



# Etude fonctionnelle de nouveaux ARN régulateurs exprimés chez *Enterococcus faecium*

Killian Le Neindre

## ► To cite this version:

Killian Le Neindre. Etude fonctionnelle de nouveaux ARN régulateurs exprimés chez *Enterococcus faecium*. Microbiologie et Parasitologie. Université de Rennes, 2021. Français. NNT : 2021REN1B036 . tel-03890076

**HAL Id: tel-03890076**

**<https://theses.hal.science/tel-03890076>**

Submitted on 8 Dec 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THESE DE DOCTORAT DE L'UNIVERSITE DE RENNES 1

ECOLE DOCTORALE N° 605

*Biologie Santé*

Spécialité : *Microbiologie, virologie, parasitologie*

Par

**Killian LE NEINDRE**

## **Etude fonctionnelle de nouveaux ARN régulateurs exprimés chez *Enterococcus faecium***

Thèse présentée et soutenue à l'université de Rennes 1, le 16 Novembre 2021

Unité de recherche : INSERM U1230 ARN régulateurs bactériens et médecine

### **Rapporteurs avant soutenance :**

Martine PESTEL-CARON      PU-PH, IRIB GRAM EA2656, Université de Rouen  
Marie-Frédérique LARTIGUE      PU-PH, ISP-BRMF INRA UMR1282, Université de Tours

### **Composition du Jury :**

Président :	Reynald GILLET	PU, IGDR CNRS UMR6290, Université Rennes 1
Examineurs :	Reynald GILLET	PU, IGDR CNRS UMR6290, Université Rennes 1
	Olga SOUTOURINA	PU, I2BC UMR9198, Université Paris-Sud
	Jean-Christophe GIARD	PU, U2RM EA4655, UniCaen
	Martine PESTEL-CARON	PU-PH, IRIB GRAM EA2656, Université de Rouen
	Marie-Frédérique LARTIGUE	PU-PH, ISP-BRMF INRA UMR1282, Université de Tours

Dir. de thèse :	Vincent CATTOIR	PU-PH, BRM INSERM UMR1230, Université Rennes 1
Co-dir. de thèse :	Brice FELDEN †	PU, BRM INSERM UMR1230, Université Rennes 1



# Table des matières

---

LISTE DES FIGURES .....	4
REMERCIEMENTS .....	8
LISTE DES ABREVIATIONS .....	11
INTRODUCTION .....	15
ANALYSE BIBLIOGRAPHIQUE .....	17
I. <i>ENTEROCOCCUS FAECIUM</i> , PATHOGENE HOSPITALIER MAJEUR .....	18
A. CARACTERISTIQUES MICROBIOLOGIQUES .....	18
B. EPIDEMIOLOGIE MOLECULAIRE : COMMENSAL OU PATHOGENE .....	19
C. PLACE EN INFECTIOLOGIE .....	20
1. <i>Epidémiologie des infections et résistances</i> .....	20
a. Infections et hôpital .....	20
b. Acquisition de résistance .....	22
2. <i>Prise en charge</i> .....	23
a. Antibiothérapie .....	23
b. Gestion épidémique .....	24
D. RESISTANCES AUX ANTIBIOTIQUES .....	25
1. <i><math>\beta</math>-lactamines</i> .....	25
2. <i>Glycopeptides</i> .....	26
3. <i>Quinolones</i> .....	28
4. <i>Antibiotiques inhibiteurs de la synthèse protéique</i> .....	28
a. Aminosides .....	35
b. Macrolides et apparentés .....	36
c. Tétracyclines .....	38
d. Linézolide .....	42
5. <i>Daptomycine</i> .....	43
E. PLASTICITE GENOMIQUE ET SUPPORTS GENETIQUES DE L'ANTIBIORESISTANCE CHEZ <i>E. FAECIUM</i> .....	44
1. <i>Eléments génétiques mobiles</i> .....	44

2. Implications dans la résistance et les épidémies .....	47
<b>II. SYSTEMES DE REGULATION BACTERIENS DEPENDANT DE LA CONFORMATION DE L'ARN .....</b>	<b>52</b>
A. ARN REGULATEUR « CONVENTIONNEL » .....	53
B. REGULATION EN 5'UTR .....	59
1. Atténuation transcriptionnelle et traductionnelle .....	59
a. Régulation de <i>mef</i> (E).....	65
b. Régulation de <i>tet</i> (M) .....	67
2. Riboswitch.....	71
a. Régulation de <i>glmS</i> chez les <i>Firmicutes</i> : ribozyme .....	73
b. Régulation de la biosynthèse des pyrimidines .....	74
c. T-box riboswitch .....	78
<b>CONTRIBUTION SCIENTIFIQUE.....</b>	<b>85</b>
<b>PROJET DE RECHERCHE .....</b>	<b>86</b>
<b>RESULTATS ET DISCUSSION .....</b>	<b>89</b>
<b>I. CARACTERISATION ET ANALYSE INITIALE DES ARN CANDIDATS .....</b>	<b>89</b>
A. MISE EN EVIDENCE DES ARN CANDIDATS D'INTERET .....	89
B. REGULATEURS EN 5'UTR.....	100
1. Antibiotique-dépendant.....	100
a. Tétracyclines et <i>ern0030</i> .....	100
b. Macrolides et <i>ern1300</i> .....	104
2. T-box riboswitch .....	106
a. L'opéron <i>hisS-aspS</i> et <i>ern2050</i> .....	106
b. Opéron de gènes codant des protéines putatives d'import de l'histidine et <i>ern1900</i> .....	107
3. Riboswitchs putatifs avec effecteur inconnu .....	109
a. Régulation du gène <i>glmU</i> par <i>ern2210</i> ? .....	109
b. Régulation d'un opéron de transporteur ABC putatif par <i>ern2410</i> ?.....	112
C. ARN REGULATEURS .....	113
1. Action en cis .....	113
a. Le couple <i>ern0280-ern0290</i> comme toxine-antitoxine de type 1 .....	113

b. Régulation d'une aldolase ( <i>fsaA</i> -like) par <i>ern1260</i> ? .....	115
2. Action en trans d' <i>ern0120</i> ? .....	117
3. Action mixte en cis et trans d' <i>ern0160</i> ? .....	119
<b>II. IMPLICATION DE L'ARN <i>ERN0030</i> DANS LA REGULATION DE LA RESISTANCE AUX</b>	
<b>TETRACYCLINES INDUIT PAR LE GENE <i>TET(M)</i> .....</b>	<b>121</b>
A. RESUME DE L'ARTICLE 1 .....	121
1. Présentation .....	121
2. Principaux résultats .....	122
B. ARTICLE 1 .....	123
<b>III. REGULATION DE DEUX AARS PAR LE T-BOX RIBOSWITCH <i>ERN2050</i> .....</b>	<b>161</b>
A. RESUME DE L'ARTICLE 2 .....	161
1. Présentation .....	161
2. Principaux résultats .....	162
B. ARTICLE 2 .....	164
<b>IV. DISCUSSION .....</b>	<b>202</b>
<b>CONCLUSION .....</b>	<b>220</b>
<b>BIBLIOGRAPHIE.....</b>	<b>222</b>
<b>ANNEXES .....</b>	<b>241</b>

## Liste des Figures

Figure 1. Pourcentage d'ERV parmi les souches invasives d' <i>E. faecium</i> dans les pays d'Europe en 2006 (A), 2013 (B) et 2019 (C).....	23
Figure 2. Description de la résistance à la vancomycine par <i>vanA</i> (Cattoir et Leclercq, 2010; Le Neindre, 2018) .....	27
Figure 3. Description des acteurs impliqués dans la synthèse peptidique.....	29
Figure 4. Synthèse peptidique : Initiation de la traduction. ....	30
Figure 5. Synthèse peptidique : Elongation.....	32
Figure 6. Synthèse peptidique : Terminaison.....	33
Figure 7. Représentation des étapes inhibées par les antibiotiques inhibiteurs de la synthèse peptidique. ....	34
Figure 8. Description du mécanisme de résistance aux tétracyclines par la protéine Tet(M) (Dönhöfer <i>et al.</i> , 2012). ....	41
Figure 9. Présentation des 3 types de transposons (García-Solache et Rice, 2019).....	45
Figure 10. Disposition schématique des différents modules composant les transposons Tn916-like (Roberts et Mullany, 2009).....	49
Figure 11. Représentation de la mobilisation des transposons de type Tn916 (Salysers <i>et al.</i> , 1995).....	50
Figure 12. Représentation des mécanismes de régulation induits par l'interaction entre un ARNreg et un ARNm. ....	55
Figure 13. Northern blot validant 10 ARN candidats exprimés par <i>E. faecium</i> (Sinél <i>et al.</i> 2017).....	57
Figure 14. Représentation du principe d'atténuation transcriptionnelle et traductionnelle. ....	60
Figure 15. Représentation de la régulation par atténuation transcriptionnelle de l'opéron <i>trp</i> dépendant de la restriction ou non de différents AA (Yanofsky, 1981). ....	62

Figure 16. Représentation d'une atténuation traductionnelle impliquant des gènes de résistance aux antibiotiques (phénicolés et macrolides) (Lovett, 1990; Ramu <i>et al.</i> , 2009). .....	64
Figure 17. Représentation des deux conformations de la région 5'UTR du gène <i>mef(E)</i> (Chancey <i>et al.</i> , 2015). ....	66
Figure 18. Structures secondaires prédites selon la stabilité des tige-boucles (Su <i>et al.</i> , 1992).....	68
Figure 19. Position des délétions retrouvées dans les souches BTR (Beabout <i>et al.</i> , 2015b). .....	70
Figure 20. Représentation du principe de riboswitch. ....	72
Figure 21. Représentation de la régulation de <i>glmS</i> par son ribozyme (Kawada-Matsuo <i>et al.</i> , 2017). ....	73
Figure 22. Représentation des gènes impliqués dans la biosynthèse des pyrimidines. ....	75
Figure 23. Mécanisme de régulation de l'opéron <i>pyr</i> médié par un riboswitch PyrR-dépendant (Zhang et Switzer, 2003).....	76
Figure 24. Représentation d'un T-box riboswitch (Kreuzer et Henkin, 2018).....	79
Figure 25. Structure secondaire du T-box riboswitch régulant <i>tyrS</i> et de l'ARN <sup>Tyr</sup> chez <i>B. subtilis</i> (Kreuzer et Henkin, 2018). ....	81
Figure 26. Arbre phylogénétique montrant l'évolution des T-box riboswitchs ayant un codon « specifier » Asp ou Asn (Vitreschak <i>et al.</i> , 2008). ....	84
Figure 27. Validation des souches d'étude d' <i>ern0160</i> . ....	91
Figure 28. Expression d'ARN altérée par la présence de ciprofloxacine.....	93
Figure 29. Représentation de l'environnement de chacun des ARN candidats. ....	96
Figure 30. Validation de l'ARN <i>ern0030</i> dans les souches E1162 et Aus0004.....	101
Figure 31. Essai de détermination de demi-vie dans la souche E1162 d' <i>E. faecium</i> . ....	102
Figure 32. Contrôle des transcrits obtenus par transcription <i>in vitro</i> après élution. ....	103



Figure 33. Alignement des séquences nucléotidiques des séquences 5'UTR du gène <i>mef(A)</i> pour la souche Aus0004 et <i>mef(E)</i> pour la souche GA17457. ....	105
Figure 34. Région non traduite en 5' (5'UTR) du gène codant la protéine hypothétique HisM. ....	107
Figure 35. Représentation de l'opéron <i>hisXYZ</i> et des gènes adjacents. ....	108
Figure 36. Séquence nucléotidique de la potentielle région non traduite en 5' (5'UTR) du gène <i>glmU</i> . ....	109
Figure 37. Représentation de la voie de biosynthèse de l'UDP-GlcNAc et des mécanismes de régulation associés chez <i>Lactobacillus casei</i> (Rodríguez-Díaz <i>et al.</i> , 2012b). ....	110
Figure 38. Représentation de l'opéron potentiellement régulé par <i>ern2410</i> et des gènes adjacents. ....	112
Figure 39. Séquence nucléotidique et peptidique contenant les 2 ARN <i>ern0280-ern0290</i> . ....	114
Figure 40. Représentation du mécanisme de régulation hypothétique d' <i>ern1260</i> . ....	115
Figure 41. Représentation de l'élongation d' <i>ern1260</i> induite par la daptomycine vis-à-vis de la position du gène <i>fsaA</i> . ....	116
Figure 42. Représentation de l'autorégulation hypothétique entre <i>ern0120</i> et <i>rnhC</i> . ....	117
Figure 43. Représentation de l'approche MAPS permettant l'identification des cibles d'un ARNreg. ....	204
Figure 44. Représentation des méthodes de validation des cibles d'ARNreg. ....	206
Figure 45. Voie de biosynthèse de la L-Lysine chez <i>E. faecium</i> Aus0004. ....	213
Figure 46. Hypothèse du mécanisme de régulation de <i>tet(M)</i> Tn916-dépendant : médiation par une protéine régulatrice. ....	215
Figure 47. Hypothèse du mécanisme de régulation de <i>tet(M)</i> Tn916-indépendant : médiation par protection ribosomale. ....	218

## Liste des Tableaux

Tableau 1. Place des deux principales espèces d'entérocoques dans les micro-organismes responsables d'infections nosocomiales. ....	21
Tableau 2. Récapitulatif des coordonnées déterminées par 5'3'RACE des 9 ARN candidats. ....	89
Tableau 3. Liste des 16 ARN sélectionnés. ....	92
Tableau 4. Ajustement des coordonnées 5' et 3' des ARN mappés par 5'3'RACE. ....	97
Tableau 5. Prédiction des coordonnées 5' et 3' des ARN non mappés par 5'3'RACE. ....	98
Tableau 6. Bilan des cibles prédites par TargetRNA2 (p-value <0,05) selon la catégorie (COG) de la cible et les différents ARN candidats. ....	99
Tableau 7. Description des résultats obtenus par l'alignement de la région 5'UTR de <i>glmU</i> . ....	111
Tableau 8. Sélection de cibles prédites par Target RNA2 et copraRNA d' <i>ern0120</i> . ....	118
Tableau 9. Sélection de cibles prédites par Target RNA2 et copraRNA d' <i>ern0160</i> . ....	120
Tableau 10. Liste récapitulative des fonctions des 16 ARN sélectionnés initialement. ....	219

## Remerciements

Après plus de dix longues années d'étude, c'est avec ce travail de thèse que se conclut officiellement ma période « étudiant ».

