zinc est le deuxième métal, le plus abondant dans la mer (Outten et al., 2001) suivi par le magnésium et le calcium (Eisenhauer, 2011). De ce fait, la vie primitive a dû se former en présence des métaux dès les premiers jours. Le zinc et le magnésium sont considérés comme les éléments essentiels pour la vie de même que le manganèse et le fer qui ont des rôles très importants dans la majorité des organismes (Merchant et al., 2012).

Selon la théorie du « RNA world », les êtres vivants utilisaient l'ARN pour tout, incluant comme outil de régulation sans avoir besoin de protéines (Breaker, 2012a). Tous les cations chargés positivement peuvent lier l'ARN de manière non spécifique, mais il y a certains cations qui se lient plus fortement à certains motifs d'ARN et ils sont plus spécifiques (Feig et al., 1999). La découverte des ribozymes (Kruger et al., 1982; Guerrier-Takada et al., 1983) renforce la théorie de « RNA world » (Robertson et al., 2012), sachant que les cations divalents sont nécessaires pour l'activité catalytique du ribozyme (Dahm et al., 1993). Cette découverte suggère que l'ARN doit agir au moins avec certains métaux de manière spécifique.

Les *riboswitchs* sont très spécifiques pour leur ligand et découvrir des nouveaux *riboswitchs* ou des ARNnc aide à comprendre l'évolution des micro-organismes et le rôle de l'ARN dans la régulation (Breaker, 2012a). Les *riboswitchs* associés avec les coenzymes et seconds messagers sont parmi les plus nombreux, ce

qui suggère l'émergence de ces molécules pendant « RNA world » (McCown et al., 2017) ; Ainsi, un des intérêts de découvrir des *riboswitchs* ou des ARNnc spécifiques aux métaux est le fait qu'ils pourraient aider à mieux comprendre la régulation des différents métaux dans le « RNA world », un peu comme des fossiles moléculaires d'une époque précédant le dernier ancêtre universel commun.

6.2 Combien de *riboswitchs* existent-ils sur la terre ?

La question de déterminer le nombre de *riboswitchs* existant, dans l'optique de combien de molécules avaient besoin de régulateurs d'ARN dans la vie à l'époque du « *RNAworld* » (Breaker, 2012a), semble très pertinente. La question pourrait se reformuler ainsi : est-ce que chaque métabolite ou chaque ion a besoin d'un seul ou plusieurs *riboswitchs* ? Pour éclaircir la question on peut d'abord répondre à la question première : existent-ils des exemples de différentes classes de structures de *riboswitchs* pour un seul ligand ? La deuxième question est de savoir si un *riboswitch* pourrait avoir deux ligands ? La réponse pour ces deux questions est un grand OUI.

Il existe au moins cinq classes de structure de *riboswitchs* qui reconnaissent de manière spécifique la molécule S-Adénosylméthionine (SAM) (Breaker, 2011). Il existe aussi des familles de structures de *riboswitchs* spécifiques pour deux molécules différentes tel que nickel et cobalt (Furukawa et al., 2015). Pour les *riboswitchs*, la spécificité de l'*aptamère* pour des ligands est très cruciale. Par exemple il a été montré qu'une seule mutation dans l'*aptamère* du *riboswitch* de l'adénine et de la guanine à distinguer deux métabolites-ligands différents (Mandal et al., 2004b). De même, montrant aussi la grande sélectivité des *riboswitchs*, le *riboswitch* SAM-I est capable de distinguer entre le SAM et le S-Adénosyl-L-Homocystéine (SAH) (Breaker, 2012a). Par contre, le *riboswitch* SAM-SAH ne fait pas de distinction entre ces deux molécules (Breaker, 2012a), indiquant que la sélectivité est sujette à des pressions de sélection qui peuvent varier.

Quant aux *riboswitchs* des métaux chez les bactéries, la problématique est un peu différente. Les métaux, principalement le magnésium, sont nécessaires pour le repliement de l'ARN (Draper, 2004) et le magnésium se retrouve en concentration élevée, de l'ordre de mM (Finney et al., 2003). Il a été montré que d'autres métaux, tels que le manganèse, sont capables de replier le ribozyme *hammerhead* (Boots et al.,

2008). Imaginer un *riboswitch* spécifique pour un métal particulier semble un peu difficile à cause de la charge positive des métaux qui interagissent facilement avec l'ARN négativement chargé. Bien qu'il soit difficile d'imaginer la spécificité d'un *riboswitch* pour le magnésium il existe deux classes distinctes de *riboswitchs* pour le magnésium (Cromie et al., 2006; Dann et al., 2007). Au moins une classe de ces deux *riboswitchs* peut se lier à d'autres ions (Ramesh et al., 2011) mais la concentration élevée du magnésium fait ce dernier son ligand. Plus tard, deux classes distinctes de *riboswitchs* manganèse (Shi et al., 2014; Dambach et al., 2015; Price et al., 2015) et une classe de *riboswitchs* Nickel/Cobalt (Furukawa et al., 2015) ont été découvertes. Le *riboswitch* Nickel/Cobalt pour l'instant est le seul *riboswitch* connu qui est spécifique pour deux ions métalliques différents. En résumé il existerait d'autres *riboswitchs* ou des sous groupes de *riboswitchs* au moins pour les cations qui reconnaîtraient un même ligand (cation) ou plusieurs ligands (cations).

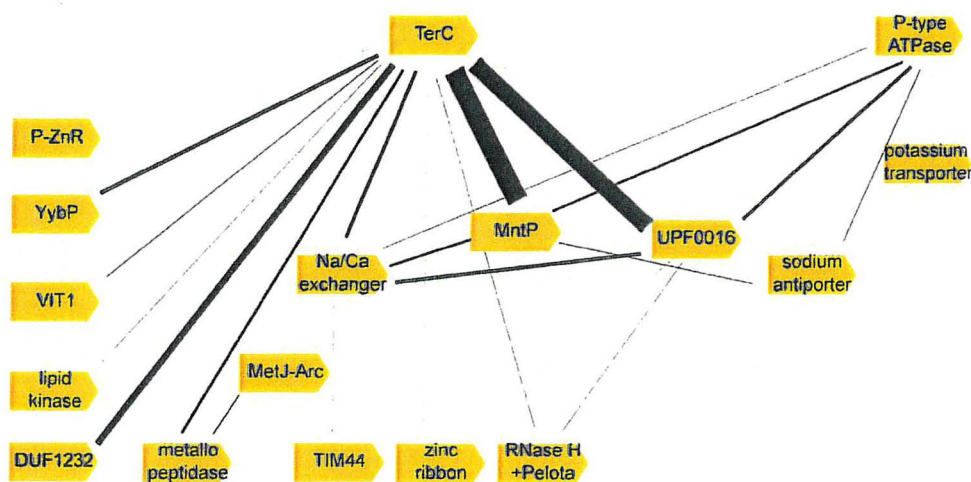
6.3 Le motif *yybP-ykoY* (*riboswitch* manganèse) :

6.3.1 Contexte génétique :

Le motif *yybP* a été identifié pour la première fois par la méthode bioinformatique (Barrick et al., 2004) et pendant longtemps il a été considéré comme un *riboswitch* orphelin (Breaker, 2011), c'est-à-dire sans ligand connu. Ce motif se retrouve dans la partie 5'-UTR du gène *alx* et pourrait réguler l'expression d'*alx* selon la variation du pH (Nechooshtan et al., 2009). Ce motif, *yybP-ykoY* se retrouve également en amont des gènes de la famille TerC (Figure 6.1) et d'ATPase transporteurs de cations (Barrick et al., 2004). Le motif *yybP-ykoY* est connu aujourd'hui comme le *riboswitch* manganèse (Dambach et al., 2015; Price et al., 2015).

La modulation des gènes en aval des motifs *yybP-ykoY* a été rapportée. Nous avons aussi relevé que ce motif est retrouvé en amont de gènes annotés comme des transporteurs de calcium. Par exemple, ce motif se retrouve en amont du gène *yoaB* chez *Lactococcus lactis*, qui est annoté comme *P-type II (calcium-transporting) ATPase* (Price et al., 2015). Il a été suggéré que cette famille d'ATPase a un rôle de protection contre la toxicité des métaux (Chan et al., 2010). D'ailleurs, il a été suggéré que la famille TerC (Figure 6.1) comprend différents groupes de protéines, y compris des *P-*

type ATPase (Dambach et al., 2015). Tout ceci pourrait suggérer un potentiel rôle du *riboswitch yybP-ykoY* comme un senseur d'ions métalliques tels que le calcium ou d'ions toxiques, en plus de son rôle connu de senseur de manganèse (Dambach et al., 2015; Price et al., 2015).

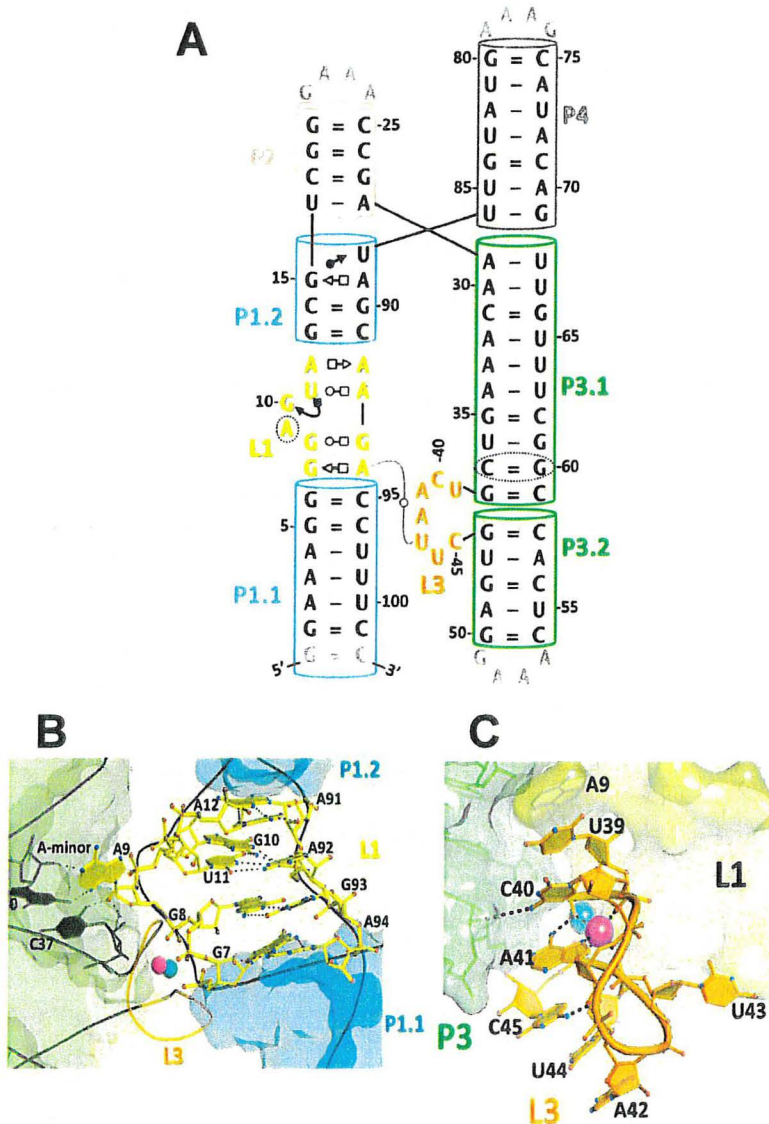


30. Figure 6.1 La famille de réseaux TerC.

Ces gènes sont prédits, impliqués dans la protection contre la toxicité des métaux. Des lignes épaisses ont le motif *yybP-ykoY* dans leurs régions intergéniques. Des groupes P-ZnR (des peptidases avec des rubans de zinc) et MetJ-Arc (*binding DNA proteins*) qui se retrouvent dans l'opéron avec le motif *yybP-ykoY*. Tiré de (Dambach et al., 2015).

6.3.2 La boucle E dans le motif *yybP-ykoY* (riboswitch manganèse) :

Le motif *yybP-ykoY* (Dambach et al., 2015; Price et al., 2015) montre l'existence d'une boucle E (*loop E*) (L1) (Figure 6.2 A) et que l'interaction de cette boucle avec la partie conservée « CA » du motif *yybP-ykoY* (L3) est responsable de la reconnaissance du manganèse, où le nucléotide « A » est spécifique pour le manganèse (Fig 6.2 B et C) (Price et al., 2015).



31. Figure 6.2 L'interaction de la boucle E avec le manganèse dans le motif *yypP-ykoY*.

A) La structure secondaire du *riboswitch yypP-ykoY*. B) La boucle L1 du *riboswitch yypP-ykoY* contient la séquence du motif RAGUA qui avec l'aide de L3 reconnaît le manganèse. Des phosphates des nucléotides G7 et G8 sont orientés vers le site de liaison des métaux, manganèse et magnésium et avec l'aide de A9 qui interagit avec le C37 et G60 de la boucle L3. C) Les position C40 et A41 de L3 ensemble contiennent le métal avec l'interaction de L1. Tiré de (Price et al., 2015).