Je souhaite tout d'abord remercier l'ensemble des membres de mon jury de thèse, les Professeurs Martine Pestel-Caron, Marie-Frédérique Lartigue, Reynald Gillet, Olga Soutourina et Jean-Christophe Giard d'avoir accepté de juger ce travail de thèse.

Je remercie également le Pr Vincent Cattoir, mon directeur de thèse, pour sa disponibilité, son expertise et nos discussions scientifiques enrichissantes. Merci de m'avoir inculqué ta rigueur scientifique.

J'ai une pensée pour le Pr Brice Felden, mon co-directeur de thèse, chercheur passionné par le monde des ARN, m'ayant accepté au sein de son équipe de recherche et permis de découvrir le monde de la recherche.

Je remercie les différents membres (titulaires U et H-U, doctorants, M2...) de l'équipe INSERM U1230 BRM à Rennes, notamment Sveta et Valérie pour leur collaboration sur l'article *ern2050* et la reprise de mes manip depuis mon départ, Yoann pour sa relecture critique de l'article *ern0030*, Marie-Laure et Gaetan pour leur aide sur l'ajustement de certains protocoles et Mohammed pour m'avoir appris les bases de la bio-informatique. Je remercie également mes co-thésards et étudiants (anciens et nouveaux), Tony, Julie, Camille, Simon, Kim, Charlotte, Claire, Audrey, David, pour leur accueil, leur aide technique précieuse et leur compagnie.

Je remercie la « Vingt centimes », Sophie et Loren, pour avoir enduré ensemble ces 3 années de thèse. Je vous souhaite bon courage pour la dernière ligne droite !

Je remercie mes deux anciens co-internes, IPR et H-U, Sébastien et Guillaume, pour nos pauses café « très » enrichissante entre autres. Egalement bon courage pour vos thèses respectives !

Je remercie les membres de mon Comité de Suivi de thèse composé des Pr Pascale Serror et Pr Olga Soutourina, et du Dr David Gilot pour leurs regards extérieurs sur mes travaux et leurs indications ayant permis l'avancée de cette thèse.

Je remercie l'équipe opérationnelle d'hygiène du CHU de Rennes (PH, IDE et techniciens), notamment le Pr Pierre-Yves Donnio, m'ayant, depuis le début de mon internat, permis d'entamer une carrière hospitalo-universitaire. Je vous remercie pour vos précieux conseils et m'avoir permis de dégager du temps pour la recherche. Je remercie également le Dr Amélie Morin-Le Bihan, pour ses conseils avisés et sa gentillesse.

Je remercie le laboratoire associé des entérocoques du CNR « résistance aux antibiotiques » (PH, ingénieur et techniciens), notamment le Dr Francois Guerin, pour ses protocoles, son expertise technique et avoir continué mes expériences sur *ern0030*. Merci à toi pour avoir résolu mes difficultés expérimentales.

Je remercie chaleureusement mes deux collaborateurs, le Pr Wilma Ziebuhr et le Dr Freya Wencker de l'université de Würzburg, pour leur implication sur la description du T-box riboswitch.

Je remercie le Dr Emmanuel Piednoir, pour m'avoir intégré dans l'équipe du CPias Bretagne lors de la crise Covid-19, m'ayant permis de bien m'occuper lors du 1<sup>er</sup> confinement et de gagner grandement en expérience. Je remercie particulièrement l'équipe de Rennes (Fabienne, Marie-Alix, Margaux, Jocelyne et Marion) pour la synergie dégagée dans la gestion de cette crise.

Je remercie le Pr Frédéric Barbut de me permettre de continuer dans la voie hospitalo-universitaire à l'hôpital St Antoine à Paris. Je vous remercie pour les projets que vous me proposez ainsi que le temps que vous m'avez laissé pour terminer cette 4<sup>e</sup> année de thèse. Je remercie également le reste de l'équipe d'hygiène hospitalière (Sarah, Jeanne, Gil, Cyril, Nathalie, Hélène) pour leur accueil chaleureux. Je remercie, au niveau des laboratoires, Christelle, Laurence, Rabab et Anlyata pour leur bienveillance et me faciliter la vie au quotidien coté hôpital.

Je remercie également l'équipe de microbiologie de mon ancienne faculté de pharmacie, Paris Descartes, pour m'avoir accueilli maintenant en tant qu'AHU.

Je remercie tous mes amis (Axelle, Romain, Jérémy, LH, Repié, Ravinet, Virginie, Jérôme Baptiste et les autres...) pour les bons moments passés ensemble et me permettre de penser à autre chose que le boulot.

Je remercie toute ma famille pour votre soutien lors de toutes ces années. Je remercie également ma belle-famille pour la place qu'elle m'a accordée au sein de celle-ci.

Enfin, je remercie ma femme, Céline, pour avoir pris le temps de relire ce manuscrit mais surtout pour le bonheur et le soutien que tu m'apportes au quotidien.

## Liste des Abréviations

5'UTR : Région non traduite en 5' (5'Untranslated region)

A : Adénine

AA : Acide aminé ou amino-acide

AAC : Aminoside acétyl-transférase

aaRS : Aminoacyal-ARNt synthétase ou ligase

ADN : Acide désoxyribonucléique

AdoCbl : Adénosyl-cobalamine

Ala : Alanine

ANT : Aminoside nucléotidyl-transférase

APH : Aminoside phospho-transférase

ARN : Acide ribonucléique

ARNm : ARN messenger

ARNr : ARN ribosomal

ARNreg : ARN régulateur

ARNt : ARN de transfert

Asn : Asparagine

Asp : Aspartate

Axe-Txe : Antitoxine protéique entéroccocque-toxine protéique entéroccocque

BHRe : Bactérie Hautement Résistante emergente

BMR : Bactérie Multi-Résistante

BTR : Bio-reactor Tigecycline-Resistant strain

C : Cytosine

CC : Complexe clonal

CHU : Centre hospitalo-universitaire

CMI : Concentration Minimal Inhibitrice

CNR : Centre National de Référence  
 COG : Cluster of orthogolous groups of protein  
 CPT : Centre de la peptidyl-transférase (Peptidyl-transferase center)  
 CRISPR : Clustered regulatory interspaced short palindromic repet  
 di-GMPc : Diguanosine monophosphate cyclique  
 EF : Facteur d'él<sup>o</sup>ngation (Elongation factor)  
 EGM : Elément Génétique Mobile  
 EPC : Entérobactérie productrice de carbapénémase  
*ern* : Enterococcus ribonucleic acid  
 ERV/ERG : E. *faecium* résistant à la vancomycine / aux glycopeptides  
 ES : Début de phase stationnaire (Early stationnary phase)  
 FC : Fold change  
 fMet : N-formylméthionine  
 FMN : Flavine mononucléotide  
 G : Guanine  
 HCSP : Haut conseil de Santé publique  
 His : Histidine  
 ICE : Integrative and conjugative element  
 IR : Séquence inversée répétée (Inverted repet)  
 IS : Séquence d'insertion (Insertion sequences)  
 IF : Facteur d'initiation (Initiation factor)  
 kb : Kilobase  
 Lac : Lactate  
 LE : Fin de phase exponentielle (Late exponential phase)  
 MAPS : MS2-affinity purification coupled with RNA sequencing  
 ME : Milieu de phase exponentielle (Middle exponential phase)  
 MFS : Major facilitator super family

MLS : Macrolides-Lincosamides-Streptogramines

nt : Nucléotide(s)

OMP : Orotidine monophosphate

ORF : Cadre ouvert de lecture (Open readinframe)

pb : Paire de bases

PhLOPS<sub>A</sub> : Phénicolé, Lincosamide, Oxazolidonone, Pleuromutiline, Streptogramine A

PLP : Protéine liant les pénicillines

PRP : Plasmide sensible aux phéromones (Pheromone-responsive plasmid)

PRPP : 5-phospho- $\alpha$ D-ribose 1-diphosphate

qPCR : Amplification génique quantitative (quantitative Polymerase Chain Reaction)

RACE : Amplification rapide des extrémités de l'ADN circulaire (Rapide amplication of cDNA ending)

RBS : Site de fixation du ribosome (Ribosomal Binding Site)

RCR : Réplication en cercle roulant (Rolling circle replication)

REP : Repetitive Extragenic Palindrome

RNAseq : Séquençage d'ARN

RPKM : Expression transcriptomique normalisée (Reads Per Kilobase per Million mapped reads)

RPP : Protéine de protection ribosomale (Ribosomal protection protein)

SAM : S-adenosyl methionine

SARM : S. aureus résistant à la méticilline

SD : Shine-Dalgarno

Ser : Sérine

T : Thymine

T-AT : Toxine-Antitoxine

TBDB : T-box riboswitch annotation database

TPP : Pyrophosphate de thiamine



Trp : Tryptophane

TSP : Tunnel de sortie peptidique (Nascent peptid exit tunnel)

Tyr : Tyrosine

U : Uracile

UMP : Uracile monophosphate

## Introduction

*Enterococcus faecium* est une bactérie largement impliquée dans les infections nosocomiales et les épidémies hospitalières, notamment par sa grande multi-résistance aux antibiotiques. Depuis des dizaines d'années, des souches résistantes aux glycopeptides (ERG) ont largement diffusé dans le monde et ne présentent que peu d'alternatives thérapeutiques. Les déterminants impliqués dans leur pouvoir épidémiogène et leur capacité d'adaptation à l'environnement hospitalier restent encore à déterminer. Une des pistes de réflexion est l'étude des ARN régulateurs, acteurs-clés dans l'adaptation dans l'environnement. Récemment, les premiers ARN régulateurs ont été décrits chez *E. faecium* en lien avec l'adaptation à la daptomycine, un antibiotique de dernier recours utilisé lors des infections aux ERG.

Ce travail de thèse s'inscrit dans la thématique de recherche de l'UMR\_S1230 « ARN régulateurs bactériens et médecine » (Université Rennes 1/Inserm). En collaboration avec le laboratoire « entérocoques » associé au Centre National de Référence « Résistance aux antibiotiques », l'unité a étendu l'exploration de la fonction des ARN régulateurs à *E. faecium* notamment en lien avec l'antibiorésistance.

L'objectif du travail présenté est l'investigation fonctionnelle de ces nouveaux ARN régulateurs décrits par une approche génétique, transcriptomique, phénotypique et mécanistique. La première partie bibliographique présentera la place d'*E. faecium* en tant que pathogène opportuniste, multi-résistant et adaptable. Nous reviendrons ensuite sur les différents mécanismes de régulation impliquant les ARN, notamment les mécanismes d'atténuation, les riboswitchs et les ARN régulateurs. La seconde partie présentera les différents résultats de nos travaux, de la caractérisation générale des différents ARN à l'investigation fonctionnelle approfondie de deux d'entre eux. Enfin, nous remettrons en perspective les différentes pistes de réflexions ouvertes par ce travail.