6.3.3 Le motif « ressemblant à *yybP-ykoY* », (*yybP-ykoY-r*) :

Nous avons prédit par la bioinformatique, avec *Graphclust* (Heyne et al., 2012) un motif qui ressemble et partage en grande partie la structure du motif *yybP-ykoY*. Ce motif se retrouve en amont de gènes annotés ayant un lien direct ou indirect avec le calcium au sein de certaines espèces. Il faut bien noter que le nouveau motif ressemblant à *yybP-ykoY* (candidat pour le *riboswitch* calcium) ne se retrouve pas systématiquement dans toutes les séquences intergéniques prédites par Rfam ayant le motif de *yybP-ykoY* (*riboswitch* manganèse) et vice-versa. Par exemple le motif de Rfam *yybP-ykoY* (RF00080) ne correspond à aucun hit dans la base de données Rfam pour *Desulfovibrio magneticus* RS-1 (NC_012796), bien qu'on y ait trouvé le motif *yybP-ykoY-r*. Le motif *yybP-ykoY* (*riboswitch* manganèse) se retrouve aussi chez les bactéries pathogènes telles que *Pseudomonas aeruginosa*.

6.3.4 Comparaison des motifs *yybP-ykoY* et *yybP-ykoY-r* :

Étant donné que le *riboswitch* *yybP-ykoY* et *yybP-ykoY-r* ont tous les deux la boucle E, cela peut suggérer, du point de vue d'évolution, qu'il y avait une super classe de *riboswitch* pour des métaux capables de reconnaître plusieurs ions métalliques en même temps dans les conditions plus ou moins différentes. Comme discuté par (Feig et al., 1999), la boucle E permet de lier différents métaux y compris le manganèse (Price et al., 2015) le magnésium (Leontis et al., 1986; Correll et al., 1997), le calcium (Leontis et al., 1986), le zinc (Ciesiolka et al., 1995), et le potassium (Auffinger et al., 2004). Les deux nucléotides conservés, « CA », dans la structure de *yybP-ykoY* sont responsables de reconnaître le manganèse, mais le changement de ces nucléotides ou un nucléotide pourrait changer la spécificité pour un autre ion. Par exemple le nucléotide A41 est considéré responsable de la liaison sélective du manganèse dans la pochette de liaison (Price et al., 2015), alors que la substitution de C40U résulte à perdre l'affinité au manganèse dans le motif *yybP-ykoY* (Price et al., 2015).

Considérons les propriétés physico-chimiques et biologiques de chaque ion métallique décrit auparavant pour cette reconnaissance pour mieux comprendre. Par exemple, le manganèse a un rôle dans le stress oxydatif (Anjem et al., 2009) avec des propriétés physico-chimiques plus ou moins comme magnésium (Kehres et al., 2003),

tandis que le calcium est impliqué dans la signalisation (Maguire et al., 2002) avec plusieurs formes de coordinations (Maguire, 2006). Il est intéressant de se rappeler que le manganèse se lie à l'ARN plus fortement que le magnésium, à cause des électrons d (Draper, 2004). Ainsi, il est possible que le motif *yybP-ykoY-r* soit plus flexible au niveau du nucléotide « C » pour détecter différentes formes de calcium, puisque le Ca^{2+} a plus de formes de coordination.

6.3.5 L'effet du pH et le motif *yybP-ykoY* et *yybP-ykoY-r* :

L'effet du pH sur les cations en général devrait aussi être pris en considération. Le gène *alx* chez *E. coli* possède le motif *yybP-ykoY* (Barrick et al., 2004). L'expression de ce gène change lorsque le pH est alcalin (Bingham et al., 1990) et plus tard il a été démontré que l'expression du gène *alx* est sous contrôle de l'ARN en pH alcalin (Nechooshtan et al., 2009). Il est intéressant d'imaginer que ce *riboswitch* pourrait réguler deux ions différents, tels que le manganèse et le calcium, dans deux pH différents sachant qu'il a déjà été montré que le *riboswitch mgtE* se lie au manganèse quatre fois mieux que le magnésium et le calcium et permet aussi le repliement de l'ARN (Ramesh et al., 2011).

6.3.6 Les données transcriptomiques vis-à-vis du motif *yybP-ykoY* :

Le rôle des métaux dans la virulence des bactéries est bien étudié. Par exemple le rôle du magnésium dans la virulence de *Salmonella typhimurium* a été bien démontré (Groisman et al., 2013). L'implication de la régulation du calcium dans la virulence a aussi été bien démontrée (Guragain et al., 2013; Broder et al., 2016). Il est possible que ce motif, *yybP-ykoY-r*, joue un rôle important pendant la régulation du calcium chez *Pseudomonas aeruginosa*. Il existe des données transcriptomiques pour la région intergénique ayant le motif *yybP-ykoY* (*riboswitch* manganèse), qui montrent la variation de l'expression de l'ARN pour cette région. Par exemple, la variation de la température (28 °C et 37 °C) changera l'expression de l'ARN dans la région intergénique du locus-tag : PA4629 chez *Pseudomonas aeruginosa* PA14 (Wurtzel et al., 2012). Il a été montré également que l'expression de l'ARN du locus_tag (UCBPP-PA14, PA14_61260), contenant le *riboswitch* manganèse, a été changée par un facteur ~ 300 dans des échantillons des patients atteint de fibrose kystique (*cystic fibrosis*, CF)

(Rossi et al., 2018). Sachant que le calcium atteint une concentration de 3 à 7 mM dans les poumons de patients atteints de CF et par conséquent cette concentration peut stimuler la virulence chez *Pseudomonas aeruginosa* (Guragain et al., 2016). Il est tentant de supposer que cette différence d'expression serait due au *riboswitch* et à son induction par le calcium. Il est intéressant de savoir que des données transcriptomiques (<http://exploration.weizmann.ac.il/TCOL/>), (Cohen et al., 2016) de *Desulfovibrio vulgaris* str. Hildenborough, (NC_002937, 720681-720371, locus_tag : DVU2617), (He et al., 2010) montrent l'expression du gène « Sodium calcium exchange family » (locus_tag : Des_vul) dans le milieu DSMZ (contient ~ 6 µM de calcium). Cette région intergénique a été prédite pour avoir le motif *yybP-ykoY-r*. Les données mentionnées pour *Pseudomonas aeruginosa* (Rossi et al., 2018), sur l'effet de la concentration élevée de calcium sur la virulence (Sarkisova et al., 2005; Guragain et al., 2016) et celle de *Desulfovibrio vulgaris* str. Hildenborough, NC_002937 (He et al., 2010) sont intéressantes, car elles montrent tout d'abord que la région 5'-UTR ayant le candidat pour le *riboswitch* calcium se transcrit toujours. Ensuite en concentration élevée de calcium (les échantillons de CF) la région codante (CDS) est réprimée, par contre en faible concentration de calcium (le milieu DSMZ) l'expression de la région codante (CDS) subit moins de changements, ce qui correspond bien à nos propres résultats (chapitre 5).

Ces données transcriptomiques et aussi des résultats obtenus par *in-line probing* indiquent que le motif *yybP-ykoY* ou *yybP-ykoY-r* pourrait être aussi un régulateur d'autres ions que le calcium ou le manganèse dans certaines conditions. Il a été suggéré que l'adénosine triphosphate (ATP) joue un rôle dans l'efflux de calcium via des ATPase chez *E. coli* (Naseem et al., 2009). En effet, plusieurs ATPases impliquées dans l'efflux de calcium ont été identifiées telles que : chez *B. subtilis* où une telle ATPase est exprimée pendant la sporulation (Raeymaekers et al., 2002) ; chez *Listeria monocytogenes*, une pompe ATPase avec un seul site de liaison a été identifiée (Faxen et al., 2011) ; chez *Synechococcus* sp. strain PCC 7942 (Berkelman et al., 1994) et chez *Mycobacterium smegmatis* (Gupta et al., 2017), de même que *Pseudomonas aeruginosa* PAO1 (Guragain et al., 2013), les pompes P-type ATPase pourraient aussi être impliquées dans l'acquisition du calcium. Ces données suggèrent que la régulation du calcium pourrait s'effectuer indirectement par d'autres mécanismes

et chercher des séquences intergéniques liées à ces mécanismes pourrait élucider davantage la régulation du calcium.

Il est intéressant de savoir que l'expression du gène *mgtA* est réprimée en présence de calcium chez *E. coli* (Naseem et al., 2009) ce qui suggère qu'il soit possible que le *riboswitch Mg sensor* (Cromie et al., 2006) régulerait aussi le gène en présence du calcium. Il a été montré également que le *riboswitch mgtE* de *Bacillus subtilis* se replie de la même façon avec le calcium et le magnésium (Ramesh et al., 2011). Il a été déjà montré qu'il existe un *riboswitch* qui reconnaît deux ions métalliques en même temps, (le *riboswitch NiCo*) (Furukawa et al., 2015). Il a été montré que le *riboswitch yybP-ykoY* se lie aussi aux autres métaux de transition tel que Ni^{2+} , Co^{2+} et Cd^{2+} et que le manganèse et le cadmium se lient au *riboswitch yybP-ykoY* avec l'heptacoordination (Bachas et al., 2018). Tout cela pourrait renforcer la possibilité d'avoir plusieurs ligands pour un *riboswitch*.

Très récemment il a été montré que le *riboswitch yybP-ykoY* chez *Streptococcus pneumoniae* se retrouve en amont du gène *mgtA* et lierait le calcium avec une affinité 50 fois plus petite (meilleure) que le manganèse (Martin et al., 2019). Il a été suggéré que dans la condition de stress le *riboswitch yybP-ykoY*, chez *Streptococcus pneumoniae*, pourrait réguler le calcium (Martin et al., 2019).

6.4 L'effet de variation des nucléotides sur la structure et la régulation du ribozyme et du *riboswitch* :

Il est intéressant de mentionner que des mutations naturelles pourraient affecter la spécificité pour un ion métallique plus qu'un autre. Le ribozyme *hammerhead* retrouvé dans le phage *Bcep176* a subi une mutation naturelle dans la séquence consensus, le nucléotide A6 → C6 (Perreault et al., 2011). Ce changement naturel du ribozyme *hammerhead* donne une préférence au manganèse plutôt qu'aux autres cations (chapitre 2) (Naghdi et al., 2020). De même, des mutations préservant la structure et changeant la spécificité pour un ion ou bien la vitesse du clivage ont été observées pour d'autres ribozymes (HDV) (Riccitelli et al., 2014). Il a été aussi montré que certaines mutations dans certaines positions du cœur catalytique du ribozyme *hammerhead* peut changer la spécificité du ribozyme pour un autre ligand comme le

cas de *glmS* (Lau et al., 2017) et aussi la mutation qui change la spécificité de Mg^{2+} pour Zn^{2+} (Mir et al., 2015).

7. Conclusion

Au cours de cette thèse, nous avons abordé la découverte et la validation d'ARNnc par des méthodes bioinformatiques et de biologie moléculaire. La découverte d'ARNnc spécifiques aux métaux et aussi l'importance de séquences et de structures de l'ARN afin de reconnaître son ligand (cation par exemple) est bien discutée et montrée. Par exemple nous avons montré que la déviation du consensus de ribozyme *hammerhead* (*Bcep176*) favorise un ion divalent plutôt que d'autres. Certains éléments seront plus élaborés au cours des lignes suivantes :

Vu que des métaux sont essentiels, et que la vie primitive, avait besoin de les utiliser, alors découvrir des nouveaux ARNnc pourrait permettre de mieux connaître la régulation des métaux dans la vie primitive. On peut imaginer qu'il aurait existé une super classe de *riboswitch* responsable pour la régulation de plusieurs métaux.

RiboGap est un outil puissant qui permet de chercher des séquences intergéniques avec la fonction de « pattern matching ». Le motif *yybP-ykoY-r* a été retrouvé à partir des régions intergéniques des procaryotes, en utilisant le mot clé (`[[[:<:]Ca[:>:]]|Calcium`), en utilisant RiboGap (Naghdi et al., 2017). Ce qui est limitant dans ce mode de recherche est l'annotation des bases de données comme NCBI. La recherche par mot clé exclue évidemment d'autres gènes qui n'ont pas été annotés comme tel, mais peut aussi inclure de mauvaises annotations. Cependant RiboGap est un outil puissant pour extraire des séquences intergéniques pour une fonction particulière et en même temps permet de rapidement déterminer la présence d'ARNnc connus déjà annotés. Présumément, les séquences non-pertinentes (mal annotées) seront exclues lors des étapes subséquentes de *clustering* et de prédiction de structures d'ARN conservées s'il n'y a pas de raison fonctionnelle pour avoir une régulation basée sur le calcium.

D'autres méthodes en bioinformatique comme *RNArobo* (Rampásek et al., 2016) permettant de chercher la variation des séquences dans la même structure pourraient être utiles, surtout lorsque la variation de séquence devient cruciale pour favoriser la liaison d'un ligand sur d'autre. Cette méthode a été utilisée auparavant pour

retrouver des cœurs catalytiques de ribozyme *hammerhead* avec une seule variation de séquence.

La méthode *in-line probing* (Regulski et al., 2008b) est couramment utilisée pour étudier le changement de structure de *riboswitchs* en présence et absence des ligands (Nahvi et al., 2002; Winkler et al., 2002b). Malgré cela, cette méthode a ses limitations pour des ligands métalliques. Le tampon du *in-line probing* à un pH de 8.3 ce qui pourrait causer des problèmes pour certains ions comme le manganèse, qui précipite en milieu alcalin. En revanche, changer la condition de pH (pH 7, ou 8) n'a pas affecté significativement le patron de dégradation d'ARN comparativement au pH original (8.3).

Le motif *yybP-ykoY* peut reconnaître et lier le calcium (Martin et al., 2019) et chapitre 5 de cette thèse et des données récentes, incluant ce que nous présentons au chapitre 2, montrent qu'en introduisant différentes mutations l'affinité ou la spécificité du *riboswitch*, *yybP-ykoY* (Bachas et al., 2018) ou du ribozyme *hammerhead* (Mir et al., 2015; Lau et al., 2017; Naghdi et al., 2020) pour leur ligand peut changer. Tout cela renforce l'idée de l'existence d'une superclasse de structures pour un *riboswitch* qui reconnaîtrait différents ligands, mais dont plusieurs sous-classes auraient chacune leur spécificité.