## Analyse bibliographique

## I. *Enterococcus faecium*, pathogène hospitalier majeur

### A. Caractéristiques microbiologiques

*E. faecium* est un cocci à Gram positif appartenant au genre *Enterococcus*. Les entérocoques sont des bactéries non sporulées, non capsulées, dépourvues d'oxydase et de catalase, avec une respiration anaérobie-aérobie tolérant, un génome comprenant entre 37 et 45 % de (G+C) et possédant généralement l'antigène D de la classification de Lancefield (Lancefield, 1933; Schleifer et Kilpper-Bälz, 1984; Murray, 1990; Klein, 2003). Les entérocoques appartiennent à la famille des *Enterococcaceae*, à l'ordre des *Lactobacillales*, à la classe des *Bacilli* et au phylum des *Firmicutes*. Ce genre bactérien était initialement inclus dans le genre *Streptococcus* mais se distingue par diverses spécificités comme leur capacité de croissance (entre 10 et 45°C, en milieu hyper-salé, en présence de 40% de sels biliaires et à pH 9,6) et leurs caractéristiques biochimiques comme l'hydrolyse de l'esculine (Sherman, 1937; Schleifer et Kilpper-Bälz, 1984; Murray, 1990; Leclerc *et al.*, 1996; Klein, 2003). Ce sont des commensaux du tube digestif chez l'homme et diverses espèces animales (Klein, 2003; Van Tyne et Gilmore, 2014).

Ces bactéries ont une bonne capacité d'adaptation, notamment au stress oxydant, (Verneuil *et al.*, 2006; Carbona *et al.*, 2007; Kajfasz *et al.*, 2012), aux variations thermiques et ioniques (Laport *et al.*, 2003; Inouye et Phadtare, 2004; Gaca et Lemos, 2019). Ces différents éléments expliquent la longue survie potentielle des entérocoques sur les surfaces, notamment *E. faecium*, pouvant aller jusqu'à 4 mois selon les conditions (Wendt *et al.*, 1998; Neely et Maley, 2000; Hanczvikkel et Tóth, 2018). Ces caractéristiques en font des bactéries ubiquitaires largement retrouvées dans l'environnement (Pinto *et al.*, 1999; Byappanahalli *et al.*, 2012).

## B. Epidémiologie moléculaire : commensal ou pathogène

Malgré le caractère commensal et ubiquitaire de ce genre bactérien, certaines espèces comme *E. faecium* sont impliquées dans les infections humaines. Une dualité existe chez les entérocoques entre intérêts et risques infectieux (Kim et Marco, 2014; Isnard, 2017). En effet, leurs caractères intrinsèques en font de bons marqueurs fécaux de l'environnement (Leclerc *et al.*, 1996; Boehm et Sassoubre, 2014) mais également de potentiels probiotiques voire oncobiotiques (Canani *et al.*, 2007; Daillère *et al.*, 2016; Celiberto *et al.*, 2017; Holzapfel *et al.*, 2018). Leurs capacités d'adaptation à l'environnement associées à la production de bactériocines permettent une utilisation possible en agro-alimentaire (Foulquié Moreno *et al.*, 2006; Bonacina *et al.*, 2017).

Cependant, cette capacité d'adaptation est aussi un élément favorisant la pathogénicité notamment via l'acquisition de facteurs de virulence pour *Enterococcus faecalis* ou de déterminants d'antibiorésistance pour *E. faecium* (Van Tyne et Gilmore, 2014). Depuis longtemps, les entérocoques se sont adaptés selon leur hôte et leur environnement (Lebreton *et al.*, 2017). Ces bactéries peuvent moduler efficacement leur métabolisme et possèdent une adaptabilité génomique importante.

Chez *E. faecium*, les souches peuvent être séparées en 3 groupes (clades) selon leurs pathogénicités et leurs niches écologiques (Lebreton *et al.*, 2013) :

- Clade A1 : souches provoquant des infections chez l'homme, largement impliquées dans les infections nosocomiales et les épidémies hospitalières (appartenant au complexe clonal 17 [CC17]),
- Clade A2 : souches retrouvées chez l'animal,
- Clade B : souches commensales humaines rarement impliquées en pathologie infectieuse.

Les déterminants impliqués dans la pathogénie d'*E. faecium* sont encore mal connus. Pourtant, certaines souches, notamment celles appartenant au CC17, se sont particulièrement adaptées à l'environnement hospitalier et ont pu largement diffuser (Lee *et al.*, 2019).

## C. Place en infectiologie

### 1. Epidémiologie des infections et résistances

#### a. Infections et hôpital

Les espèces d'entérocoques principalement impliquées dans les infections humaines sont *E. faecalis* et *E. faecium* avec respectivement 80% et 20% d'implication. Ces pathogènes sont majoritairement responsables d'infections urinaires, d'infections intraabdominales (souvent polymicrobiennes), de bactériémies et d'endocardites (Selton-Suty *et al.*, 1997; Murdoch *et al.*, 2009; Vincent *et al.*, 2009; Arias et Murray, 2012; Ayobami *et al.*, 2020). Leur adaptabilité et ubiquité en font des pathogènes opportunistes majeurs dans les infections nosocomiales. La dernière enquête nationale de prévalence des infections nosocomiales de 2017 situe *E. faecalis* à la 3<sup>ème</sup> position et *E. faecium* à la 11<sup>ème</sup> position des micro-organismes causant respectivement 6,5 % et 1,5 % des infections nosocomiales en France. En cinq ans, leurs prévalences ont augmenté d'environ 60 % en France (Santé Publique France, 2019) [Tableau 1]. Aux Etats-Unis, ces deux espèces se retrouvent respectivement à la 5<sup>ème</sup> et 9<sup>ème</sup> position des micro-organismes causant des infections nosocomiales (Weiner *et al.*, 2016) [Tableau 1].

Tableau 1. Place des deux principales espèces d'entérocoques dans les micro-organismes responsables d'infections nosocomiales.

Rang		<i>E. faecalis</i>		<i>E. faecium</i>	
		2012	2017	2012	2017
France	Infections sur site opératoire	5 <sup>e</sup>	4 <sup>e</sup>	13 <sup>e</sup>	8 <sup>e</sup>
	Bactériémie	8 <sup>e</sup>	6 <sup>e</sup>	13 <sup>e</sup>	11 <sup>e</sup>
	Infections urinaires	4 <sup>e</sup>	2 <sup>e</sup>	14 <sup>e</sup>	13 <sup>e</sup>
		2011-2014			
Etats-Unis	Infections sur site opératoire	4 <sup>e</sup>		11 <sup>e</sup>	
	Bactériémies sur cathéter central	5 <sup>e</sup>		11 <sup>e</sup>	
	Infections urinaires sur sonde	3 <sup>e</sup>		5 <sup>e</sup>	

La diversité des complexes clonaux d'*E. faecalis* retrouvés responsables d'infections chez l'homme démontre un certain potentiel virulent de l'espèce (McBride *et al.*, 2007). En effet, *E. faecalis* possède plusieurs facteurs de virulence comme des cytolysines, des adhésines ou des protéases favorisant sa pathogénèse (Jett *et al.*, 1994). Contrairement à *E. faecalis*, les souches d'*E. faecium* prédominantes dans les infections nosocomiales appartiennent à un seul complexe clonal (CC17), souches particulièrement adaptées à l'environnement hospitalier (Arias et Murray, 2012; Van Tyne et Gilmore, 2014; Lee *et al.*, 2019). Ces souches semblent contenir les éléments génétiques nécessaires pour dominer les infections à *E. faecium* comme un îlot de pathogénicité (contenant les gènes de virulence *esp* et *hyl<sub>Efm</sub>*) et une résistance de haut niveau à l'ampicilline et aux fluoroquinolones (Leavis *et al.*, 2004; Cattoir et Leclercq, 2010). Mais les déterminants expliquant le fort potentiel épidémique de ce complexe clonal ne sont pas connus, montrant l'intérêt de s'intéresser particulièrement à CC17 (Lee *et al.*, 2019). Si *E. faecalis* est le versant « virulence » des entérocoques, *E. faecium* en est le versant « résistance ».



### *b. Acquisition de résistance*

L'acquisition de résistance aux anti-infectieux par les micro-organismes reste une problématique majeure aujourd'hui. Cette acquisition de résistance est un facteur majeur d'échec thérapeutique. Cette omniprésence dans les infections nosocomiales des bactéries multi-résistantes (BMR), dont *E. faecium* fait partie, augmente le risque de morbi-mortalité (Lautenbach *et al.*, 1999; Klevens *et al.*, 2007; Colomb-Cotin *et al.*, 2016).

Contrairement à *E. faecalis*, *E. faecium* est globalement plus difficile à éradiquer avec une résistance acquise à l'ampicilline largement répandue (80 à 90 % des souches cliniques). De plus, *E. faecium* semble avoir une meilleure capacité d'acquisition et d'expression des gènes de résistance, dont les plus inquiétantes actuellement sont en lien avec la résistance aux glycopeptides (Lee *et al.*, 2019). Ces souches dites ERG ne laissent que peu d'alternatives thérapeutiques. Elles sont apparues dans les années 1980 en Europe et ont depuis largement diffusé, notamment aux Etats-unis (Willems et Bonten, 2007). En France, ces souches sont moins répandues mais certains pays d'Europe connaissent une augmentation drastique de ces souches [Figure 1]. Aux Etats-Unis, ces souches représentent aujourd'hui près de 80 % des infections aux entérocoques (O'Driscoll et Crank, 2015). D'autres résistances sont souvent associées, notamment pour les souches appartenant au CC17. Depuis quelques années, ces souches d'*E. faecium* du CC17 tendent à dépasser en nombre d'infections associées aux soins celles provoquées par *E. faecalis* (Top *et al.*, 2007). Les souches d'*E. faecium* résistant à la vancomycine (ERV/ERG) appartiennent au groupe de bactéries pathogènes ESKAPE [*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, *Enterobacter* spp.] pour lequel il devient urgent de développer de nouvelles thérapies anti-infectieuses afin d'éviter une impasse thérapeutique (Rice, 2008; Oliveira *et al.*, 2020).

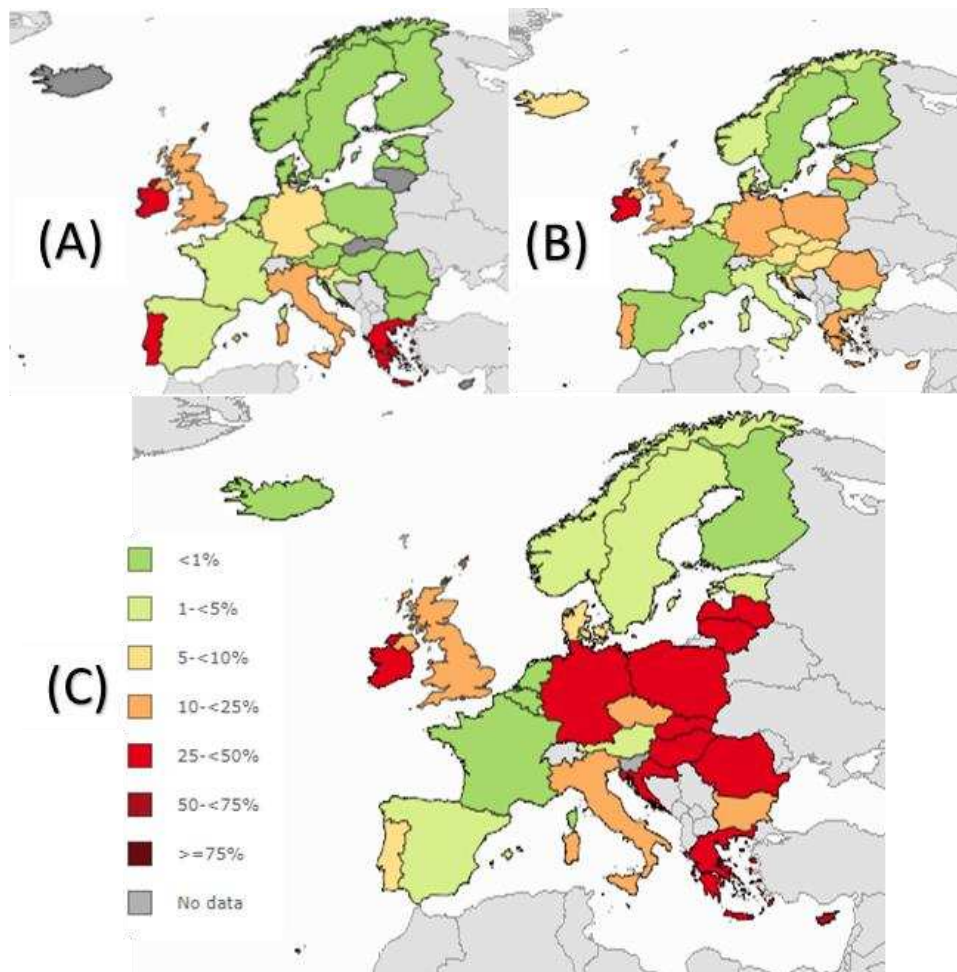


Figure 1. Pourcentage d'ERV parmi les souches invasives d'*E. faecium* dans les pays d'Europe en 2006 (A), 2013 (B) et 2019 (C) (<http://atlas.ecdc.europa.eu/public/index.aspx>)

## 2. Prise en charge

### a. Antibiothérapie

Le traitement de référence des infections à *E. faecalis* et à *E. faecium* est respectivement l'amoxicilline et la vancomycine. Pour les souches résistantes aux glycopeptides, les antibiotiques de deuxième ligne pouvant être utilisés sont principalement la daptomycine, le linézolide ou les tétracyclines de 3<sup>ème</sup> génération (ex. tigécycline). Selon la gravité de l'infection, une association avec un aminoside (notamment la gentamicine) est possible (Kristich *et al.*, 2014; O'Driscoll et Crank, 2015).