8. Perspectives

Ce chapitre a pour but de suggérer des travaux et des études qui pourraient être effectués pour faire suite au travail présenté dans cette thèse notamment pour le chapitre 5.

8.1 Étudier l'expression du gène rapporteur avec une version mutante non-fonctionnelle du *riboswitch yybP-ykoY-r* :

Bien que l'effet de la régulation du *riboswitch yybP-ykoY-r* soit comparée avec celle du *riboswitch* SAM, il serait plus pertinent d'introduire quelques mutations afin de défaire certaines tiges ou substituer des nucléotides dans la boucle E ou ceux de la boucle 3 (par exemple C54U et A55G, Figure 5.4 D), ce qui permettrait d'obtenir des résultats plus pertinents pour un contrôle négatif.

8.2 L'effet du pH :

L'interaction de la boucle E avec les ions métalliques est un sujet intéressant qui pourrait aider à mieux comprendre comment l'ARN peut reconnaître spécifiquement des ions métalliques et mieux déterminer l'évolution d'ARN vers une fonction particulière. Pour élucider davantage le fonctionnement du motif *yybP-ykoY* dans la régulation et l'effet du pH. Il serait possible de faire l'expérience avec le milieu LB pH (6.8), et LB alcalin, LBK pH (8.4), et comparer l'expression du gène rapporteur (Nechooshtan et al., 2009) en présence des ions tels que le manganèse et calcium. Le milieu LBK : contient du KCl à la place de NaCl, et 100 mM 3-[(1,1-diméthyl-2-hydroxyéthyl) amino]-2-hydroxypropanesulfonic acid (AMPSO) (Amresco) rajouté et ensuite en utilisant le KOH le pH sera ajusté au pH 8.4 (Stancik et al., 2002; Nechooshtan et al., 2009). Cette expérience pourrait aider à clarifier si l'observation précédente d'un lien avec le pH serait indirecte *via* l'effet du pH sur la solubilité du Mn^{2+} par exemple.

8.3 Le motif RAGUA :

En définissant un consensus comme « RAGUA » on peut chercher toutes les séquences intergéniques des bactéries ayant potentiellement une boucle E en utilisant RiboGap. Il serait alors intéressant de regarder aussi si les séquences intergéniques récupérées ont un lien direct ou indirect avec les métaux dans leurs descriptions. Faire la prédiction de la structure secondaire permettrait de retrouver le motif *yybP-ykoY* comme contrôle positif. Choisir un nouveau candidat et le tester pour son interaction avec des métaux pourrait permettre d'approfondir l'hypothèse que la boucle E peut servir d'une « pochette de liaison généraliste » pour les métaux et donc faire partie d'autres *riboswitchs* impliqués dans la régulation des métaux. Il faut savoir que les séquences d'ARN ribosomiaux comme 5S et 23S doivent être éliminés, car elles donneraient beaucoup de « positifs ». Il existe quelques autres exemples d'ARN de *riboswitchs* qui ne sont pas spécifiques aux métaux, mais qui contiennent aussi la boucle E, comme le *riboswitch* Lysine. Il faudrait donc aussi les éliminer avec les ARN ribosomiaux.

8.4 Terminaison « Rho dépendant » *yybP-ykoY P. aeruginosa* :

Il serait intéressant de déterminer le mécanisme de régulation du *riboswitch* *yybP-ykoY* chez *Pseudomonas aeruginosa*. Deux mécanismes différents ont été observés pour différentes versions du *riboswitch* *yybP-ykoY*, la régulation par le terminateur « Rho » indépendant et le mécanisme de l'initiation de la traduction (Dambach et al., 2015; Price et al., 2015). Découvrir un autre mécanisme comme la régulation par le terminateur « Rho » dépendant et la distribution de ce mécanisme au sein d'espèces bactériennes serait intéressant. Sachant que le terminateur « Rho » indépendant n'a pas été prédit dans la séquence intergénique pour le locus-tag PA14_61260 chez *Pseudomonas aeruginosa* PA14, il semble que la régulation pourrait être effectuée *via* le mécanisme « Rho » dépendant chez *Pseudomonas aeruginosa* PA14. De plus, des résultats transcriptomiques (Wurtzel et al., 2012) montrent un arrêt de transcription entre le *riboswitch* et le codon *start*. Étudier l'effet de la bicyclomycin (antibiotique qui affecte la terminaison « Rho » dépendante) (Zwiefka et al., 1993) chez *Pseudomonas aeruginosa* PA14 ou d'autres bactéries moins résistante pourrait aider à élucider le mécanisme de régulation du motif *yybP-ykoY*. Par contre, il faudrait trouver

des souches sensibles à la bicyclomycine. Alternativement, des souches KO (*Knock Out*) pour le gène *rho* pourraient potentiellement fournir des informations similaires.

8.5 Sous famille L3 :

La mutation dans la boucle 3 (L3) du *riboswitch* *yybP-ykoY* change l'affinité pour les ions (Bachas et al., 2018), sachant que la boucle 3 (L3) du *riboswitch* manganèse est variable d'une espèce à l'autre : *L. lactis*, (*UCAAUUC*) (Price et al., 2015). *E. coli* (*ACAUACU*) (Dambach et al., 2015) et *S. pneumoniae*, (*UCAUUC*) (Martin et al., 2019). La boucle L3 dans les candidats retrouvés de *P. aeruginosa* (*ACACACU*) et celle de *D. magneticus* (*ACAUCAC*) sont différentes, ce qui suggère que cette différence pourrait jouer sur la régulation des ions. Pour évaluer cela, des tests préliminaires en faisant du *in-line probing* pour des motifs nommés ci-dessus, en présence d'ions tel que Mn^{2+} et Ca^{2+} , aideraient à comprendre l'importance des positions de chaque nucléotide. Il est pertinent mentionner qu'il y a des précédents pour des sous-classes de *riboswitchs* de structure similaire qui reconnaissent chacune un ligand différent, en particulier le *riboswitch* *ykkC* a été sous groupé selon la reconnaissance des ligands : guanidine, tel que guanosine tetrphosphate (ppGpp), phosphoribosyl pyrophosphate (PRPP), de même que l'adénosine-5'-diphosphate (l'ADP) ou la cytidine-5'-diphosphate (CDP) (Sherlock et al., 2019).

L'homéostasie et le rôle des métaux dans différentes activité vitales et aussi la virulence des bactéries montrent l'importance d'étudier davantage les méthodes qu'ont les bactéries pour la régulation des concentrations intracellulaires de métaux. La régulation du calcium et d'autres ions par le *riboswitch* *yybP-ykoY-r* et les conditions de sélectivité des *riboswitch* présumés entre différents ions reste à élucider.

9. Bibliographie

- Aigner, K., Dreßen, F. and Steger, G. (2012). Methods for Predicting RNA Secondary Structure. 19-41. 10.1007/978-3-642-25740-7_3.
- Alatossava, T., Jutte, H., Kuhn, A. and Kellenberger, E. (1985). Manipulation of intracellular magnesium content in polymyxin B nonapeptide-sensitized *Escherichia coli* by ionophore A23187. *J. Bacteriol.* (162): 413-419.
- Altman, S. (1990). Nobel lecture. Enzymatic cleavage of RNA by RNA. *Biosci. Rep.* (10): 317-337.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* (215): 403-410.
- Ames, T. D. and Breaker, R. R. (2011). Bacterial aptamers that selectively bind glutamine. *RNA Biol.* (8): 82-89.
- Ames, T. D., Rodionov, D. A., Weinberg, Z. and Breaker, R. R. (2010). A eubacterial riboswitch class that senses the coenzyme tetrahydrofolate. *Chem. Biol.* (17): 681-685.
- Andrews, S. C., Robinson, A. K. and Rodriguez-Quinones, F. (2003). Bacterial iron homeostasis. *FEMS Microbiol. Rev.* (27): 215-237.
- Anjem, A., Varghese, S. and Imlay, J. A. (2009). Manganese import is a key element of the OxyR response to hydrogen peroxide in *Escherichia coli*. *Mol. Microbiol.* (72): 844-858.
- Archibald, F. S. and Fridovich, I. (1981). Manganese, superoxide dismutase, and oxygen tolerance in some lactic acid bacteria. *J. Bacteriol.* (146): 928-936.
- Atilho, R. M., Mirihana Arachchilage, G., Greenlee, E. B., Knecht, K. M. and Breaker, R. R. (2019). A bacterial riboswitch class for the thiamin precursor HMP-PP employs a terminator-embedded aptamer. *Elife.* (8):
- Auffinger, P., Bielecki, L. and Westhof, E. (2004). Symmetric K⁺ and Mg²⁺ ion-binding sites in the 5S rRNA loop E inferred from molecular dynamics simulations. *J. Mol. Biol.* (335): 555-571.
- Bachas, S. T. and Ferre-D'Amare, A. R. (2018). Convergent Use of Heptacoordination for Cation Selectivity by RNA and Protein Metalloregulators. *Cell Chem Biol.* (25): 962-973.
- Backofen, R., Amman, F., Costa, F., Findeiss, S., Richter, A. S. and Stadler, P. F. (2014). Bioinformatics of prokaryotic RNAs. *RNA Biol.* (11): 470-483.
- Badelt, S., Hammer, S., Flamm, C. and Hofacker, I. L. (2015). Thermodynamic and kinetic folding of riboswitches. *Meth. Enzymol.* (553): 193-213.
- Bagg, A. and Neilands, J. B. (1987). Ferric uptake regulation protein acts as a repressor, employing iron (II) as a cofactor to bind the operator of an iron transport operon in *Escherichia coli*. *Biochemistry.* (26): 5471-5477.
- Baglivo, I. et al. (2009). The structural role of the zinc ion can be dispensable in prokaryotic zinc-finger domains. *Proc. Natl. Acad. Sci. U.S.A.* (106): 6933-6938.
- Baker, J. L., Sudarsan, N., Weinberg, Z., Roth, A., Stockbridge, R. B. and Breaker, R. R. (2012). Widespread genetic switches and toxicity resistance proteins for fluoride. *Science.* (335): 233-235.
- Barrick, J. E. and Breaker, R. R. (2007). The distributions, mechanisms, and structures

of metabolite-binding riboswitches. *Genome Biol.* (8): R239.

Barrick, J. E. et al. (2004). New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc. Natl. Acad. Sci. U.S.A.* (101): 6421-6426.

Bastet, L., Chauvier, A., Singh, N., Lussier, A., Lamontagne, A. M., Prevost, K., Masse, E., Wade, J. T. and Lafontaine, D. A. (2017). Translational control and Rho-dependent transcription termination are intimately linked in riboswitch regulation. *Nucleic Acids Res.* (45): 7474-7486.

Becher, A. and Schweizer, H. P. (2000). Integration-proficient *Pseudomonas aeruginosa* vectors for isolation of single-copy chromosomal lacZ and lux gene fusions. *BioTechniques*. (29): 948-950.

Benoit, S. L., Miller, E. F. and Maier, R. J. (2013). *Helicobacter pylori* stores nickel to aid its host colonization. *Infect. Immun.* (81): 580-584.

Berkelman, T., Garret-Engele, P. and Hoffman, N. E. (1994). The pacL gene of *Synechococcus* sp. strain PCC 7942 encodes a Ca(2+)-transporting ATPase. *J. Bacteriol.* (176): 4430-4436.

Bernhart, S. H., Hofacker, I. L., Will, S., Gruber, A. R. and Stadler, P. F. (2008). RNAalifold: improved consensus structure prediction for RNA alignments. *BMC Bioinformatics*. (9): 474.

Bilecen, K. and Yildiz, F. H. (2009). Identification of a calcium-controlled negative regulatory system affecting *Vibrio cholerae* biofilm formation. *Environ. Microbiol.* (11): 2015-2029.

Bingham, R. J., Hall, K. S. and Slonczewski, J. L. (1990). Alkaline induction of a novel gene locus, alx, in *Escherichia coli*. *J. Bacteriol.* (172): 2184-2186.

Blaha, D., Arous, S., Bleriot, C., Dorel, C., Mandrand-Berthelot, M. A. and Rodrigue, A. (2011). The *Escherichia coli* metallo-regulator RcnR represses rcnA and rcnR transcription through binding on a shared operator site: Insights into regulatory specificity towards nickel and cobalt. *Biochimie*. (93): 434-439.

Blount, K. F. et al. (2015). Novel riboswitch-binding flavin analog that protects mice against *Clostridium difficile* infection without inhibiting cecal flora. *Antimicrob. Agents Chemother.* (59): 5736-5746.

Boots, J. L., Canny, M. D., Azimi, E. and Pardi, A. (2008). Metal ion specificities for folding and cleavage activity in the *Schistosoma* hammerhead ribozyme. *RNA*. (14): 2212-2222.

Boyd, J., Oza, M. N. and Murphy, J. R. (1990). Molecular cloning and DNA sequence analysis of a diphtheria tox iron-dependent regulatory element (dtxR) from *Corynebacterium diphtheriae*. *Proc. Natl. Acad. Sci. U.S.A.* (87): 5968-5972.

Breaker, R. R. (2011). Prospects for riboswitch discovery and analysis. *Mol. Cell*. (43): 867-879.

Breaker, R. R. (2012b). New insight on the response of bacteria to fluoride. *Caries Res.* (46): 78-81.