### *b. Gestion épidémique*

Les souches d'*E. faecium* résistantes sont particulièrement surveillées en France pour limiter leur dissémination. La gestion des ERV est importante à cause du risque d'impasse thérapeutique, de l'augmentation de la morbi-mortalité associée à l'acquisition de BMR et du risque de transfert horizontal de résistance aux staphylocoques (Lautenbach *et al.*, 1999; Colomb-Cotin *et al.*, 2016; McGuinness *et al.*, 2017). Les ERV, particulièrement ceux appartenant au CC17, sont la cause de nombreuses épidémies liées aux transmissions croisées par les soignants et l'environnement (Arias et Murray, 2012). La maîtrise de ce risque épidémique repose majoritairement sur le bon usage des antibiotiques et le respect des mesures de prévention des transmissions croisées (Goossens *et al.*, 2005). En France, des mesures spécifiques (mesures BHRe soit Bactérie Hautement Résistante émergente) s'appliquent aux ERV (HCSP, 2019). Elles reposent principalement sur une politique de dépistage des patients contacts, le port d'équipements de protection individuelle adaptés aux professionnels de santé, une gestion maîtrisée des transferts ainsi qu'un bio-nettoyage accru (HCSP, 2013, 2019). Ces mesures permettent de limiter le risque de diffusion épidémique, expliquant peut-être la relative maîtrise de la diffusion des ERV en France comparativement à d'autres pays [Figure 1].

## D. Résistances aux antibiotiques

Les antibiotiques sont des molécules pouvant inhiber la croissance bactérienne (bactériostatique) ou provoquer la destruction (bactéricide) des bactéries. Ces antibiotiques agissent selon 2 grands types de mécanismes : soit une altération du métabolisme bactérien, soit une altération directe de la membrane bactérienne. La plupart des antibiotiques appartiennent à la première catégorie : ils inhibent une enzyme impliquée dans une voie métabolique indispensable pour la croissance ou la survie de la bactérie comme la synthèse de la paroi bactérienne, la synthèse des acides nucléiques ou la synthèse des protéines. Ces antibiotiques ont des spectres d'activité définis par les caractéristiques intrinsèques des espèces bactériennes dites « résistances naturelles ». Pour échapper à ces molécules, les bactéries ont pu développer au fil du temps des mécanismes dits « de résistance acquise ». La résistance des micro-organismes est dérivée de trois mécanismes généraux de résistance :

- La diminution de la concentration intracellulaire d'antibiotique par imperméabilité membranaire et/ou efflux actif.
- L'inactivation et/ou la dégradation de l'antibiotique par une enzyme.
- Modification et/ou protection de la cible de l'antibiotique réduisant son affinité pour sa cible.

### 1. $\beta$ -lactamines

La famille des  $\beta$ -lactamines est la première famille d'antibiotiques utilisée en médecine humaine. Ces antibiotiques inhibent la synthèse du peptidoglycane, important composant de la paroi bactérienne. Cette inhibition se fait par fixation des  $\beta$ -lactamines sur des enzymes nommées PLP (Protéines liant les pénicillines) impliquées dans les étapes « d'assemblage » du peptidoglycane comme les transpeptidases, les transglycosylases et les carboxypeptidases. Chez *E. faecium*, la résistance aux  $\beta$ -lactamines est principalement

médiée par un mécanisme de modification de cible. En effet, les entérocoques produisent une PLP de faible affinité (PLP-5) pour les  $\beta$ -lactamines, qui est notamment impliquée dans la résistance intrinsèque de ces espèces aux céphalosporines (Fontana *et al.*, 1994; Sifaoui *et al.*, 2001; Rice *et al.*, 2004). Une résistance de haut niveau à l'amoxicilline peut être acquise chez *E. faecium* en cas de modifications qualitatives et/ou quantitatives de cette PLP-5 (Montealegre *et al.*, 2017). Cette résistance est acquise pour une majorité (80-90 %) des souches cliniques d'*E. faecium* en France.

## 2. Glycopeptides

Les glycopeptides sont des molécules de haut poids moléculaire (1500 à 1900 kDa) bloquant la synthèse du peptidoglycane en fixant l'extrémité D-Ala-D-Ala des précurseurs du peptidoglycane. Cet encombrement stérique empêche l'action des enzymes impliquées dans la synthèse de la membrane. Les entérocoques peuvent acquérir une résistance aux glycopeptides par modification des extrémités D-Ala-D-Ala de ces précurseurs par des motifs D-Ala-D-Lac ou D-Ala-D-Ser. Chez *E. faecium*, la résistance aux glycopeptides est majoritairement médiée par les opérons *vanA* ou *vanB* respectivement portés par les transposons Tn1546 et de type Tn1549 (Noble *et al.*, 1992; Arthur *et al.*, 1993; Rice *et al.*, 1998; Garnier *et al.*, 2000; Courvalin, 2006). Ces opérons comprennent 5 gènes impliqués dans le mécanisme de résistance et 2 gènes de régulation. Le mécanisme de résistance repose sur une modification de la cible en empêchant la formation de précurseurs naturels (avec une extrémité D-Ala-D-Ala) par VanX et VanY [Figure 2]. Les enzymes VanH et VanA/B vont respectivement permettre la déshydrogénation de l'Alanine en Lactate et la ligation sur l'extrémité du précurseur de tel sorte à obtenir une extrémité D-Ala-D-Lac (Cattoir et Leclercq, 2010).

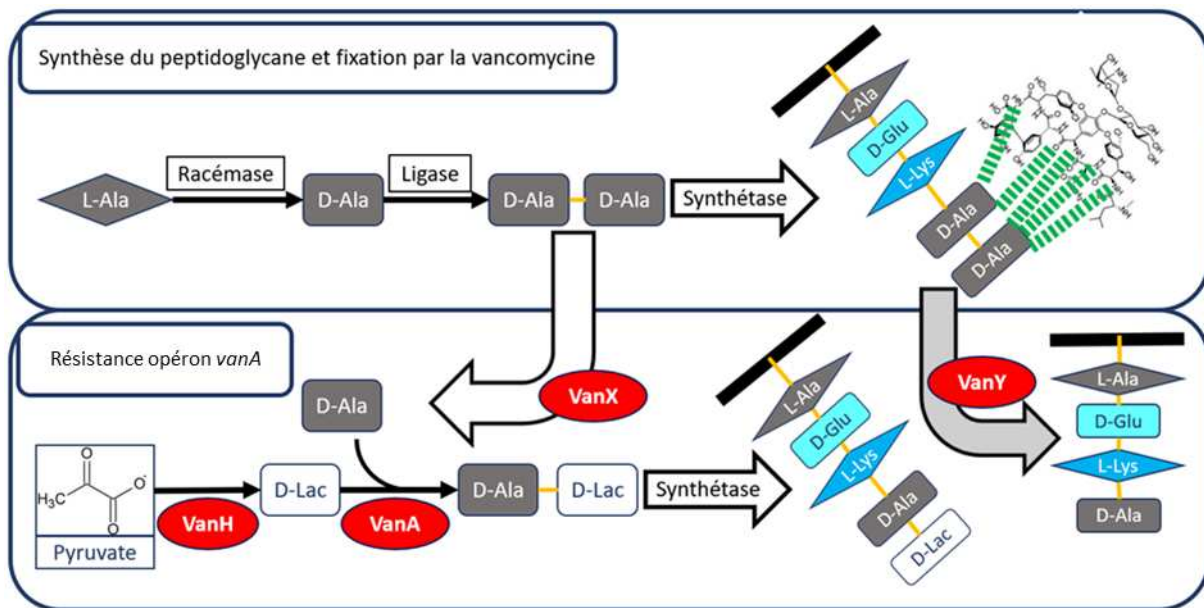


Figure 2. Description de la résistance à la vancomycine par *vanA* (Cattoir et Leclercq, 2010; Le Neindre, 2018)

Cette extrémité D-Ala-D-Lac provoque une diminution d'affinité à la vancomycine d'environ 1000 fois mais peut toujours être reconnue par les enzymes impliquées dans la synthèse du peptidoglycane (Cattoir et Leclercq, 2013). Le gène *vanZ* de fonction inconnue retrouvé dans l'opéron *vanA* semble être impliqué dans la résistance à la téicoplanine (Arthur *et al.*, 1995; Vimberg *et al.*, 2020). Cette résistance à la téicoplanine n'est pas acquise avec l'opéron *vanB* dû à une absence d'induction de l'opéron par la téicoplanine. L'apparition de ces souches d'ERV pose problème à cause du manque d'alternatives thérapeutiques et des transferts génétiques horizontaux pouvant exister vers d'autres pathogènes bactériens comme *S. aureus* (Noble *et al.*, 1992; McGuinness *et al.*, 2017).

### 3. Quinolones

Les fluoroquinolones sont des antibiotiques initialement utilisés comme anti-Gram négatifs. Vers les années 2000, des fluoroquinolones dites à spectre étendu, ont été développées pour avoir une activité sur les bactéries à Gram positif, notamment le pneumocoque. Ces molécules inhibent l'activité des topoisomérases de type II comme l'ADN gyrase (GyrA<sub>2</sub>GyrB<sub>2</sub>) et la topoisomérase IV (ParC<sub>2</sub>ParE<sub>2</sub>), altérant ainsi la réplication bactérienne. L'activité de ces molécules est limitée vis-à-vis des entérocoques, notamment par la production intrinsèque de protéines Qnr-like (Arsène et Leclercq, 2007; Hegde *et al.*, 2011; Sinel *et al.*, 2017b). Les protéines Qnr sont des protéines chaperonnes protégeant les enzymes cibles pour empêcher la fixation des fluoroquinolones sur celles-ci (Tran et Jacoby, 2002; Hooper et Jacoby, 2015). Les entérocoques peuvent également acquérir une haute résistance aux fluoroquinolones par modifications de la cible accompagnées ou non de l'hyper-expression de systèmes d'efflux (Oyamada *et al.*, 2006).

### 4. Antibiotiques inhibiteurs de la synthèse protéique

La synthèse protéique est indispensable pour le développement et le fonctionnement d'une cellule bactérienne. Cette voie est donc une cible idéale pour les antibiotiques (McCoy *et al.*, 2011). La synthèse protéique nécessite l'intervention de différents acteurs avec comme élément-clé, le ribosome. C'est une structure particulière composée d'ARN ribosomaux et de protéines ribosomales. Le ribosome permet la liaison d'acides aminés à partir d'une matrice, l'ARN messager.

Le ribosome bactérien 70S est composé de 2 sous-unités [Figure 3] :

- La sous-unité 50S formée des ARNr 23S et 5S et d'un ensemble de protéines ribosomales. L'ARNr 23S peut être délimité en 6 domaines. Cette sous-unité est impliquée

dans la liaison peptidique entre les différents acides aminés apportés par les ARN de transfert (ARNt).

- La sous-unité 30S : composée de l'ARNr 16S et d'un ensemble de protéines. Cette sous-unité est impliquée dans la présentation des codons aux anticodons des ARNt amino-acylés et fait donc le lien entre les ARNt et l'ARNm. La sous-unité 30S peut être séparée en 3 sites :

- le site P permet l'initiation de la traduction par fixation de l'ARNt initiateur (ARNt<sup>fMet</sup>) et accueille l'ARNt chargé par le peptide en cours de formation lors de l'élongation.
- le site A « vérifie » la cohérence entre le codon et l'ARNt amino-acylé présenté.
- le site E accueille l'ARNt « dé-aminoacylé » avant de le « relarguer ».