Breaker, R. R. (2012a). Riboswitches and the RNA world. *Cold Spring Harb Perspect Biol*. (4):

Broder, U. N., Jaeger, T. and Jenal, U. (2016). LadS is a calcium-responsive kinase that induces acute-to-chronic virulence switch in *Pseudomonas aeruginosa*. *Nat Microbiol*. (2): 16184.

Brunel, C., Romby, P., Westhof, E., Ehresmann, C. and Ehresmann, B. (1991). Three-dimensional model of *Escherichia coli* ribosomal 5 S RNA as deduced from structure probing in solution and computer modeling. *J. Mol. Biol.* (221): 293-308.

Bruni, G. N., Weekley, R. A., Dodd, B. J. T. and Kralj, J. M. (2017). Voltage-gated

calcium flux mediates Escherichia coli mechanosensation. *Proc. Natl. Acad. Sci. U.S.A.* (114): 9445-9450.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. and Madden, T. L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics*. (10): 421.

Capdevila, D. A., Wang, J. and Giedroc, D. P. (2016). Bacterial Strategies to Maintain Zinc Metallostatics at the Host-Pathogen Interface. *J. Biol. Chem.* (291): 20858-20868.

Caron, M. P., Bastet, L., Lussier, A., Simoneau-Roy, M., Masse, E. and Lafontaine, D. A. (2012). Dual-acting riboswitch control of translation initiation and mRNA decay. *Proc. Natl. Acad. Sci. U.S.A.* (109): E3444-3453.

Carter, E. L., Flugga, N., Boer, J. L., Mulrooney, S. B. and Hausinger, R. P. (2009). Interplay of metal ions and urease. *Metallomics*. (1): 207-221.

Cech, T. R. (1990). Nobel lecture. Self-splicing and enzymatic activity of an intervening sequence RNA from Tetrahymena. *Biosci. Rep.* (10): 239-261.

Chan, H. et al. (2010). The p-type ATPase superfamily. *J. Mol. Microbiol. Biotechnol.* (19): 5-104.

Chan, K. G., Priya, K., Chang, C. Y., Abdul Rahman, A. Y., Tee, K. K. and Yin, W. F. (2016). Transcriptome analysis of Pseudomonas aeruginosa PAO1 grown at both body and elevated temperatures. *PeerJ*. (4): e2223.

Chandrangsu, P., Rensing, C. and Helmann, J. D. (2017). Metal homeostasis and resistance in bacteria. *Nat. Rev. Microbiol.* (15): 338-350.

Chauvier, A., Picard-Jean, F., Berger-Dancause, J. C., Bastet, L., Naghdi, M. R., Dube, A., Turcotte, P., Perreault, J. and Lafontaine, D. A. (2017). Transcriptional pausing at the translation start site operates as a critical checkpoint for riboswitch regulation. *Nat Commun*. (8): 13892.

Chivers, P. T. and Sauer, R. T. (2000). Regulation of high affinity nickel uptake in bacteria. Ni²⁺-Dependent interaction of NikR with wild-type and mutant operator sites. *J. Biol. Chem.* (275): 19735-19741.

Chou, A. Y., Archdeacon, J. and Kado, C. I. (1998). Agrobacterium transcriptional regulator Ros is a prokaryotic zinc finger protein that regulates the plant oncogene ipt. *Proc. Natl. Acad. Sci. U.S.A.* (95): 5293-5298.

Choudhary, P. K. and Sigel, R. K. (2014). Mg²⁺-induced conformational changes in the btuB riboswitch from E. coli. *RNA*. (20): 36-45.

Ciesiolka, J., Gorski, J. and Yarus, M. (1995). Selection of an RNA domain that binds Zn²⁺. *RNA*. (1): 538-550.

Cohen, O., Doron, S., Wurtzel, O., Dar, D., Edelheit, S., Karunker, I., Mick, E. and Sorek, R. (2016). Comparative transcriptomics across the prokaryotic tree of life. *Nucleic Acids Res.* (44): 46-53.

Collier, R. W. (1984). Particulate and dissolved vanadium in the North Pacific Ocean. *Nature*. (309): 441-444. 10.1038/309441a0.

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

Coppins, R. L., Hall, K. B. and Groisman, E. A. (2007). The intricate world of riboswitches. *Curr. Opin. Microbiol.* (10): 176-181.

Cornelis, P., Wei, Q., Andrews, S. C. and Vinckx, T. (2011). Iron homeostasis and management of oxidative stress response in bacteria. *Metallomics*. (3): 540-549.

Correll, C. C., Freeborn, B., Moore, P. B. and Steitz, T. A. (1997). Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain. *Cell*. (91): 705-712.

Cromie, M. J., Shi, Y., Latifi, T. and Groisman, E. A. (2006). An RNA sensor for

- intracellular Mg(2+). *Cell*. (125): 71-84.
- Dahm, S. C., Derrick, W. B. and Uhlenbeck, O. C. (1993). Evidence for the role of solvated metal hydroxide in the hammerhead cleavage mechanism. *Biochemistry*. (32): 13040-13045.
- Dambach, M., Sandoval, M., Updegrove, T. B., Anantharaman, V., Aravind, L., Waters, L. S. and Storz, G. (2015). The ubiquitous yybP-ykoY riboswitch is a manganese-responsive regulatory element. *Mol. Cell*. (57): 1099-1109.
- Dann, C. E., Wakeman, C. A., Sieling, C. L., Baker, S. C., Irnov, I. and Winkler, W. C. (2007). Structure and mechanism of a metal-sensing regulatory RNA. *Cell*. (130): 878-892.
- Das, S. and Reusch, R. N. (2001). pH regulates cation selectivity of poly-(R)-3-hydroxybutyrate/polyphosphate channels from *E. coli* in planar lipid bilayers. *Biochemistry*. (40): 2075-2079.
- Deng, J., Xiong, Y. and Sundaralingam, M. (2001). X-ray analysis of an RNA tetraplex (UGGGU)₄ with divalent Sr(2+) ions at subatomic resolution (0.61 Å). *Proc. Natl. Acad. Sci. U.S.A.*. (98): 13665-13670.
- Ding, Y., Chan, C. Y. and Lawrence, C. E. (2005). RNA secondary structure prediction by centroids in a Boltzmann weighted ensemble. *RNA*. (11): 1157-1166.
- Do, C. B., Woods, D. A. and Batzoglou, S. (2006). CONTRAfold: RNA secondary structure prediction without physics-based models. *Bioinformatics*. (22): e90-98.
- Dominguez, D. C. (2004). Calcium signalling in bacteria. *Mol. Microbiol.* (54): 291-297.
- Dowell, R. D. and Eddy, S. R. (2004). Evaluation of several lightweight stochastic context-free grammars for RNA secondary structure prediction. *BMC Bioinformatics*. (5): 71.
- Draper, D. E. (2004). A guide to ions and RNA structure. *RNA*. (10): 335-343.
- Eddy, S. R. (2001). Non-coding RNA genes and the modern RNA world. *Nat. Rev. Genet.* (2): 919-929.
- Eisenhauer, A. (2011). *Encyclopedia of Geobiology*. Springer Netherlands, Dordrecht, 331-336. https://doi.org/10.1007/978-1-4020-9212-1_75
- Eitinger, T., Suhr, J., Moore, L. and Smith, J. A. C. (2005). Secondary Transporters for Nickel and Cobalt Ions: Theme and Variations. *Biometals*. (18): 399-405. [10.1007/s10534-005-3714-x](https://doi.org/10.1007/s10534-005-3714-x).
- Endo, Y., Gluck, A. and Wool, I. G. (1993). Ribosomal RNA identity elements for recognition by ricin and by alpha-sarcin: mutation in the putative CG pair that closes a GAGA tetraloop. *Nucleic Acids Symp. Ser.* 165-166.
- Epshtein, V., Mironov, A. S. and Nudler, E. (2003). The riboswitch-mediated control of sulfur metabolism in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* (100): 5052-5056.
- Farnham, P. J. and Platt, T. (1981). Rho-independent termination: dyad symmetry in DNA causes RNA polymerase to pause during transcription in vitro. *Nucleic Acids Res.* (9): 563-577.
- Faxen, K., Andersen, J. L., Gourdon, P., Fedosova, N., Morth, J. P., Nissen, P. and Møller, J. V. (2011). Characterization of a *Listeria monocytogenes* Ca(2+) pump: a SERCA-type ATPase with only one Ca(2+)-binding site. *J. Biol. Chem.* (286): 1609-1617.
- Feig, A. and Uhlenbeck, O. (1999). 12 The Role of Metal Ions in RNA Biochemistry. *Cold Spring Harbor Monograph Archive*. (37):
- Ferre-D'Amare, A. R. and Winkler, W. C. (2011). The roles of metal ions in regulation by riboswitches. *Met Ions Life Sci*. (9): 141-173.
- Finney, L. A. and O'Halloran, T. V. (2003). Transition metal speciation in the cell: insights from the chemistry of metal ion receptors. *Science*. (300): 931-936.

- Fox, G. E. and Woese, C. R. (1975). The architecture of 5S rRNA and its relation to function. *J. Mol. Evol.* (6): 61-76.
- Frieda, K. L. and Block, S. M. (2012). Direct observation of cotranscriptional folding in an adenine riboswitch. *Science*. (338): 397-400.
- Furtig, B., Nozinovic, S., Reining, A. and Schwalbe, H. (2015). Multiple conformational states of riboswitches fine-tune gene regulation. *Curr. Opin. Struct. Biol.* (30): 112-124.
- Furukawa, K., Ramesh, A., Zhou, Z., Weinberg, Z., Vallery, T., Winkler, W. C. and Breaker, R. R. (2015). Bacterial riboswitches cooperatively bind Ni(2+) or Co(2+) ions and control expression of heavy metal transporters. *Mol. Cell*. (57): 1088-1098.
- Gaballa, A. and Helmann, J. D. (1998). Identification of a zinc-specific metalloregulatory protein, Zur, controlling zinc transport operons in *Bacillus subtilis*. *J. Bacteriol.* (180): 5815-5821.
- Gangola, P. and Rosen, B. P. (1987). Maintenance of intracellular calcium in *Escherichia coli*. *J. Biol. Chem.* (262): 12570-12574.
- Garcia Vescovi, E., Soncini, F. C. and Groisman, E. A. (1996). Mg²⁺ as an extracellular signal: environmental regulation of *Salmonella* virulence. *Cell*. (84): 165-174.
- Gardner, P. P., Wilm, A. and Washietl, S. (2005). A benchmark of multiple sequence alignment programs upon structural RNAs. *Nucleic Acids Res.* (33): 2433-2439.
- Garrison-Schilling, K. L., Grau, B. L., McCarter, K. S., Olivier, B. J., Comeaux, N. E. and Pettis, G. S. (2011). Calcium promotes exopolysaccharide phase variation and biofilm formation of the resulting phase variants in the human pathogen *Vibrio vulnificus*. *Environ. Microbiol.* (13): 643-654.
- Garst, A. D. and Batey, R. T. (2009). A switch in time: detailing the life of a riboswitch. *Biochim. Biophys. Acta*. (1789): 584-591.
- Garst, A. D., Edwards, A. L. and Batey, R. T. (2011). Riboswitches: structures and mechanisms. *Cold Spring Harb Perspect Biol.* (3):
- Ge, R. G., Wang, D. X., Hao, M. C. and Sun, X. S. (2013). Nickel trafficking system responsible for urease maturation in *Helicobacter pylori*. *World J. Gastroenterol.* (19): 8211-8218.
- Gelfand, M. S., Mironov, A. A., Jomantas, J., Kozlov, Y. I. and Perumov, D. A. (1999). A conserved RNA structure element involved in the regulation of bacterial riboflavin synthesis genes. *Trends Genet.* (15): 439-442.
- Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A. and Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*. (6): 343-345.
- Gorodkin, J., Heyer, L. J. and Stormo, G. D. (1997). Finding the most significant common sequence and structure motifs in a set of RNA sequences. *Nucleic Acids Res.* (25): 3724-3732.
- Grilley, D., Soto, A. M. and Draper, D. E. (2006). Mg²⁺-RNA interaction free energies and their relationship to the folding of RNA tertiary structures. *Proc. Natl. Acad. Sci. U.S.A.* (103): 14003-14008.
- Groisman, E. A., Hollands, K., Kriner, M. A., Lee, E. J., Park, S. Y. and Pontes, M. H. (2013). Bacterial Mg²⁺ homeostasis, transport, and virulence. *Annu. Rev. Genet.* (47): 625-646.
- Grundy, F. J., Lehman, S. C. and Henkin, T. M. (2003). The L box regulon: lysine sensing by leader RNAs of bacterial lysine biosynthesis genes. *Proc. Natl. Acad. Sci. U.S.A.* (100): 12057-12062.
- Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. and Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*. (35): 849-857.

- Gupta, H. K., Shrivastava, S. and Sharma, R. (2017). A Novel Calcium Uptake Transporter of Uncharacterized P-Type ATPase Family Supplies Calcium for Cell Surface Integrity in *Mycobacterium smegmatis*. *MBio*. (8):
- Guragain, M., King, M. M., Williamson, K. S., Perez-Osorio, A. C., Akiyama, T., Khanam, S., Patrauchan, M. A. and Franklin, M. J. (2016). The *Pseudomonas aeruginosa* PAO1 Two-Component Regulator CarSR Regulates Calcium Homeostasis and Calcium-Induced Virulence Factor Production through Its Regulatory Targets CarO and CarP. *J. Bacteriol.* (198): 951-963.
- Guragain, M., Lenaburg, D. L., Moore, F. S., Reutlinger, I. and Patrauchan, M. A. (2013). Calcium homeostasis in *Pseudomonas aeruginosa* requires multiple transporters and modulates swarming motility. *Cell Calcium*. (54): 350-361.
- Hamada, M. (2014). Fighting against uncertainty: an essential issue in bioinformatics. *Brief. Bioinformatics*. (15): 748-767.
- Hamada, M. (2015). RNA secondary structure prediction from multi-aligned sequences. *Methods Mol. Biol.* (1269): 17-38.
- Hammann, C., Luptak, A., Perreault, J. and de la Pena, M. (2012). The ubiquitous hammerhead ribozyme. *RNA*. (18): 871-885.
- Harmanci, A. O., Sharma, G. and Mathews, D. H. (2008). PARTS: probabilistic alignment for RNA joint secondary structure prediction. *Nucleic Acids Res.* (36): 2406-2417.
- Havgaard, J. H. and Gorodkin, J. (2014). RNA structural alignments, part I: Sankoff-based approaches for structural alignments. *Methods Mol. Biol.* (1097): 275-290.
- Havgaard, J. H., Torarinsson, E. and Gorodkin, J. (2007). Fast pairwise structural RNA alignments by pruning of the dynamical programming matrix. *PLoS Comput. Biol.* (3): 1896-1908.
- Hawtin, P. R., Delves, H. T. and Newell, D. G. (1991). The demonstration of nickel in the urease of *Helicobacter pylori* by atomic absorption spectroscopy. *FEMS Microbiol. Lett.* (61): 51-54.
- He, S. et al. (2010). Validation of two ribosomal RNA removal methods for microbial metatranscriptomics. *Nat. Methods*. (7): 807-812.
- Helmann, J. D. (2014). Specificity of metal sensing: iron and manganese homeostasis in *Bacillus subtilis*. *J. Biol. Chem.* (289): 28112-28120.
- Helmling, C., Klotzner, D. P., Sochor, F., Mooney, R. A., Wacker, A., Landick, R., Furtig, B., Heckel, A. and Schwalbe, H. (2018). Life times of metastable states guide regulatory signaling in transcriptional riboswitches. *Nat Commun*. (9): 944.
- Helmling, C., Wacker, A., Wolfinger, M. T., Hofacker, I. L., Hengesbach, M., Furtig, B. and Schwalbe, H. (2017). NMR Structural Profiling of Transcriptional Intermediates Reveals Riboswitch Regulation by Metastable RNA Conformations. *J. Am. Chem. Soc.* (139): 2647-2656.
- Hendrix, D. K., Brenner, S. E. and Holbrook, S. R. (2005). RNA structural motifs: building blocks of a modular biomolecule. *Q. Rev. Biophys.* (38): 221-243.
- Henkin, T. M. and Grundy, F. J. (2006). Sensing metabolic signals with nascent RNA transcripts: the T box and S box riboswitches as paradigms. *Cold Spring Harb. Symp. Quant. Biol.* (71): 231-237.
- Herbaud, M. L., Guiseppi, A., Denizot, F., Haiech, J. and Kilhoffer, M. C. (1998). Calcium signalling in *Bacillus subtilis*. *Biochim. Biophys. Acta*. (1448): 212-226.
- Hertel, K. J., Herschlag, D. and Uhlenbeck, O. C. (1994). A kinetic and thermodynamic framework for the hammerhead ribozyme reaction. *Biochemistry*. (33): 3374-3385.
- Heyne, S., Costa, F., Rose, D. and Backofen, R. (2012). GraphClust: alignment-free

structural clustering of local RNA secondary structures. *Bioinformatics*. (28): i224-232.

Hiron, A., Posteraro, B., Carriere, M., Remy, L., Delporte, C., La Sorda, M., Sanguinetti, M., Juillard, V. and Borezee-Durant, E. (2010). A nickel ABC-transporter of *Staphylococcus aureus* is involved in urinary tract infection. *Mol. Microbiol.* (77): 1246-1260.

Hofacker, I. L. (2003). Vienna RNA secondary structure server. *Nucleic Acids Res.* (31): 3429-3431.

Hofacker, I. L., Fekete, M. and Stadler, P. F. (2002). Secondary structure prediction for aligned RNA sequences. *J. Mol. Biol.* (319): 1059-1066.

Hollands, K., Proshkin, S., Sklyarova, S., Epshtein, V., Mironov, A., Nudler, E. and Groisman, E. A. (2012). Riboswitch control of Rho-dependent transcription termination. *Proc. Natl. Acad. Sci. U.S.A.* (109): 5376-5381.

Hood, G., Ramachandran, V., East, A. K., Downie, J. A. and Poole, P. S. (2017). Manganese transport is essential for N₂ -fixation by *Rhizobium leguminosarum* in bacteroids from galeoid but not phaseoloid nodules. *Environ. Microbiol.* (19): 2715-2726.

Hu, H. Q., Johnson, R. C., Merrell, D. S. and Maroney, M. J. (2017). Nickel Ligation of the N-Terminal Amine of HypA Is Required for Urease Maturation in *Helicobacter pylori*. *Biochemistry*. (56): 1105-1116.

Huang, X., Shin, J. H., Pinochet-Barros, A., Su, T. T. and Helmann, J. D. (2017). *Bacillus subtilis* MntR coordinates the transcriptional regulation of manganese uptake and efflux systems. *Mol. Microbiol.* (103): 253-268.

Imazawa, R., Takahashi, Y., Aoki, W., Sano, M. and Ito, M. (2016). A novel type bacterial flagellar motor that can use divalent cations as a coupling ion. *Sci Rep.* (6): 19773.

Ippolito, J. A. and Steitz, T. A. (1998). A 1.3-Å resolution crystal structure of the HIV-1 trans-activation response region RNA stem reveals a metal ion-dependent bulge conformation. *Proc. Natl. Acad. Sci. U.S.A.* (95): 9819-9824.

Iwig, J. S., Leitch, S., Herbst, R. W., Maroney, M. J. and Chivers, P. T. (2008). Ni(II) and Co(II) sensing by *Escherichia coli* RcnR. *J. Am. Chem. Soc.* (130): 7592-7606.

Juttukonda, L. J. and Skaar, E. P. (2015). Manganese homeostasis and utilization in pathogenic bacteria. *Mol. Microbiol.* (97): 216-228.

Kadner, R. J. (2005). Regulation by iron: RNA rules the rust. *J. Bacteriol.* (187): 6870-6873.

Karkare, S., Yousafzai, F., Mitchenall, L. A. and Maxwell, A. (2012). The role of Ca²⁺ in the activity of *Mycobacterium tuberculosis* DNA gyrase. *Nucleic Acids Res.* (40): 9774-9787.

Kehres, D. G. and Maguire, M. E. (2003). Emerging themes in manganese transport, biochemistry and pathogenesis in bacteria. *FEMS Microbiol. Rev.* (27): 263-290.

Kim, J. N. and Breaker, R. R. (2008). Purine sensing by riboswitches. *Biol. Cell.* (100): 1-11.

Kim, J. N., Roth, A. and Breaker, R. R. (2007). Guanine riboswitch variants from *Mesoplasma florum* selectively recognize 2'-deoxyguanosine. *Proc. Natl. Acad. Sci. U.S.A.* (104): 16092-16097.

Kim, P. B., Nelson, J. W. and Breaker, R. R. (2015). An ancient riboswitch class in bacteria regulates purine biosynthesis and one-carbon metabolism. *Mol. Cell.* (57): 317-328.

Klein, D. J. and Ferre-D'Amare, A. R. (2006). Structural basis of glmS ribozyme activation by glucosamine-6-phosphate. *Science*. (313): 1752-1756.

- Kobayashi, M. and Shimizu, S. (1999). Cobalt proteins. *Eur. J. Biochem.* (261): 1-9.
- Kolev, S. K., Petkov, P. S., Rangelov, M. A., Trifonov, D. V., Milenov, T. I. and Vayssilov, G. N. (2018). Interaction of Na⁺, K⁺, Mg²⁺ and Ca²⁺ counter cations with RNA. *Metallomics*. (10): 659-678.
- Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E. and Cech, T. R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell*. (31): 147-157.
- Lau, M. W., Trachman, R. J. and Ferre-D'Amare, A. R. (2017). A divalent cation-dependent variant of the glmS ribozyme with stringent Ca²⁺ selectivity co-opts a preexisting nonspecific metal ion-binding site. *RNA*. (23): 355-364.
- Lee, E. R., Baker, J. L., Weinberg, Z., Sudarsan, N. and Breaker, R. R. (2010). An allosteric self-splicing ribozyme triggered by a bacterial second messenger. *Science*. (329): 845-848.
- Lemay, J. F., Penedo, J. C., Tremblay, R., Lilley, D. M. and Lafontaine, D. A. (2006). Folding of the adenine riboswitch. *Chem. Biol.* (13): 857-868.
- Leontis, N. B., Ghosh, P. and Moore, P. B. (1986). Effect of magnesium ion on the structure of the 5S RNA from Escherichia coli. An imino proton magnetic resonance study of the helix I, IV, and V regions of the molecule. *Biochemistry*. (25): 7386-7392.
- Leontis, N. B. and Westhof, E. (1998). The 5S rRNA loop E: chemical probing and phylogenetic data versus crystal structure. *RNA*. (4): 1134-1153.
- Li, S. and Breaker, R. R. (2017). Identification of 15 candidate structured noncoding RNA motifs in fungi by comparative genomics. *BMC Genomics*. (18): 785.
- Li, S., Hwang, X. Y., Stav, S. and Breaker, R. R. (2016). The yjdF riboswitch candidate regulates gene expression by binding diverse azaaromatic compounds. *RNA*. (22): 530-541.
- Li, Y., Zhong, C. and Zhang, S. (2014). Finding consensus stable local optimal structures for aligned RNA sequences and its application to discovering riboswitch elements. *Int J Bioinform Res Appl*. (10): 498-518.
- Lippa, G. M., Liberman, J. A., Jenkins, J. L., Krucinska, J., Salim, M. and Wedekind, J. E. (2012). Crystallographic analysis of small ribozymes and riboswitches. *Methods Mol. Biol.* (848): 159-184.
- Loh, E. et al. (2009). A trans-acting riboswitch controls expression of the virulence regulator PrfA in Listeria monocytogenes. *Cell*. (139): 770-779.
- Lorenz, R., Wolfinger, M. T., Tanzer, A. and Hofacker, I. L. (2016). Predicting RNA secondary structures from sequence and probing data. *Methods*. (103): 86-98.
- Maguire, M. E. (2006). Magnesium transporters: properties, regulation and structure. *Front. Biosci.* (11): 3149-3163.
- Maguire, M. E. and Cowan, J. A. (2002). Magnesium chemistry and biochemistry. *Biomaterials*. (15): 203-210.
- Makthal, N. and Kumaraswami, M. (2017). Zinc'ing it out: zinc homeostasis mechanisms and their impact on the pathogenesis of human pathogen group A streptococcus. *Metallomics*. (9): 1693-1702.
- Mandal, M., Boese, B., Barrick, J. E., Winkler, W. C. and Breaker, R. R. (2003). Riboswitches control fundamental biochemical pathways in Bacillus subtilis and other bacteria. *Cell*. (113): 577-586.
- Mandal, M. and Breaker, R. R. (2004b). Adenine riboswitches and gene activation by disruption of a transcription terminator. *Nat. Struct. Mol. Biol.* (11): 29-35.
- Mandal, M. and Breaker, R. R. (2004a). Gene regulation by riboswitches. *Nat. Rev. Mol. Cell Biol.* (5): 451-463.

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

Martin, J. E., Le, M. T., Bhattarai, N., Capdevila, D. A., Shen, J., Winkler, M. E. and Giedroc, D. P. (2019). A Mn-sensing riboswitch activates expression of a Mn²⁺/Ca²⁺ ATPase transporter in *Streptococcus*. *Nucleic Acids Res.* (47): 6885-6899.

Martin, J. E., Waters, L. S., Storz, G. and Imlay, J. A. (2015). The *Escherichia coli* small protein MntS and exporter MntP optimize the intracellular concentration of manganese. *PLoS Genet.* (11): e1004977.

Masse, E., Salvail, H., Desnoyers, G. and Arguin, M. (2007). Small RNAs controlling iron metabolism. *Curr. Opin. Microbiol.* (10): 140-145.

Masse, E., Vanderpool, C. K. and Gottesman, S. (2005). Effect of RyhB small RNA on global iron use in *Escherichia coli*. *J. Bacteriol.* (187): 6962-6971.

Mathews, D. H., Moss, W. N. and Turner, D. H. (2010). Folding and finding RNA secondary structure. *Cold Spring Harb Perspect Biol.* (2): a003665.

McCaskill, J. S. (1990). The equilibrium partition function and base pair binding probabilities for RNA secondary structure. *Biopolymers*. (29): 1105-1119.

McCluskey, K., Boudreault, J., St-Pierre, P., Perez-Gonzalez, C., Chauvier, A., Rizzi, A., Beauregard, P. B., Lafontaine, D. A. and Penedo, J. C. (2019). Unprecedented tunability of riboswitch structure and regulatory function by sub-millimolar variations in physiological Mg²⁺. *Nucleic Acids Res.* (47): 6478-6487.

McCown, P. J., Corbino, K. A., Stav, S., Sherlock, M. E. and Breaker, R. R. (2017). Riboswitch diversity and distribution. *RNA*. (23): 995-1011.

McDaniel, B. A., Grundy, F. J., Artsimovitch, I. and Henkin, T. M. (2003). Transcription termination control of the S box system: direct measurement of S-adenosylmethionine by the leader RNA. *Proc. Natl. Acad. Sci. U.S.A.* (100): 3083-3088.