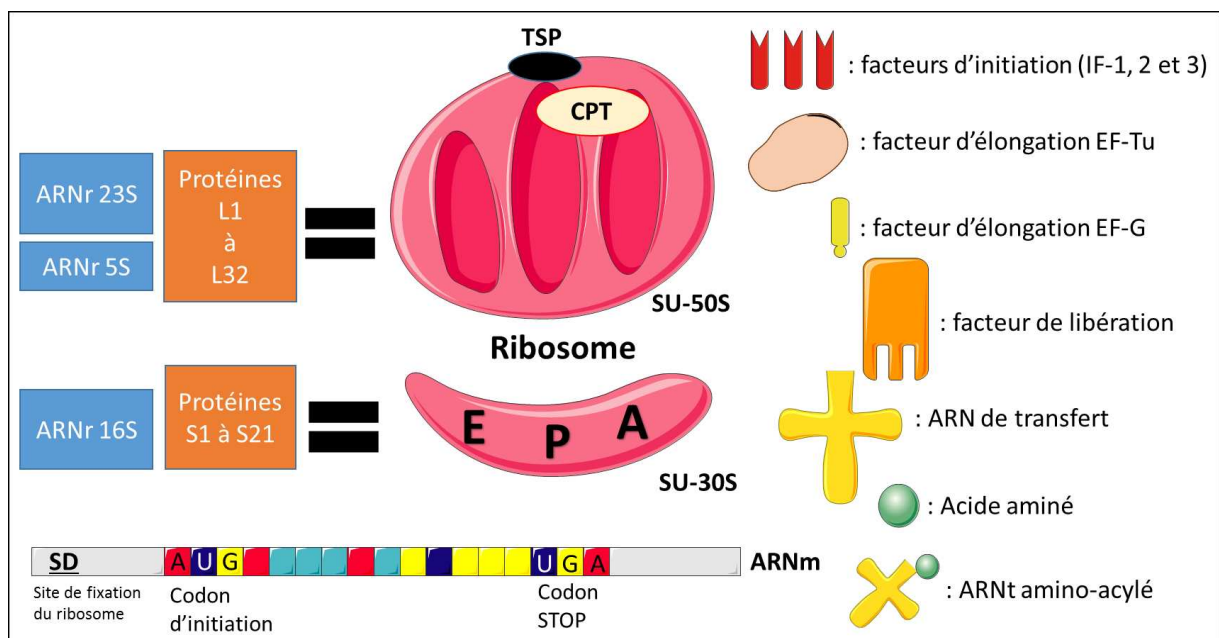


Figure 3. Description des acteurs impliqués dans la synthèse peptidique.

SD : Shine-Dalgarno ; TSP : Tunnel de sortie peptidique ; CPT : Centre de la peptidyl-transférase.



La traduction comporte 3 étapes majeures : l'initiation, l'élongation et la terminaison. L'initiation sera enclenchée avec la reconnaissance de la séquence d'initiation, nommée RBS (Ribosomal Binding Site), sur l'ARNm par l'ARNr 16S et des facteurs d'initiation [Figure 4, étape 1]. Cette séquence spécifique comprend la séquence Shine-Dalgarno (SD) riche en purines et le codon d'initiation (Londei, 2005). L'ARN de transfert chargé par l'acide aminé formyl-méthionine ( $\text{ARNt}^{\text{fMet}}$ ) va fixer le codon d'initiation au niveau du site P de la sous-unité 30S [Figure 4, étape 2]. La sous-unité 50S sera alors recrutée formant ainsi le complexe ribosomal 70S [Figure 4, étapes 3 et 4].

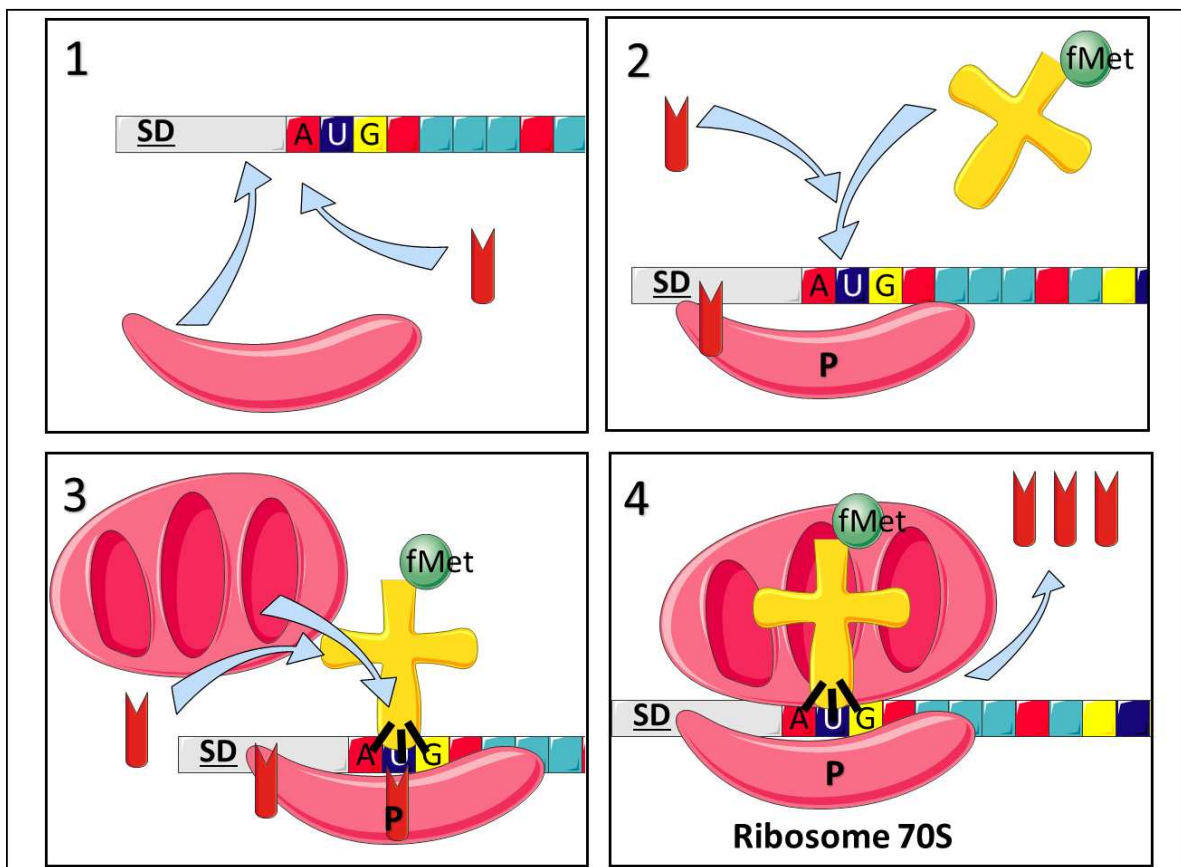


Figure 4. Synthèse peptidique : Initiation de la traduction.

- (1) Fixation de la sous-unité 30S au niveau de la séquence de Shine-Dalgarno à l'aide d'IF-1.
- (2) Fixation de l'ARNt amino-acylé avec fMet ( $\text{ARNt}^{\text{fMet}}$ ) sur le codon d'initiation à l'aide d'IF-2.
- (3) Recrutement de la sous-unité 50S à l'aide d'IF-3.
- (4) Complexe ribosomal 70S stabilisé.

L'élongation peut être séparée en différentes étapes [Figure 5] : le décodage de l'ARN, l'accommodation de l'ARNt, la liaison peptidique et la translocation (Schmeing et Ramakrishnan, 2009). La première étape se situe au niveau du site A de la sous-unité 30S avec le recrutement de l'ARNt amino-acylé correspondant au codon situé sur le site A [Figure 5, étape 1]. L'ARNt amino-acylé recruté forme déjà un complexe avec un facteur d'élongation EF-Tu : aaARNt/EF-Tu (Nissen *et al.*, 1996). La bonne correspondance entre l'ARNt et le codon va enclencher l'étape d'accommodation. Cette étape induit un changement conformationnel de l'ARNr 16S, notamment à l'aide du facteur EF-Tu, stabilisant le complexe pour l'étape de translocation (Ogle *et al.*, 2001; Zaher et Green, 2009). En l'absence d'une interaction suffisante entre le codon de l'ARNm et l'anticodon de l'ARNt, l'ARNt sera délogé pour faire place à un autre ARNt plus « adapté ». L'étape suivante comprend la liaison peptidique, le déplacement de l'ARNt vers le site E et le passage du peptide en cours de formation *via* le tunnel de sortie peptidique. La catalyse de la liaison peptidique s'effectue au niveau du centre de la peptidyl transférase dans la sous-unité 50S [Figure 5, étape 2]. La conformation de la structure ribosomique permet une proximité des 2 AA chargés sur les ARNt. L'ARNr 23S va alors catalyser la liaison peptidique de l'AA chargé par l'ARNt du site P avec l'AA chargé par l'ARNt du site A (Nissen *et al.*, 2000). Ce peptide néo-synthétisé est positionné au niveau du tunnel de sortie peptidique (constitué d'une partie de l'ARNr 23S et de protéines ribosomales notamment L4 et L22) permettant de faire sortir du ribosome l'extrémité N-terminale du peptide. La présence du peptide au niveau de l'ARNt positionné sur le site A n'est pas thermodynamiquement favorable. Cela va induire une délocalisation catalysée par un facteur d'élongation EF-G de l'ARNt libre du site P au site E et de l'ARNt chargé du peptide néo-synthétisé du site A vers le site P [Figure 5, étape 3] (Moazed et Noller, 1989; Rodnina *et al.*, 1997). Cette translocation va libérer le site A du ribosome au niveau du codon suivant [Figure 5, étape 4]. Cette étape d'élongation va se répéter jusqu'à l'arrivée d'un codon stop [Figure 5, étapes 5 et 6].

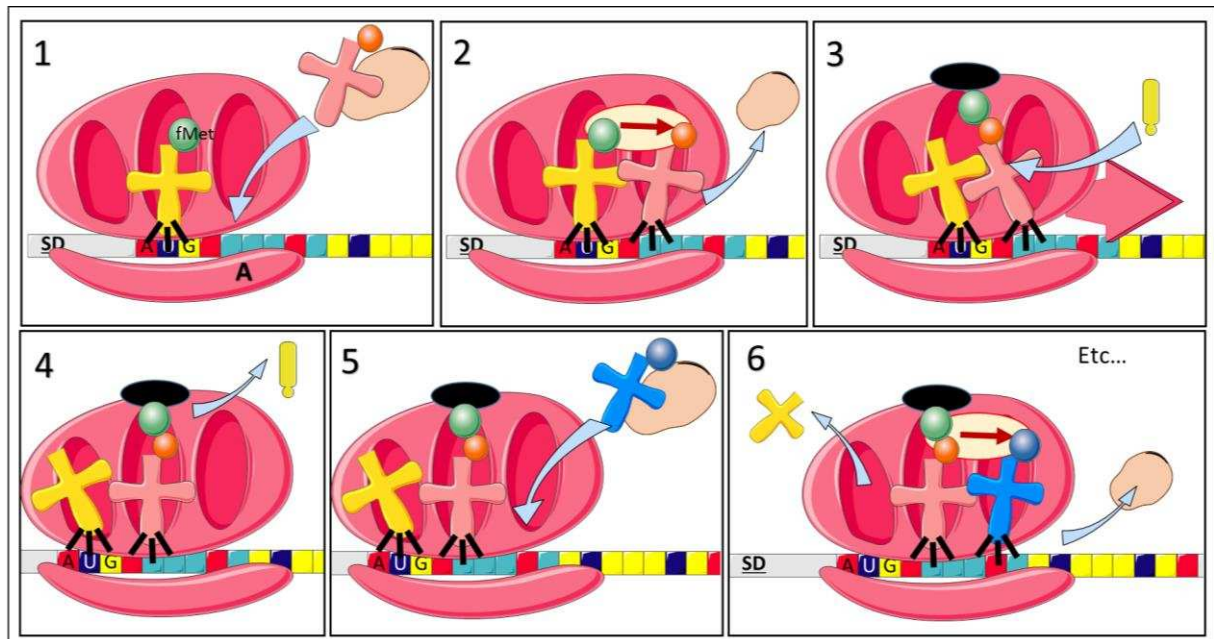


Figure 5. Synthèse peptidique : Elongation.

(1) Décodage de l'ARN et accommodation de l'ARNt à l'aide de EF-Tu.