Mellin, J. R., Tiensuu, T., Becavin, C., Gouin, E., Johansson, J. and Cossart, P. (2013). A riboswitch-regulated antisense RNA in *Listeria monocytogenes*. *Proc. Natl. Acad. Sci. U.S.A.* (110): 13132-13137.

Merchant, S. S. and Helmann, J. D. (2012). Elemental economy: microbial strategies for optimizing growth in the face of nutrient limitation. *Adv. Microb. Physiol.* (60): 91-210.

Meyer, M. M., Roth, A., Chervin, S. M., Garcia, G. A. and Breaker, R. R. (2008). Confirmation of a second natural preQ1 aptamer class in Streptococcaceae bacteria. *RNA*. (14): 685-695.

Mir, A., Chen, J., Robinson, K., Lendy, E., Goodman, J., Neau, D. and Golden, B. L. (2015). Two Divalent Metal Ions and Conformational Changes Play Roles in the Hammerhead Ribozyme Cleavage Reaction. *Biochemistry*. (54): 6369-6381.

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

Misra, V. K. and Draper, D. E. (2001). A thermodynamic framework for Mg²⁺ binding to RNA. *Proc. Natl. Acad. Sci. U.S.A.* (98): 12456-12461.

Mizuno, T., Chou, M. Y. and Inouye, M. (1984). A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proc. Natl. Acad. Sci. U.S.A.* (81): 1966-1970.

Mulrooney, S. B. and Hausinger, R. P. (2003). Nickel uptake and utilization by microorganisms. *FEMS Microbiol. Rev.* (27): 239-261.

Murray, J. B., Seyhan, A. A., Walter, N. G., Burke, J. M. and Scott, W. G. (1998). The hammerhead, hairpin and VS ribozymes are catalytically proficient in monovalent cations

alone. *Chem. Biol.* (5): 587-595.

Naghdi, M. R., Boutet, E., Mucha, C., Ouellet, J. and Perreault, J. (2020). Single Mutation in Hammerhead Ribozyme Favors Cleavage Activity with Manganese over Magnesium. *Noncoding RNA*. (6):

Naghdi, M. R., Smail, K., Wang, J. X., Wade, F., Breaker, R. R. and Perreault, J. (2017). Search for 5'-leader regulatory RNA structures based on gene annotation aided by the RiboGap database. *Methods*. (117): 3-13.

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

Naseem, R., Davies, S. R., Jones, H., Wann, K. T., Holland, I. B. and Campbell, A. K. (2007). Cytosolic Ca²⁺ regulates protein expression in *E. coli* through release from inclusion bodies. *Biochem. Biophys. Res. Commun.* (360): 33-39.

Naseem, R., Wann, K. T., Holland, I. B. and Campbell, A. K. (2009). ATP regulates calcium efflux and growth in *E. coli*. *J. Mol. Biol.* (391): 42-56.

Navarro, C., Wu, L. F. and Mandrand-Berthelot, M. A. (1993). The *nik* operon of *Escherichia coli* encodes a periplasmic binding-protein-dependent transport system for nickel. *Mol. Microbiol.* (9): 1181-1191.

Nawrocki, E. P. and Eddy, S. R. (2013). Computational identification of functional RNA homologs in metagenomic data. *RNA Biol.* (10): 1170-1179.

Nechooshtan, G., Elgrably-Weiss, M., Sheaffer, A., Westhof, E. and Altuvia, S. (2009). A pH-responsive riboregulator. *Genes Dev.* (23): 2650-2662.

Nelson, J. W., Atilho, R. M., Sherlock, M. E., Stockbridge, R. B. and Breaker, R. R. (2017). Metabolism of Free Guanidine in Bacteria Is Regulated by a Widespread Riboswitch Class. *Mol. Cell*. (65): 220-230.

Nelson, J. W., Sudarsan, N., Furukawa, K., Weinberg, Z., Wang, J. X. and Breaker, R. R. (2013). Riboswitches in eubacteria sense the second messenger c-di-AMP. *Nat. Chem. Biol.* (9): 834-839.

Nelson, J. W., Sudarsan, N., Phillips, G. E., Stav, S., Lunse, C. E., McCown, P. J. and Breaker, R. R. (2015). Control of bacterial exoelectrogenesis by c-AMP-GMP. *Proc. Natl. Acad. Sci. U.S.A.* (112): 5389-5394.

Niegowski, D. and Eshaghi, S. (2007). The CorA family: structure and function revisited. *Cell. Mol. Life Sci.* (64): 2564-2574.

Nies, D. H. (1995). The cobalt, zinc, and cadmium efflux system CzcABC from *Alcaligenes eutrophus* functions as a cation-proton antiporter in *Escherichia coli*. *J. Bacteriol.* (177): 2707-2712.

Noeske, J., Richter, C., Stiral, E., Schwalbe, H. and Wohnert, J. (2006). Phosphate-group recognition by the aptamer domain of the thiamine pyrophosphate sensing riboswitch. *Chembiochem*. (7): 1451-1456.

Noeske, J., Schwalbe, H. and Wohnert, J. (2007). Metal-ion binding and metal-ion induced folding of the adenine-sensing riboswitch aptamer domain. *Nucleic Acids Res.* (35): 5262-5273.

Norris, V., Grant, S., Freestone, P., Canvin, J., Sheikh, F. N., Toth, I., Trinei, M., Modha, K. and Norman, R. I. (1996). Calcium signalling in bacteria. *J. Bacteriol.* (178): 3677-3682.

Nussinov, R. and Jacobson, A. B. (1980). Fast algorithm for predicting the secondary structure of single-stranded RNA. *Proc. Natl. Acad. Sci. U.S.A.* (77): 6309-6313.

Olson, J. W., Mehta, N. S. and Maier, R. J. (2001). Requirement of nickel metabolism proteins HypA and HypB for full activity of both hydrogenase and urease in *Helicobacter pylori*. *Mol. Microbiol.* (39): 176-182.

- Osman, D. and Cavet, J. S. (2010). Bacterial metal-sensing proteins exemplified by ArsR-SmtB family repressors. *Nat Prod Rep.* (27): 668-680.
- Outten, C. E. and O'Halloran, T. V. (2001). Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science.* (292): 2488-2492.
- Pallan, P. S., Marshall, W. S., Harp, J., Jewett, F. C., Wawrzak, Z., Brown, B. A., Rich, A. and Egli, M. (2005). Crystal structure of a luteoviral RNA pseudoknot and model for a minimal ribosomal frameshifting motif. *Biochemistry.* (44): 11315-11322.
- Palmer, L. D. and Skaar, E. P. (2016). Transition Metals and Virulence in Bacteria. *Annu. Rev. Genet.* (50): 67-91.
- Papp-Wallace, K. M. and Maguire, M. E. (2006). Manganese transport and the role of manganese in virulence. *Annu. Rev. Microbiol.* (60): 187-209.
- Patzer, S. I. and Hantke, K. (1998). The ZnuABC high-affinity zinc uptake system and its regulator Zur in *Escherichia coli*. *Mol. Microbiol.* (28): 1199-1210.
- de la Pena, M. and Garcia-Robles, I. (2010). Ubiquitous presence of the hammerhead ribozyme motif along the tree of life. *RNA.* (16): 1943-1950.
- de la Pena, M., Garcia-Robles, I. and Cervera, A. (2017). The Hammerhead Ribozyme: A Long History for a Short RNA. *Molecules.* (22):
- Perdrizet, G. A., Artsimovitch, I., Furman, R., Sosnick, T. R. and Pan, T. (2012). Transcriptional pausing coordinates folding of the aptamer domain and the expression platform of a riboswitch. *Proc. Natl. Acad. Sci. U.S.A.* (109): 3323-3328.
- Perreault, J., Weinberg, Z., Roth, A., Popescu, O., Chartrand, P., Ferbeyre, G. and Breaker, R. R. (2011). Identification of hammerhead ribozymes in all domains of life reveals novel structural variations. *PLoS Comput. Biol.* (7): e1002031.
- Perry, R. D., Bobrov, A. G. and Fetherston, J. D. (2015). The role of transition metal transporters for iron, zinc, manganese, and copper in the pathogenesis of *Yersinia pestis*. *Metallomics.* (7): 965-978.
- Pi, H., Patel, S. J., Arguello, J. M. and Helmann, J. D. (2016). The *Listeria monocytogenes* Fur-regulated virulence protein FrvA is an Fe(II) efflux P1B4 -type ATPase. *Mol. Microbiol.* (100): 1066-1079.
- de Pina, K., Navarro, C., McWalter, L., Boxer, D. H., Price, N. C., Kelly, S. M., Mandrand-Berthelot, M. A. and Wu, L. F. (1995). Purification and characterization of the periplasmic nickel-binding protein NikA of *Escherichia coli* K12. *Eur. J. Biochem.* (227): 857-865.
- Price, I. R., Gaballa, A., Ding, F., Helmann, J. D. and Ke, A. (2015). Mn(2+)-sensing mechanisms of yybP-ykoY orphan riboswitches. *Mol. Cell.* (57): 1110-1123.
- Prody, G. A., Bakos, J. T., Buzayan, J. M., Schneider, I. R. and Bruening, G. (1986). Autolytic processing of dimeric plant virus satellite RNA. *Science.* (231): 1577-1580.
- Que, Q. and Helmann, J. D. (2000). Manganese homeostasis in *Bacillus subtilis* is regulated by MntR, a bifunctional regulator related to the diphtheria toxin repressor family of proteins. *Mol. Microbiol.* (35): 1454-1468.
- Raeymaekers, L., Wuytack, E., Willems, I., Michiels, C. W. and Wuytack, F. (2002). Expression of a P-type Ca(2+)-transport ATPase in *Bacillus subtilis* during sporulation. *Cell Calcium.* (32): 93.
- Ragsdale, S. W. and Kumar, M. (1996). Nickel-Containing Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase(.). *Chem. Rev.* (96): 2515-2540.
- Ramesh, A., Wakeman, C. A. and Winkler, W. C. (2011). Insights into metalloregulation by M-box riboswitch RNAs via structural analysis of manganese-bound complexes. *J. Mol. Biol.* (407): 556-570.
- Rampášek, L., Jimenez, R. M., Lupták, A., Vinař, T. and Brejová, B. (2016). RNA motif

search with data-driven element ordering. *BMC Bioinformatics*. (17): 216.

Regulski, E. E. and Breaker, R. R. (2008b). In-line probing analysis of riboswitches. *Methods Mol. Biol.* (419): 53-67.

Regulski, E. E., Moy, R. H., Weinberg, Z., Barrick, J. E., Yao, Z., Ruzzo, W. L. and Breaker, R. R. (2008a). A widespread riboswitch candidate that controls bacterial genes involved in molybdenum cofactor and tungsten cofactor metabolism. *Mol. Microbiol.* (68): 918-932.

Riccitelli, N. J., Delwart, E. and Luptak, A. (2014). Identification of minimal HDV-like ribozymes with unique divalent metal ion dependence in the human microbiome. *Biochemistry*. (53): 1616-1626.

Rieder, R., Lang, K., Graber, D. and Micura, R. (2007). Ligand-induced folding of the adenosine deaminase A-riboswitch and implications on riboswitch translational control. *ChemBiochem*. (8): 896-902.

Rieder, U., Kreutz, C. and Micura, R. (2010). Folding of a transcriptionally acting preQ1 riboswitch. *Proc. Natl. Acad. Sci. U.S.A.* (107): 10804-10809.

Rivas, E. and Eddy, S. R. (2001a). Noncoding RNA gene detection using comparative sequence analysis. *BMC Bioinformatics*. (2): 8.

Rivas, E., Klein, R. J., Jones, T. A. and Eddy, S. R. (2001b). Computational identification of noncoding RNAs in *E. coli* by comparative genomics. *Curr. Biol.* (11): 1369-1373.

Robertson, M. P. and Joyce, G. F. (2012). The origins of the RNA world. *Cold Spring Harb Perspect Biol*. (4):

Rodionov, D. A., Hebbeln, P., Gelfand, M. S. and Eitinger, T. (2006). Comparative and functional genomic analysis of prokaryotic nickel and cobalt uptake transporters: evidence for a novel group of ATP-binding cassette transporters. *J. Bacteriol.* (188): 317-327.

Romer, R. and Hach, R. (1975). tRNA conformation and magnesium binding. A study of a yeast phenylalanine-specific tRNA by a fluorescent indicator and differential melting curves. *Eur. J. Biochem.* (55): 271-284.

Rosch, J. W., Sublett, J., Gao, G., Wang, Y. D. and Tuomanen, E. I. (2008). Calcium efflux is essential for bacterial survival in the eukaryotic host. *Mol. Microbiol.* (70): 435-444.

Rossi, E., Falcone, M., Molin, S. and Johansen, H. K. (2018). High-resolution in situ transcriptomics of *Pseudomonas aeruginosa* unveils genotype independent patho-phenotypes in cystic fibrosis lungs. *Nat Commun*. (9): 3459.

Roth, A. et al. (2007). A riboswitch selective for the queuosine precursor preQ1 contains an unusually small aptamer domain. *Nat. Struct. Mol. Biol.* (14): 308-317.

Ruzzo, W. L. and Gorodkin, J. (2014). De novo discovery of structured ncRNA motifs in genomic sequences. *Methods Mol. Biol.* (1097): 303-318.

Sankoff, D. (1985). Simultaneous Solution of the RNA Folding, Alignment and Protosequence Problems. *SIAM Journal on Applied Mathematics*. (45): 810-825. 10.1137/0145048.

Sarkisova, S., Patrauchan, M. A., Berglund, D., Nivens, D. E. and Franklin, M. J. (2005). Calcium-induced virulence factors associated with the extracellular matrix of mucoid *Pseudomonas aeruginosa* biofilms. *J. Bacteriol.* (187): 4327-4337.

Schaedler, T. A., Tong, Z. and van Veen, H. W. (2012). The multidrug transporter LmrP protein mediates selective calcium efflux. *J. Biol. Chem.* (287): 27682-27690.

Schmitt, M. P. and Holmes, R. K. (1991). Iron-dependent regulation of diphtheria toxin and siderophore expression by the cloned *Corynebacterium diphtheriae* repressor gene dtxR in *C. diphtheriae* C7 strains. *Infect. Immun.* (59): 1899-1904.

- Schwalbe, H., Buck, J., Furtig, B., Noeske, J. and Wohnert, J. (2007). Structures of RNA switches: insight into molecular recognition and tertiary structure. *Angew. Chem. Int. Ed. Engl.* (46): 1212-1219.
- Seetin, M. G. and Mathews, D. H. (2012b). RNA structure prediction: an overview of methods. *Methods Mol. Biol.* (905): 99-122.
- Seetin, M. G. and Mathews, D. H. (2012a). TurboKnot: rapid prediction of conserved RNA secondary structures including pseudoknots. *Bioinformatics*. (28): 792-798.
- Serganov, A. and Nudler, E. (2013). A decade of riboswitches. *Cell*. (152): 17-24.
- Serganov, A. and Patel, D. J. (2008). Towards deciphering the principles underlying an mRNA recognition code. *Curr. Opin. Struct. Biol.* (18): 120-129.
- Sherlock, M. E., Sadeeshkumar, H. and Breaker, R. R. (2019). Variant Bacterial Riboswitches Associated with Nucleotide Hydrolase Genes Sense Nucleoside Diphosphates. *Biochemistry*. (58): 401-410.
- Sherlock, M. E., Sudarsan, N. and Breaker, R. R. (2018). Riboswitches for the alarmone ppGpp expand the collection of RNA-based signaling systems. *Proc. Natl. Acad. Sci. U.S.A.* (115): 6052-6057.
- Shi, H. and Moore, P. B. (2000). The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: a classic structure revisited. *RNA*. (6): 1091-1105.
- Shi, Y., Zhao, G. and Kong, W. (2014). Genetic analysis of riboswitch-mediated transcriptional regulation responding to Mn²⁺ in *Salmonella*. *J. Biol. Chem.* (289): 11353-11366.
- Shine, J. and Dalgarno, L. (1975). Determinant of cistron specificity in bacterial ribosomes. *Nature*. (254): 34-38.
- Smith, R. L., Banks, J. L., Snavely, M. D. and Maguire, M. E. (1993). Sequence and topology of the CorA magnesium transport systems of *Salmonella typhimurium* and *Escherichia coli*. Identification of a new class of transport protein. *J. Biol. Chem.* (268): 14071-14080.
- Spinelli, S. V., Pontel, L. B., Garcia Vescovi, E. and Soncini, F. C. (2008). Regulation of magnesium homeostasis in *Salmonella*: Mg(2+) targets the *mgtA* transcript for degradation by RNase E. *FEMS Microbiol. Lett.* (280): 226-234.
- Stancik, L. M., Stancik, D. M., Schmidt, B., Barnhart, D. M., Yoncheva, Y. N. and Slonczewski, J. L. (2002). pH-dependent expression of periplasmic proteins and amino acid catabolism in *Escherichia coli*. *J. Bacteriol.* (184): 4246-4258.
- Steffen, P., Voss, B., Rehmsmeier, M., Reeder, J. and Giegerich, R. (2006). RNASHAPES: an integrated RNA analysis package based on abstract shapes. *Bioinformatics*. (22): 500-503.
- Stein, A. and Crothers, D. M. (1976b). Conformational changes of transfer RNA. The role of magnesium(II). *Biochemistry*. (15): 160-168.
- Stein, A. and Crothers, D. M. (1976a). Equilibrium binding of magnesium(II) by *Escherichia coli* tRNA^{fMet}. *Biochemistry*. (15): 157-160.
- Steinert, H. et al. (2017). Pausing guides RNA folding to populate transiently stable RNA structures for riboswitch-based transcription regulation. *Elife*. (6):
- Sudarsan, N., Hammond, M. C., Block, K. F., Welz, R., Barrick, J. E., Roth, A. and Breaker, R. R. (2006). Tandem riboswitch architectures exhibit complex gene control functions. *Science*. (314): 300-304.
- Sudarsan, N., Lee, E. R., Weinberg, Z., Moy, R. H., Kim, J. N., Link, K. H. and Breaker, R. R. (2008). Riboswitches in eubacteria sense the second messenger cyclic di-GMP. *Science*. (321): 411-413.
- Sudarsan, N., Wickiser, J. K., Nakamura, S., Ebert, M. S. and Breaker, R. R. (2003). An

mRNA structure in bacteria that controls gene expression by binding lysine. *Genes Dev.* (17): 2688-2697.

Tanaka, Y., Yamagata, S., Kitago, Y., Yamada, Y., Chimnaronk, S., Yao, M. and Tanaka, I. (2009). Deduced RNA binding mechanism of ThiI based on structural and binding analyses of a minimal RNA ligand. *RNA*. (15): 1498-1506.

Tao, X., Schiering, N., Zeng, H. Y., Ringe, D. and Murphy, J. R. (1994). Iron, DtxR, and the regulation of diphtheria toxin expression. *Mol. Microbiol.* (14): 191-197.

Tian, S., Yesselman, J. D., Cordero, P. and Das, R. (2015). Primerize: automated primer assembly for transcribing non-coding RNA domains. *Nucleic Acids Res.* (43): W522-526.

Troxell, B. and Yang, X. F. (2013a). Metal-dependent gene regulation in the causative agent of Lyme disease. *Front Cell Infect Microbiol.* (3): 79.

Troxell, B., Ye, M., Yang, Y., Carrasco, S. E., Lou, Y. and Yang, X. F. (2013b). Manganese and zinc regulate virulence determinants in *Borrelia burgdorferi*. *Infect. Immun.* (81): 2743-2752.

Tseng, H. H., Weinberg, Z., Gore, J., Breaker, R. R. and Ruzzo, W. L. (2009). Finding non-coding RNAs through genome-scale clustering. *J Bioinform Comput Biol.* (7): 373-388.

Turner, A. G., Ong, C. L., Gillen, C. M., Davies, M. R., West, N. P., McEwan, A. G. and Walker, M. J. (2015). Manganese homeostasis in group A *Streptococcus* is critical for resistance to oxidative stress and virulence. *MBio.* (6):

Uhm, H., Kang, W., Ha, K. S., Kang, C. and Hohng, S. (2018). Single-molecule FRET studies on the cotranscriptional folding of a thiamine pyrophosphate riboswitch. *Proc. Natl. Acad. Sci. U.S.A.* (115): 331-336.

Vecerek, B., Moll, I. and Blasi, U. (2007). Control of Fur synthesis by the non-coding RNA RyhB and iron-responsive decoding. *EMBO J.* (26): 965-975.

Velasco, E., Wang, S., Sanet, M., Fernandez-Vazquez, J., Jove, D., Gllaria, E., Valledor, A. F., O'Halloran, T. V. and Balsalobre, C. (2018). A new role for Zinc limitation in bacterial pathogenicity: modulation of α -hemolysin from uropathogenic *Escherichia coli*. *Sci Rep.* (8): 6535.

Vernon, W. B. (1988). The role of magnesium in nucleic-acid and protein metabolism. *Magnesium.* (7): 234-248.

Veyrier, F. J., Boneca, I. G., Cellier, M. F. and Taha, M. K. (2011). A novel metal transporter mediating manganese export (MntX) regulates the Mn to Fe intracellular ratio and *Neisseria meningitidis* virulence. *PLoS Pathog.* (7): e1002261.

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

Waldron, K. J., Rutherford, J. C., Ford, D. and Robinson, N. J. (2009). Metalloproteins and metal sensing. *Nature.* (460): 823-830.

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

Washietl, S., Hofacker, I. L. and Stadler, P. F. (2005). Fast and reliable prediction of noncoding RNAs. *Proc. Natl. Acad. Sci. U.S.A.* (102): 2454-2459.

Waters, L. S., Sandoval, M. and Storz, G. (2011). The *Escherichia coli* MntR miniregulon includes genes encoding a small protein and an efflux pump required for manganese homeostasis. *J. Bacteriol.* (193): 5887-5897.

Waters, L. S. and Storz, G. (2009). Regulatory RNAs in bacteria. *Cell.* (136): 615-628.

Weinberg, Z. et al. (2007). Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. *Nucleic Acids Res.* (35): 4809-4819.

Weinberg, Z., Lunse, C. E., Corbino, K. A., Ames, T. D., Nelson, J. W., Roth, A., Perkins, K. R., Sherlock, M. E. and Breaker, R. R. (2017). Detection of 224 candidate structured RNAs by comparative analysis of specific subsets of intergenic regions. *Nucleic Acids Res.* (45): 10811-10823.

Weinberg, Z., Wang, J. X., Bogue, J., Yang, J., Corbino, K., Moy, R. H. and Breaker, R. R. (2010). Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol.* (11): R31.

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

Wickiser, J. K., Cheah, M. T., Breaker, R. R. and Crothers, D. M. (2005). The kinetics of ligand binding by an adenine-sensing riboswitch. *Biochemistry*. (44): 13404-13414.

Will, S., Reiche, K., Hofacker, I. L., Stadler, P. F. and Backofen, R. (2007). Inferring noncoding RNA families and classes by means of genome-scale structure-based clustering. *PLoS Comput. Biol.* (3): e65.

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

Winkler, W. C., Cohen-Chalamish, S. and Breaker, R. R. (2002b). An mRNA structure that controls gene expression by binding FMN. *Proc. Natl. Acad. Sci. U.S.A.* (99): 15908-15913.

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

Winkler, W. C., Nahvi, A., Sudarsan, N., Barrick, J. E. and Breaker, R. R. (2003). An mRNA structure that controls gene expression by binding S-adenosylmethionine. *Nat. Struct. Biol.* (10): 701-707.

Winsor, G. L., Griffiths, E. J., Lo, R., Dhillon, B. K., Shay, J. A. and Brinkman, F. S. (2016). Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database. *Nucleic Acids Res.* (44): D646-653.

Winsor, G. L., Khaira, B., Van Rossum, T., Lo, R., Whiteside, M. D. and Brinkman, F. S. (2008). The Burkholderia Genome Database: facilitating flexible queries and comparative analyses. *Bioinformatics*. (24): 2803-2804.

Wurtzel, O., Yoder-Himes, D. R., Han, K., Dandekar, A. A., Edelheit, S., Greenberg, E. P., Sorek, R. and Lory, S. (2012). The single-nucleotide resolution transcriptome of *Pseudomonas aeruginosa* grown in body temperature. *PLoS Pathog.* (8): e1002945.

Yamauchi, T., Miyoshi, D., Kubodera, T., Nishimura, A., Nakai, S. and Sugimoto, N. (2005). Roles of Mg²⁺ in TPP-dependent riboswitch. *FEBS Lett.* (579): 2583-2588.

Yao, Z., Barrick, J., Weinberg, Z., Neph, S., Breaker, R., Tompa, M. and Ruzzo, W. L. (2007). A computational pipeline for high-throughput discovery of cis-regulatory noncoding RNA in prokaryotes. *PLoS Comput. Biol.* (3): e126.

Yao, Z., Weinberg, Z. and Ruzzo, W. L. (2006). CMfinder--a covariance model based RNA motif finding algorithm. *Bioinformatics*. (22): 445-452.

Zhao, B., Guffy, S. L., Williams, B. and Zhang, Q. (2017). An excited state underlies gene regulation of a transcriptional riboswitch. *Nat. Chem. Biol.* (13): 968-974.

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

Zuker, M. and Stiegler, P. (1981). Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res.* (9): 133-148.

Zwiefka, A., Kohn, H. and Widger, W. R. (1993). Transcription termination factor rho: the

site of bicyclomycin inhibition in *Escherichia coli*. *Biochemistry*. (32): 3564-3570.

10. Annexe I (Des séquences utilisées pour différentes constructions.)

Tableaux 1. Annexe I. Des séquences utilisées pour construire différents essais *in vitro* (*Desulfovibrio* sp.) et *in vivo* (*Pseudomonas aeruginosa*).