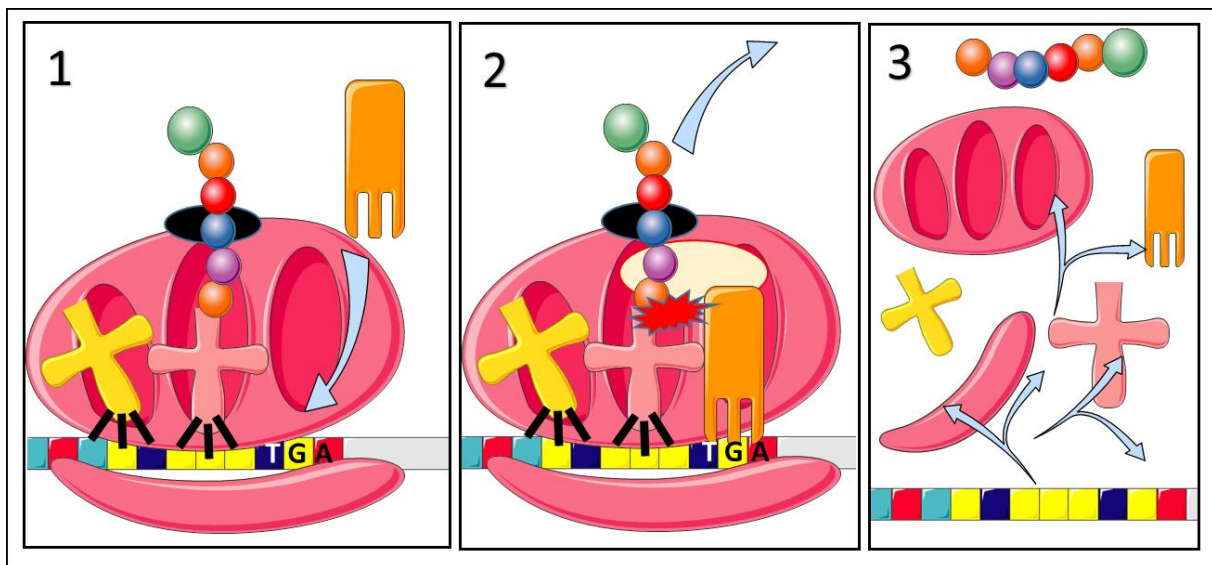
(2) Liaison peptidique catalysée par CPT.

(3) Translocation médiée par EF-G.

(4) Complexe ribosomique stable avec le peptide néo-synthétisé positionné au niveau du TSP.

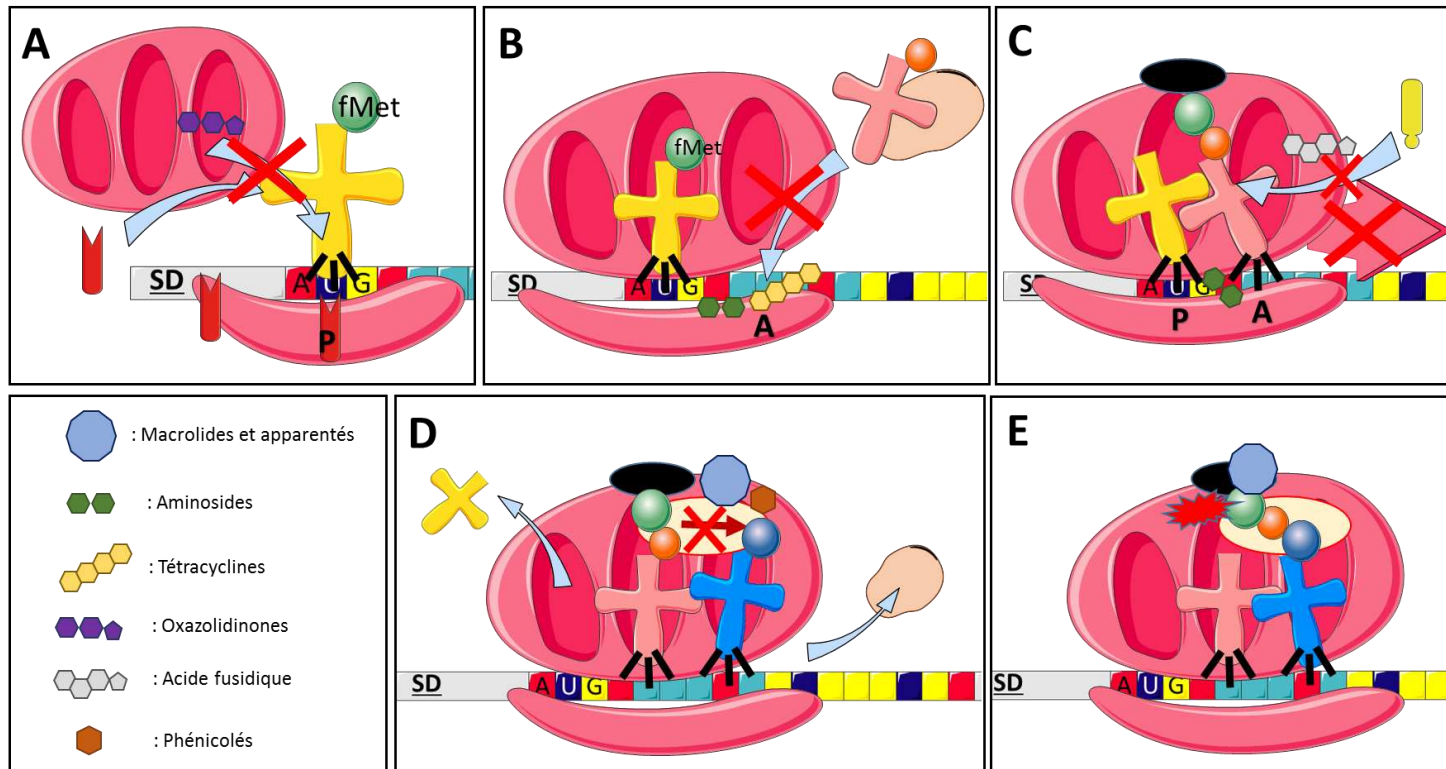
(5 et 6) Redémarrage de l'étape 1 de l'élongation.

La dernière étape de la traduction est la terminaison (Laurberg *et al.*, 2008) [Figure 6]. Le codon stop est reconnu par un facteur de libération au niveau du site A [Figure 6, étape 1]. Cette interaction va provoquer des changements conformationnels favorisant l'hydrolyse de la liaison entre l'ARNt situé sur le site P et le peptide [Figure 6, étape 2]. Le peptide est alors libéré par le tunnel de sortie peptidique et le complexe ribosomal 70S se dissocie [Figure 6, étape 3].



**Figure 6. Synthèse peptidique : Terminaison.**  
 (1) Fixation du facteur de libération sur le codon stop.  
 (2) Clivage entre le peptide néosynthétisé et l'ARNt.  
 (3) Dissociation de tous les acteurs.

Une majorité d'antibiotiques peuvent interférer sur les différentes étapes du processus de traduction [Figure 7]. Les principales familles d'antibiotique impliquées dans ces mécanismes sont les aminosides, les macrolides et apparentés (MLS, Macrolides-Lincosamides-Streptogramines), les tétracyclines, les phénicolés et les oxazolidinones (dont le linézolide).



### a. Aminosides

Les aminosides sont des antibiotiques inhibant la synthèse protéique au niveau de la sous-unité 30S du ribosome bactérien (Becker et Cooper, 2013). Plus précisément, la majorité des aminosides vont se fixer au niveau du site A [Figure 7.B]. Cette fixation va induire un changement conformationnel de l'ARNr 16S. Ce changement conformationnel va mimer l'accommodation induisant par la suite l'étape de translocation. Cette fixation provoque une altération de l'étape de décodage de l'ARNm, induisant la possibilité d'incorporation de « mauvais » AA dans le peptide en cours de synthèse. Le peptide produit contiendra donc des erreurs induisant la production de protéines non fonctionnelles (McCoy *et al.*, 2011). En fonction de la localisation de la fixation sur le site A, l'effet induit peut être différent selon l'aminoside. Par exemple, la fixation sur le site A de la spectinomycine empêchera la translocation de l'ARNt du site A au site P [Figure 7.C] (Borovinskaya *et al.*, 2007).

Les entérocoques ont naturellement une résistance intrinsèque à bas niveau aux aminosides causée par l'entrée cellulaire de ces molécules par des systèmes de transport O<sub>2</sub> dépendants. Ils peuvent tout de même être utilisés en association synergique avec des antibiotiques agissant sur la paroi bactérienne ( $\beta$ -lactamines, glycopeptides), facilitant ainsi l'accès des aminosides au ribosome. *E. faecium* peut également acquérir une résistance de haut niveau par un mécanisme d'inactivation, notamment portée par des éléments génétiques mobiles. Ces résistances sont médiées par l'acquisition d'enzymes de type phospho-transférase (APH), acétyl-transférase (AAC) et nucléotidyl-transférase (ANT) (Ramirez et Tolmasky, 2010; Becker et Cooper, 2013). Cette acquisition de résistance se retrouve souvent chez les souches responsables d'infections hospitalières. Les aminosides inactivés dépendent de l'enzyme impliquée : toto-résistance excepté streptomycine avec l'enzyme bifonctionnelle AAC(6')-Ie-APH(2'')-Ia ; résistance à la tobramycine, kanamycine et netilmicine avec l'enzyme ANT(6')-Ii ; résistance à la tobramycine, kanamycine et

amikacine avec l'enzyme ANT(4')-Ia (Ferretti *et al.*, 1986; Carlier et Courvalin, 1990; Costa *et al.*, 1993). La présence de telles souches à l'hôpital limite l'utilisation des aminosides.

### *b. Macrolides et apparentés*

Les macrolides et apparentés dits MLS (Macrolides-Lincosamides-Streptogramines) sont la classe d'antibiotiques inhibant la traduction bactérienne comportant le plus de molécules différentes. Les molécules apparentées, bien que structurellement différentes, sont regroupées dans ce même groupe MLS du fait de leurs mécanismes d'action proches et de résistances croisées. Ces molécules inhibent la synthèse protéique par fixation à la sous-unité 50S du ribosome et ont des sites de fixation qui se superposent.

Les macrolides se fixent sur l'ARNr 23S au niveau du tunnel de sortie peptidique (TSP) empêchant, par encombrement stérique, l'élongation peptidique quand le peptide néosynthétisé devient trop grand [Figure 7.E] (Hansen *et al.*, 2002; Tenson *et al.*, 2003; McCoy *et al.*, 2011; Vázquez-Laslop et Mankin, 2018). Cela bloque le ribosome sur sa position et peut induire une dissociation prématurée du peptide « atrophié » de l'ARNt. La spiramycine et les lincosamides peuvent fixer des sites supplémentaires au niveau du centre de la peptidyl transférase (CPT) bloquant l'ajout d'AA sur le peptide en cours de synthèse [Figure 7.D] (Poulsen *et al.*, 2000; Tenson *et al.*, 2003).

Les streptogramines sont des associations de 2 composants A et B qui se fixent à des sites différents au niveau de l'ARNr 23S (Reissier et Cattoir, 2020). Les streptogramines A agissent préférentiellement au niveau du CPT (Cocito *et al.*, 1997). La fixation des streptogramines A permet un changement conformationnel favorisant la fixation des streptogramines B (Hansen *et al.*, 2003). Les streptogramines B ont les mêmes sites de fixation que les macrolides et obstruent donc le TSP (Tenson *et al.*, 2003).



Chez *E. faecium*, les résistances aux MLS peuvent être médiées par des mécanismes de modification ou mutation de la cible, d'inactivation enzymatique ou d'efflux actif.

Les résistances par modification de la cible induisent potentiellement des résistances croisées dues à certains sites de fixation communs entre les différentes molécules du groupe des MLS. Cette résistance est notamment médiée par les gènes *erm* codant pour des méthylases modifiant l'A<sub>2058</sub> de l'ARNr 23S bloquant l'accès aux macrolides, lincosamides et streptogramines B (Leclercq, 2002). Chez *E. faecium*, ce mécanisme de résistance est très fréquent et la méthylase impliquée est principalement codée par le gène *erm(B)* provoquant le phénotype de résistance dit MLS<sub>B</sub> (Portillo *et al.*, 2000; Cattoir et Giard, 2014). Des mutations ponctuelles au niveau de la zone d'interaction principalement dans le domaine V de l'ARNr 23S sont également impliquées dans la résistance aux antibiotiques MLS, mais restent rares chez *E. faecium* (Vester et Douthwaite, 2001).

Plusieurs types d'enzymes inactivatrices induisent une résistance isolée selon la structure de l'antibiotique incriminé. Chez *E. faecium*, les enzymes retrouvées sont principalement (Roberts, 2008; Reissier et Cattoir, 2020) :

- des nucléotidyl-transférases codées par *lnu(B)* induisant une résistance aux lincosamides,
- des acétyl-transférases codées par *vat(D)* ou *vat(E)* induisant une résistance aux streptogramines A,
- des lactonases codées par *vgb(A)* ou *vgb(B)* induisant une résistance aux streptogramines B.



Enfin, on peut retrouver des systèmes d'efflux de type MFS qui sont principalement codés par les gènes de type *mef* largement retrouvés chez les streptocoques mais rares chez *E. faecium* (Leclercq, 2002). Chez *E. faecium*, il existe des analogues aux ABC transporteurs codés principalement par les gènes *msr*, *vga* ou *lsa*, dont le mécanisme de résistance est la protection ribosomale (Portillo *et al.*, 2000; Reynolds *et al.*, 2003; Novotna et Janata, 2006).

### c. Tétracyclines

Les tétracyclines sont une famille d'antibiotiques bactériostatiques caractérisés par un noyau naphtacène. La première molécule de cette famille (auréomycine) extrait de souches de *Streptomyces* avait été décrite comme le premier antibiotique à large spectre en 1948 (Nelson et Levy, 2011). Des dérivés hémi-synthétiques comme la doxycycline ou la minocycline ont été obtenus pour améliorer leurs caractéristiques pharmacocinétiques/pharmacodynamiques (PK/PD) et diminuer les effets secondaires de cette classe d'antibiotiques (Nguyen *et al.*, 2014). Une 3<sup>ème</sup> génération de tétracyclines a été développée pour surmonter les résistances aux tétracyclines acquises par les bactéries au fil du temps avec comme premier représentant une glycylcycline : la tigécycline (Bhattacharya *et al.*, 2009). Cette génération a une conformation différente des molécules des 2 premières générations apportant une affinité supérieure sur les zones d'interaction habituelles des tétracyclines. Cette spécificité restreint la flexibilité au niveau de l'ARNr 16S nécessaire pour le bon fonctionnement des protéines de protection ribosomale (RPP, Ribosomal protection protein) (Bauer *et al.*, 2004; Schedlbauer *et al.*, 2015).

Les tétracyclines présentent des caractéristiques pharmacocinétiques intéressantes avec une absorption généralement bonne (surtout pour les molécules de 2<sup>ème</sup> génération), une distribution tissulaire importante et une faible métabolisation (exceptée pour la minocycline et la tigécycline) (Agwuh et MacGowan, 2006). Elles ont également un large

spectre d'activité avec une activité sur les cocci à Gram positif, la plupart des entérobactéries, certains bacilles à Gram négatif non fermentants et les bactéries intracellulaires. Elles ont également une activité sur les protozoaires (Eliopoulos et Roberts, 2003). Ces diverses caractéristiques font de la famille des tétracyclines, une classe d'antibiotiques largement utilisée. Malheureusement, cette large utilisation a favorisé l'émergence de souches résistantes limitant ainsi drastiquement leur utilisation. Aujourd'hui, les tétracyclines de 1<sup>ère</sup> génération sont surtout utilisées en médecine vétérinaire (Daghrir et Drogui, 2013). Celles de seconde génération sont principalement utilisées pour les infections à bactéries intracellulaires comme les rickettsies, les bartonelles, les *Chlamydiae* ou les mycoplasmes (Eliopoulos et Roberts, 2003). Les molécules de 3<sup>ème</sup> génération (tigécycline, éravacycline, omadacycline) sont réservées à certaines infections à bactéries multi-résistantes comme les entérobactéries productrices de carbapénémases (EPC), *S. aureus* résistant à la méticilline (SARM) et les ERG (Bhattacharya *et al.*, 2009; Sutcliffe *et al.*, 2013; Bender *et al.*, 2018; Gallagher, 2019; Boukthir *et al.*, 2020). Dans le cas des ERG, l'activité bactériostatique et la faible concentration plasmatique de ces molécules limitent leur utilisation à certaines situations particulières.

Le mécanisme d'action antibactérien des tétracyclines est l'inhibition de la traduction par fixation à la sous-unité 30S du ribosome bactérien. Plus précisément, les tétracyclines se fixent au niveau du site A de la sous-unité 30S empêchant l'accès du complexe aaARNt/EF-Tu [Figure 7.B] (Brodersen *et al.*, 2000). Les interactions entre l'ARNr 16S et les tétracyclines se situent sur différentes positions dans le site A et ont un niveau variable d'affinité. Elles nécessitent souvent des ions  $Mg^{2+}$  en tant que co-facteur (Brodersen *et al.*, 2000; Pioletti *et al.*, 2001; Anokhina *et al.*, 2004). Les tétracyclines sembleraient pouvoir interagir avec des séquences aléatoires d'ARN double brin, pouvant donner des pistes de mécanisme d'action expliquant les autres propriétés thérapeutiques (antiprotozoaire,

antivirale et anti-inflammatoire) des tétracyclines (Wei et Bechhofer, 2002; Chukwudi et Good, 2016).

Comme d'autres bactéries, *E. faecium* a acquis des résistances aux tétracyclines principalement médiées par des protéines Tet impliquant des mécanismes d'efflux et/ou de protection ribosomale (Chopra et Roberts, 2001; Thaker *et al.*, 2010; Grossman, 2016). Les gènes le plus souvent retrouvés chez *E. faecium* sont *tet(M)* et *tet(L)*, gènes pouvant être parfois co-localisés et co-transcrits (Aarestrup *et al.*, 2000; Nishimoto *et al.*, 2005; Hummel *et al.*, 2007; Fiedler *et al.*, 2016; Boukthir *et al.*, 2020).

Les protéines d'efflux impliquées dans la résistance aux tétracyclines appartiennent à la famille des MFS, grande famille de pompes à efflux actif possédant 12 ou 14 domaines transmembranaires (TMS) (Paulsen *et al.*, 1996; Ranaweera *et al.*, 2015). Les gènes *tet* d'efflux peuvent être séparés en 7 groupes basés sur l'analogie de leurs séquences peptidiques (Guillaume *et al.*, 2004). Chez *E. faecium*, les gènes d'efflux appartiennent au groupe 2 incluant *tet(L)* ou *tet(K)*. Ces gènes sont principalement retrouvés chez les bactéries à Gram positif et souvent portés par des plasmides (Chopra et Roberts, 2001; Cattoir, 2004). Ils codent des MFS à 14 TMS permettant la diminution intracellulaire des tétracyclines de 1<sup>ère</sup> et 2<sup>ème</sup> génération (excepté minocycline) en échange d'ions H<sup>+</sup> (Thaker *et al.*, 2010).

L'autre principal type de résistance aux tétracyclines est causé par des protéines de protection ribosomale (RPP) codées chez *E. faecium* par les gènes *tet(M)*, *tet(O)*, *tet(S)* ou *tet(T)* (Thaker *et al.*, 2010; Boukthir *et al.*, 2020). La protéine la plus retrouvée est Tet(M) (Burdett, 1991). Les RPP ont une structure secondaire proche du complexe aaARNt/EF-Tu (Dantley *et al.*, 1998; Dönhöfer *et al.*, 2012). Cette protéine de type Tet(M) peut s'insérer au niveau du site A de la sous-unité 30S du ribosome malgré la présence de la tétracycline, empêchant la fixation du complexe aaARNt/EF-Tu [Figure 8] (Dönhöfer *et al.*, 2012). La

protéine provoque un changement conformationnel GTP-dépendant au niveau des positions A<sub>1492</sub>, A<sub>1493</sub> et C<sub>1054</sub> induisant l'expulsion de la tétracycline (Burdett, 1996; Dantley *et al.*, 1998) [Figure 8.B]. La nouvelle conformation en C<sub>1054</sub> protège alors le site d'une fixation ultérieure par la tétracycline. Le complexe aaARNt/EF-Tu peut donc s'insérer pour continuer la traduction [Figure 8.C].

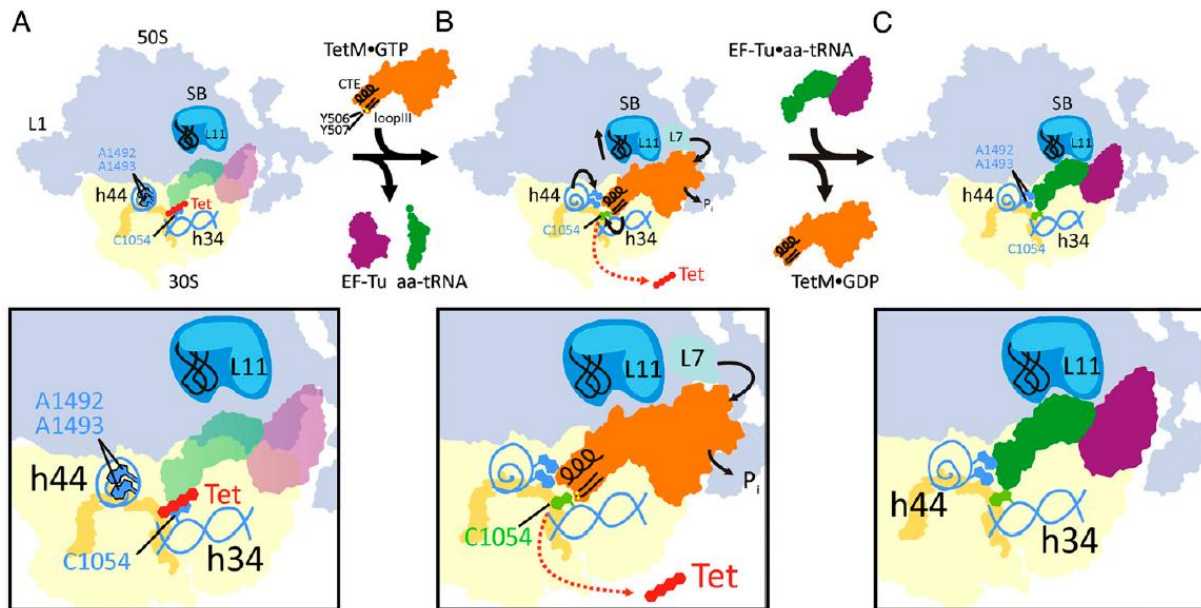


Figure 8. Description du mécanisme de résistance aux tétracyclines par la protéine Tet(M) (Dönhöfer *et al.*, 2012).

(A) Interaction de la tétracycline (rouge) empêchant l'insertion du complexe aaARNt/EF-Tu (vert et violet) au niveau de la sous-unité 30S du ribosome (jaune).

(B) Insertion de la protéine Tet(M) (orange) provoquant un changement conformationnel, expulsant la tétracycline.

(C) Insertion du complexe aaARNt/EF-Tu au niveau de la sous-unité 30S du ribosome.

Comme pour les systèmes d'efflux, les tétracyclines de 3<sup>ème</sup> génération échappent au mécanisme médié par les RPP, sauf si celles-ci sont surexprimées (Beabout *et al.*, 2015b; Fiedler *et al.*, 2016). Certaines souches ont acquis une résistance à la tigécycline par mutation/délétion au sein du gène *rpsJ* codant pour la protéine ribosomale S10 (Beabout *et al.*, 2015a; Cattoir *et al.*, 2015). Cette protéine ribosomale est localisée au niveau de l'ARNr 16S, proche du site de fixation de la tigécycline (Luo *et al.*, 2008).

#### d. Linézolide

Le linézolide est le premier et principal représentant de la classe des oxazolidinones (Hutchinson, 2003). Ces antibiotiques sont des inhibiteurs de la traduction par fixation à la sous-unité 50S du ribosome au niveau du CPT (Lin *et al.*, 1997). Plus précisément, l'antibiotique interagit avec l'ARNr 23S empêchant l'aaARNt de se positionner au niveau du site A [Figure 7.A] (Wilson *et al.*, 2008; Mendes *et al.*, 2014).

Les mécanismes de résistance étaient principalement liés à des mutations de la cible (ARNr 23S ou protéines ribosomales) ou une méthylation de la cible par la méthyltransférase Cfr (Marshall *et al.*, 2002; Diaz *et al.*, 2012). Ces deux mécanismes de résistance sont retrouvés chez *E. faecium*, mais le second est plus problématique car porté par un gène transférable (Diaz *et al.*, 2012; Lazaris *et al.*, 2017). De plus, l'acquisition du gène *cfr* confère une résistance croisée à d'autres antibiotiques comme les phénicolés ou les lincosamides (phénotype PhLOPS<sub>A</sub>) (Long *et al.*, 2006). Plus récemment, deux autres mécanismes de résistance aux oxazolidinones ont été mis en évidence, étant dus à des protéines ABC codées par les gènes *optrA* et *poxA*. Ces protéines agissent probablement par protection ribosomale (comme Msr, Vga et Lsa) et confèrent une résistance croisée avec les phénicolés (Wang *et al.*, 2015; Antonelli *et al.*, 2018). Ces gènes sont également transférables car généralement portés par des plasmides (Li *et al.*, 2016; Lazaris *et al.*, 2017; Sassi *et al.*, 2019).