Seq ID	Sequences
Desulfovibrio magneticus RS1 : NC_012796:5019180-5019383	ggCAAGCCCTTGACGCCCCCGGCTCCCGTGACTATGGTGGCCTTCC CAAAGGGGAGTAGCTGTTTCGGGCAGGGTCAACATCACTGCGGCACG TCAGCCGCCTGGCCCTGTCGTCGGTCCCTCGGAACCGATCGGCGAGA C
RN_0001_Desmag_F1	TTCTAATACGACTCACTATAGGCAAGCCCTTGACGCCCCCGGCTC
RN_0002_Desmag_R2	CCCTTTGGGAAGGCCACCATAGTCACGGGAGCCGGGGGCGT
RN_0003_Desmag_F3	TGGCCTTCCCAAAGGGGAGTAGCTGTTTCGGGCAGGGTCAACATCAC TGGGCACGTCAG
RN_0004_Desmag_R4	GTCTCGCCGATCGGTTCCGAGGGACCGACGACAGGGCCAGGCGGCTG ACGTGCCGAGT
Desulfovibrio_F1.1	ggCCCCGGCTCCCGTGACTATGGTGGCCTTCCCAAAGGGGAGTAGCT GTTTCGGGCAGGGTCAACATCACTGCGGCACGTCAGCCGCCTGGCCC TGTCGTCGGTCCCTCGGAACCGATCGGCGAGAC
Desulfovibrio_F1.2	ggACTATGGTGGCCTTCCCAAAGGGGAGTAGCTGTTTCGGGCAGGGT CAACATCACTGCGGCACGTCAGCCGCCTGGCCCTGTCGTCGGTCCCT CGGAACCGATCGGCGAGAC
RN_0005_Desmag_F1.1	TTCTAATACGACTCACTATAGGCCCGGCTCCCGTGACTATGGTGG
RN_0006_Desmag_F1.2	TTCTAATACGACTCACTATAGGCGTGACTATGGTGGCCTTCCCAAAG G
Desulfovibrio_R5	ggCAAGCCCTTGACGCCCCCGGCTCCCGTGACTATGGTGGCCTTCC CAAAGGGGAGTAGCTGTTTCGGGCAGGGTCAACATCACTGCGGCACG TCAGCCGCCTGGCCCTGTCGTCGGTCCCTCGGAACCGATCGGCGAGA CCTTTGG
RN_0068R5.1 (La construction avec F1.2)	GGTCTCGCCGATCGGTTCCGAGGGACCGACGACAGGGCCAGGCGGCT GACGTGCCGAGT

RN_0069_R5.1 (La construction avec F1.2)	AGGTCTCGCCGATCGGTTCCGAGGGACCGACGACAGGGCCAGGCGGCTGACGTGCCGCAGT
RN_0070_R5.1 (La construction avec F1.2)	AAGGTCTCGCCGATCGGTTCCGAGGGACCGACGACAGGGCCAGGCGGCTGACGTGCCGCAGT
RN_0071_R5.1 (La construction avec F1.2)	AAAGGTCTCGCCGATCGGTTCCGAGGGACCGACGACAGGGCCAGGCGCTGACGTGCCGCAGT
RN_0072_R5.1 (La construction avec F1.2)	CAAAGGTCTCGCCGATCGGTTCCGAGGGACCGACGACAGGGCCAGGCGGCTGACGTGCCGCAGT
RN_0073_R5.1 (La construction avec F1.2)	CCAAAGGTCTCGCCGATCGGTTCCGAGGGACCGACGACAGGGCCAGGCGGCTGACGTGCCGCAGT
RN_0060_Des_F3_mut_G2C	TGGCCTCCCAAAGGGCACTCGCTGTTTCGGGCAGGGTCAACATCAC TCGGGCACGTACG
RN_0061_Ds_F3_mut_ccgt	TGGCCTCCCAAAGGGGAGTAGCTGTTTCCCGTGGGTCAACATCACT GCGGCACGTACG
RN_0062_Ds_F3_mut_ca2gg	TGGCCTCCCAAAGGGGAGTAGCTGTTTCGGGCAGGGTCAACATGGC TCGGGCACGTACG
<i>Desulfovibrio desulfuricans</i> subsp. <i>desulfuricans</i> str., NC_011883:720681-720371	ggGUUGACAGCGUUGCUGAGUGGCCUAUAAGCUGGCUGCUACAAGGGG AGUAACUGGGUUCUUGAAAACCCGGGCGGUAUCAACAGCAUGGCUUUC GGGCCUGGUGCCGCCGUCGGUAAGCCGAUUGAGACCUUGGCGUC A
RN_0064_D.des_F1	TTCTAATACGACTCACTATAGGGTTGACAGCGTTGCTGAGTGGCTAT AAGCTGGCTGC
RN_0065_D.des_R2	ACCGCCCGGGTTTCAAGAACCAGTTACTCCCCTTGTAGCAGCCAGC TTATAGCCAC
RN_0066_D.des_F3	AACCGGGCGGTATCAACAGCATGGCTTTCGGGCCTGGTGC
RN_0067_D.des_R4	TGACGCCAAGGTCTACTAATCGGCTTACCGACGGCGGCACCAGGCC CGAA
<i>Pseudomonas aeruginosa</i> UCBPP_PA14_NC_008463.1_5468813_5469164	TTCACCTGCCGCCAGCGATTCCCTCGCGGATCGCTGCGGATGCCGG CACGTTTCTTGCATCAGGCTCCCTCGCCGCCCTCCCGACGAAAG TTTCCGTCTGGGCTTGCCCAAGCCGCCGGCTTGGGCAGAATGCGC CCCGTCTAGGGGAGTAGTCTCCCGCAGCGCCGACCGCCTCGCC CGGCACGCGTCAACACACTTGGTCCGAGACCATGGCGCGCGGACC CAAGCTGCCAGACAGGCCAAGGGTTTGACAAGACCTATGACACGC ACGATCTGACCCGGGGCGGCAGATGGCGCGTGTATGGTATTGGT CGACCCGCCCTTAGGAAGCCTG
RN_0079_PA14_rs_fw	CCATGGCCTGCAGGAGTACTCGCCCCGTCTAGGGGA
RN_0079.1_PA14_cds_rev	CTGGCCGTTAATAATGAATGAAATTTTTTAGTttccgcgaggggcaa cg

Des séquences de *Desulfovibrio.sp* et *Pseudomonas aeruginosa* avec différentes constructions pour le *riboswitch yybP-ykoY* ressemblant. Les deux gg (en rouge) ont été rajoutés aux séquences pour faciliter la transcription par l'enzyme de transcription de l'ARN T7. Les séquences de *Desulfovibrio magneticus* RS1 : NC_012796:5019180-5019383 et *Desulfovibrio desulfuricans* subsp. *desulfuricans* str., NC_011883:720681-720371 montrent la région intergénique avec le motif *yybP-ykoY* ressemblant. Des oligonucléotides de RN_0001 à RN_0004 et de RN_0064 à RN_0067 ont été utilisés pour la construction de *Desulfovibrio magneticus* RS1 et *Desulfovibrio desulfuricans* subsp. *desulfuricans* str. respectivement. Des constructions F1.1 et F1.2 sont 15 et 29 nucléotides plus courtes respectivement de la séquence initiale de *Desulfovibrio magneticus* RS1. La construction de *Desulfovibrio_R5* a 6 nucléotides de plus, (la boîte soulignée) dans sa 3'-UTR. Des séquences suivantes montrent des constructions itératives en rajoutant d'un nucléotide de plus de cette boîte soulignée. Des constructions de RN_0060 à RN_0062 introduisent des mutations dans le moi *yybP-ykoY* ressemblant.


```
#
# Post-search info for CM 1 : yybP-ykoY
#
# number of hits surv fraction
# -----
# rnd mod alg cfg beta expected actual expected actual
# --- --- --- --- ---
1 cm ins loc 1e-15 1.000 4 0.4417 0.4183
#
# expected time actual time
# -----
00:00:01.70 00:00:00.00
//
#
# CPU time : 1.42u 0.02 s 00:00:01.44 Elapsed : 00:00:01
```


B)

Alignement des séquences de Desulfovibrop.spp avec le modèle prédit par Cmfinder dans la cluster 1.1

```
cmsearch :: search a sequence database with an RNA CM
# INFERNAL 1.0.2 (October 2009)
# Copyright (C) 2009 HHMI Janelia Farm Research Campus
# Freely distributed under the GNU General Public License (GPLv3)
# -----
# command : cmsearch model_calcium.cm desulfo_sequences_from_cluster_list.fa
# date : Wed May 30 15:43:03 2018
# num seqs : 7
# dbsize(Mb) : 0.004600
#
# Pre-search info for CM 1 : seq_Ca_cm_finder-1
#
# cutoffs predictions
# -----
# rnd mod alg cfg beta E value bit sc surv run time
# -----
1 cm ins loc 1e-15 1.000 4.54 0.0228 00:00:04.34
B.1
```

```
CM : seq_Ca_cm_finder-1
>NC_012796:5019180-5019383 : Desulfovibrio magneticus RS-1, (NC_012796).
```

Plus strand results :

```
Query = 1 - 112, Target = 43 - 140
Score = 45.96, E = 9.74e-11, P = 1.531e-13, GC = 65
```

```
::::::::::::::(((,<<<<<<<_____>>>>>><<<<<<<<-----<<<
1 aGUCAUUGGGGAGUAGCcCaucCucggggccacacccgagGaGcccgCGUCAACAUACUUGg 60
+ CA GGGGAGUAGC +++ G C : G : GUCAACAU G :
43 UCCCAAAGGGGAGUAGCUG-----UUUCG-----GGCAGGGUCAACAUACUCG 86
```

```
<<<____.>>>>>>>>>>>>,.....)))) ::::
61 Ggcuaa.augcCcCaUGGCGcgggCG... Gccuuug.caaCgGcaaaauugGCgAGAC 112
GGC+ +GCC : C UGGC : C : G CG G+C U AAC G AU+ GCGAGAC
87 GGCACGuCAGCCGCCUGGCCCUGUCGucgGUCCUCgGAACCG---AUCGGCGAGAC 140
```

B.2

```
>NC_011883:720681-720371 : (Desulfovibrio desulfuricans subsp. desulfuricans str. ATCC 27774)
```

Plus strand results :

```
Query = 1 - 112, Target = 175 - 264
Score = 46.30, E = 8.066e-11, P = 1.268e-13, GC = 59
```

```
::::::::::::::(((,<<<<<<<_____>>>>>><<<<<<<<-----<<<
1 aGUCAUUGGGGAGUAGCcCaucCucggggccacacccgagGaGcccgCGUCAACAUACUUGg 60
A GGGGAGUA : C : : G:: C++ A:: C : G G : C : G : UCAACA U
175 A-----GGGGAGUAACUG----GCUUCUUGAAACCCG--GGCGGUAUCAACAGCAU--- 219
```

```
<<<____.>>>>>>>>>>>>,.....)))) ::::
61 Ggcuaa.augcCcCaUGGCGcgggCGGccuuugcaaCgGcaaaauugGCgAGAC 112
GGCU++ +GCC UGG : C : G : CG C + +GC AUU : G : GAGAC
220 GGCUUUcGGGCC---UGGUGCCGCCGUC----GGUAAGCCGAUUGUGAGAC 264
```

12. Annexe III (matériels supplémentaires pour l'article :

Single mutation in hammerhead ribozyme favors cleavage

activity with manganese over magnesium.

Les données se trouvent liées au document dans le dépôt institutionnelle.

**13. Annexe IV (matériels supplémentaires pour l'article :
Search for 5'-leader regulatory RNA structures based on
gene annotation aided by the RiboGap database).**

Les données se trouvent liées au document dans le dépôt institutionnelle.

**14. Annexe V (matériels supplémentaires pour l'article :
Survey of non-coding RNAs involved in bacterial virulence).**

Les données se trouvent liées au document dans le dépôt institutionnelle.

15. Annexe VI (Article Methods in Molecular Biology).

Les données se trouvent liées au document dans le dépôt institutionnelle.

16. Annexe VII (Article Nature Communications).

Les données se trouvent liées au document dans le dépôt institutionnelle.

17. Annexe VIII (Des résultats de ONE way ANOVA).

ANOVA - Single Factor

Alpha 0.05

Groups	Count	Sum	Mean	Variance
SAM	30	938229.4	31274.31333	294238372
EDTA	30	58287.37	1942.912333	390351.1542
EDTA+1mM Ca²⁺	30	578014.54	19267.15133	159678243
EDTA+300µM Mn²⁺	30	71532.04	2384.401333	226291.8794

Source of Variation	SS	df	MS	F
Between Groups	18183610922.352	3	6061203641	53.34002331
Within Groups	13181464475.276	116	113633314.4	
Total	31365075397.628	119		

P-value 9.6655586E-22 F critical 2.682809407

18. Annexe IX (droit d'auteurs de journaux).

1 Kim, J. N. and Breaker, R. R. (2008). Purine sensing by riboswitches. *Biol. Cell.* (100): 1-11.

License Number :4765570064544

2 Serganov, A. and Nudler, E. (2013). A decade of riboswitches. *Cell.* (152): 17-24.

License Number : 4765570722952

3 Breaker, R. R. (2011). Prospects for riboswitch discovery and analysis. *Mol. Cell.* (43): 867-879.

License Number : 4765580743577

4 Breaker, R. R. (2012b). New insight on the response of bacteria to fluoride. *Caries Res.* (46): 78-81.

Open access

5 Hammann, C., Luptak, A., Perreault, J. and de la Pena, M. (2012). The ubiquitous hammerhead ribozyme. *RNA.* (18): 871-885.

Waiting

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

License Number :4765590645088.

7 Kolev, S. K., Petkov, P. S., Rangelov, M. A., Trifonov, D. V., Milenov, T. I. and Vayssilov, G. N. (2018). Interaction of Na⁺, K⁺, Mg²⁺ and Ca²⁺ counter cations with RNA. *Metallomics.* (10): 659-678.

ISSN :1756-591X

8 Feig, A. and Uhlenbeck, O. (1999). 12 The Role of Metal Ions in RNA Biochemistry. *Cold Spring Harbor Monograph Archive.* (37):

Open access

9 Price, I. R., Gaballa, A., Ding, F., Helmann, J. D. and Ke, A. (2015). Mn(2⁺)-sensing mechanisms of yybP-ykoY orphan riboswitches. *Mol. Cell.* (57): 1110-1123.

License Number :4765611169126.

10 Dambach, M., Sandoval, M., Updegrove, T. B., Anantharaman, V., Aravind, L., Waters, L. S. and Storz, G. (2015). The ubiquitous yybP-ykoY riboswitch is a manganese-responsive regulatory element. *Mol. Cell.* (57): 1099-1109.

License Number :4765611415590.

11 Search for 5'-leader regulatory RNA structures based on gene annotation aided by the RiboGap database.

Non permission required for the authors.