## 5. Daptomycine

La daptomycine est un lipopeptide cyclique ayant la particularité d'être bactéricide envers les ERG (Rybak *et al.*, 2000). L'activité de la daptomycine est  $\text{Ca}^{2+}$ -dépendant et restreinte aux bactéries à Gram positif (Ho *et al.*, 2008; Taylor et Palmer, 2016). La structure de cette molécule est composée d'un peptide cyclique de 10 AA avec une chaîne de 3 AA contenant un acide décanoïque (Debono *et al.*, 1988). Le mécanisme d'action repose sur une interaction entre la chaîne carbonée et les phospholipides de la membrane bactérienne. L'ion divalent capté au niveau du peptide cyclique va provoquer une dépolarisation de la membrane favorisant une oligomérisation des molécules de daptomycine. L'oligomérisation aura pour conséquence une déstabilisation de la membrane et la formation de pores induisant une perte ionique de la cellule bactérienne et la lyse cellulaire (Heidary *et al.*, 2018).

Les acquisitions de résistance à la daptomycine restent encore sporadiques (Bender *et al.*, 2018). Chez *E. faecium*, une diminution de la sensibilité à la daptomycine peut être observée par une modification des charges et de la composition en phospholipides de la membrane bactérienne. Ces modifications sont généralement dues à des mutations multiples au niveau des gènes *liaFSR*, *cls*, *gdpD*, *yycFG* ou *cfa* (Mishra *et al.*, 2012; Diaz *et al.*, 2014; Sinel *et al.*, 2016).

## E. Plasticité génomique et supports génétiques de l'antibiorésistance chez *E. faecium*

*E. faecium* montre une grande plasticité génomique pouvant expliquer sa capacité d'adaptation. Le « core-génome », c'est-à-dire l'ensemble des gènes communs à toutes les souches d'espèce bactérienne, représente moins de 40% du nombre moyen de gènes retrouvé dans une souche (Zhong *et al.*, 2019). Cela signifie que la majorité des gènes retrouvés chez *E. faecium* a été acquise. Cette capacité de transfert horizontal de gènes est un atout majeur en termes d'adaptation et d'évolution (Boto, 2010). Cette plasticité est favorisée par un manque de CRISPR (Clustered regulatory interspaced short palindromic repeats) chez *E. faecium*. En effet, les CRISPR sont impliqués dans la protection du génome contre les insertions d'éléments génétiques étrangers comme les phages (Horvath et Barrangou, 2010). Chez *E. faecium*, l'acquisition de gènes de résistance semble corrélée à une absence de CRISPR (Palmer et Gilmore, 2010). En effet, une défaillance de ce système favorise les échanges génétiques horizontaux.

### 1. Eléments génétiques mobiles

Les éléments génétiques mobiles (EGM) principaux impliqués dans le transfert génétique horizontal sont les plasmides et les transposons. Les plasmides sont extra-chromosomiques et contiennent divers gènes pouvant être impliqués dans la résistance, la virulence ou la conjugaison. Les plasmides dit conjuguatifs possèdent les gènes de conjugaison nécessaires pour leur transfert horizontal et font donc partie des ICE (Integrative and conjugative elements) (Johnson et Grossman, 2015). Le transposon est un élément génétique chromosomique ou plasmidique comportant généralement plusieurs gènes, pouvant se transposer (couper ou copier) et s'intégrer au niveau d'autres positions génétiques. Ils sont séparés en 3 groupes [Figure 9] : composite, non-composite et conjugatif (Hegstad *et al.*, 2010). Le transposon composite est entouré de deux séquences d'insertion. Une séquence d'insertion (IS) est un petit élément transposable composé d'une transposase encadrée de

deux séquences inversées répétées (IR). Le transposon non-composite, dit de type Tn3, est entouré de deux IR et contient une transposase. Si le transposon contient une résolvase, la transposition sera alors répllicative (transposition d'une copie du transposon). Le transposon conjugatif contient des gènes de conjugaison (recombinase...) et est donc un ICE (Ciric *et al.*, 2013).

Trois mécanismes sont impliqués dans le transfert génétique horizontal chez les bactéries : la conjugaison, la transduction et la transformation.

Chez *E. faecium*, le mécanisme de transfert génétique horizontal le plus fréquent est la conjugaison. Le second utilisé est la transduction phagique (García-Solache et Rice, 2019). Bien que largement utilisé *in vitro*, les phénomènes de transformation bactérienne n'ont pas encore pu être observés naturellement chez *E. faecium* (Clewell *et al.*, 2014).

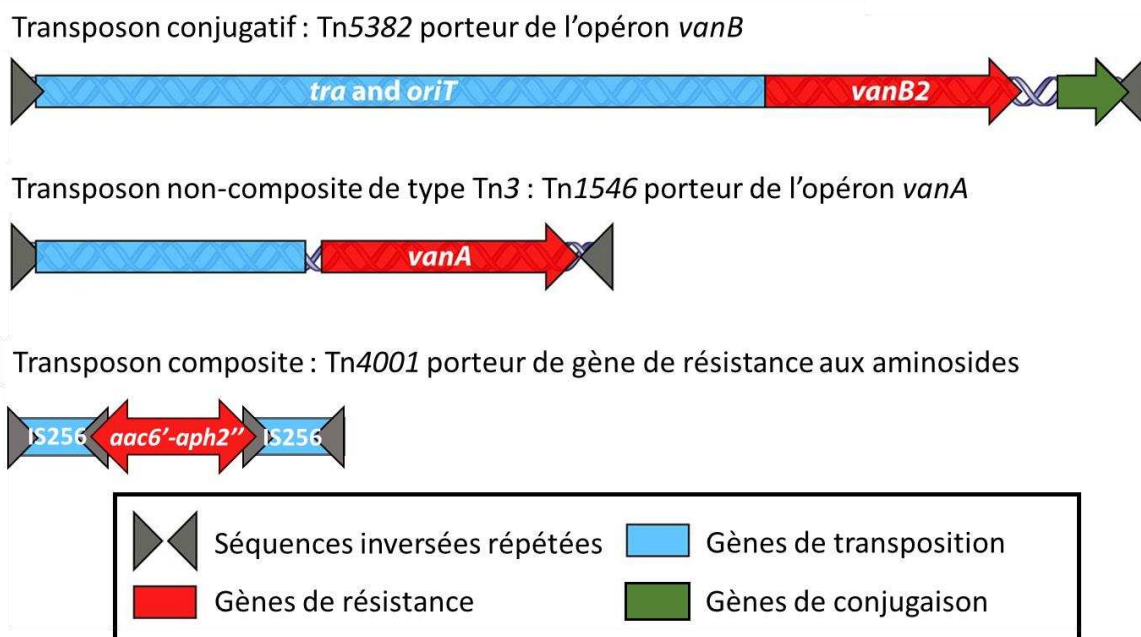


Figure 9. Présentation des 3 types de transposons (García-Solache et Rice, 2019).



On retrouve trois types de plasmides chez les entérocoques (Weaver *et al.*, 2002) :

- RCR plasmides (Rolling circle replicating ou réplication en cercle roulant) (Ruiz-Masó *et al.*, 2015) : plasmides pouvant être retrouvés chez toutes les espèces bactériennes mais majoritairement chez les bactéries à Gram positif. La réplication du brin principal démarre à partir de l'origine de réplication double-brin et l'élongation s'effectuera entièrement sur ce brin. Quand le premier brin est entièrement synthétisé, la réplication du second brin s'effectue de même en démarrant au niveau de l'origine de réplication simple brin.
- PRP (Pheromone-responsive plasmids) (Wardal *et al.*, 2010) : plasmides exclusivement retrouvés chez les entérocoques. La bactérie réceptrice va produire des petits peptides (phéromones) pouvant être reconnus par des protéines codées par le PRP contenu dans la bactérie donatrice. Cela induit alors la production de facteur d'agrégation fixant les substances de fixation entre entérocoques favorisant le contact entre les 2 cellules. Lors de ce processus, le plasmide est transféré de la cellule donatrice à la cellule réceptrice par un système de sécrétion de type IV. Après le transfert, des peptides inhibiteurs sont produits par le plasmide transféré induisant une baisse de la production de phéromone et terminant le mécanisme de transfert génétique horizontal. Ce type de plasmide se retrouve principalement chez *E. faecalis* (Rosvoll *et al.*, 2010).
- Inc-18 plasmides (Groupe d'incompatibilité 18) (Kohler *et al.*, 2018) : plasmides seulement retrouvés chez certains *Firmicutes* comme les entérocoques, streptocoques et staphylocoques. Les plasmides d'un même groupe d'incompatibilité ne peuvent pas cohabiter dans une même cellule. Ce type de plasmide est le plus courant chez *E. faecium* (Rosvoll *et al.*, 2010).

Ces plasmides peuvent contenir un système toxine-antitoxine (T-AT) permettant la stabilisation et le maintien du plasmide (Van Melderen et Saavedra De Bast, 2009). Le

plasmide exprime une toxine étant inhibée par une antitoxine (ARN ou protéine). En cas de perte du plasmide, l'antitoxine sera plus vite dégradée que la toxine, permettant à celle-ci d'inhiber voire de tuer la cellule bactérienne. Dans ce cas, seule la population bactérienne gardant le plasmide avec le système T-AT peut subsister. Chez *E. faecium*, le premier système T-AT décrit a été Axe-Txe (Antitoxine protéine entérococcique - toxine protéique entérococcique) porté par le plasmide pRUM (Grady et Hayes, 2003). Il existe d'autres systèmes T-AT pouvant être retrouvés chez les entérocoques, comme le système  $\omega$ - $\epsilon$ - $\zeta$  dans les plasmides pRE25 ou le *locus par* codant les toxines Fst dans les plasmides pAD1 (Schwarz *et al.*, 2001; Zielenkiewicz et Cegłowski, 2005; Weaver *et al.*, 2009; Rosvoll *et al.*, 2010).

## 2. Implications dans la résistance et les épidémies

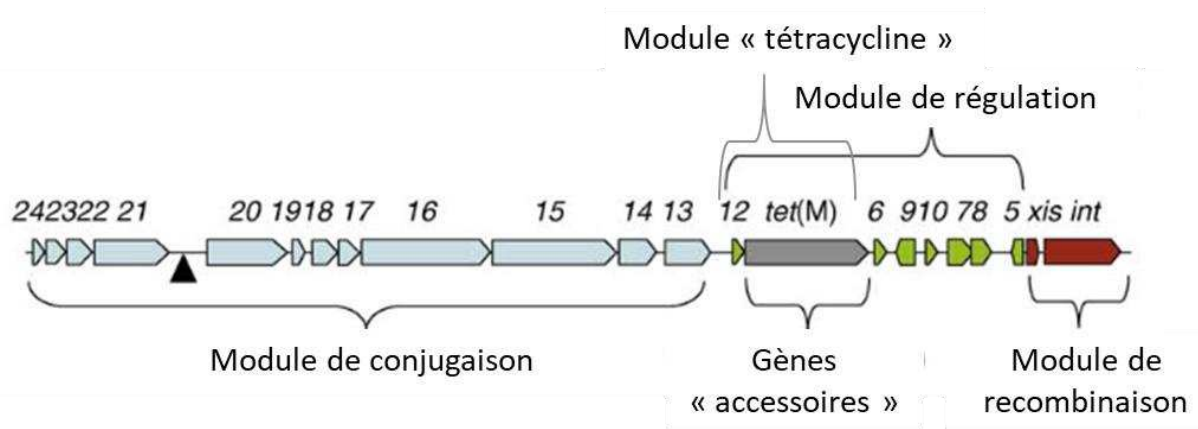
Certains éléments génétiques transférables mobiles comme les plasmides et les transposons ont largement diffusé grâce à l'avantage sélectif qu'ils peuvent apporter par le biais de gènes de résistance ou de facteurs de virulence (Wozniak et Waldor, 2010). Ces outils sont particulièrement importants chez les entérocoques, leur permettant depuis quelques années de devenir des pathogènes de plus en plus prévalents (Hegstad *et al.*, 2010; Palmer *et al.*, 2010). Dans une étude analysant un total de 94 souches d'*E. faecium*, les 2 plasmides les plus fréquemment retrouvés sont pRE25 et pRUM (Rosvoll *et al.*, 2010). Le plasmide pRE25 est présent dans des souches d'origine diverse (animaux, humains communautaires et hospitaliers). Au contraire, le plasmide pRUM est plus fréquent dans les souches hospitalières, notamment les souches responsables d'infections. Le plasmide pRUM d'origine est composé du système T-AT Axe-Txe, de gènes de résistance aux UV (*uvrA*) et aux antibiotiques (aminosides, macrolides et chloramphénicol) (Grady et Hayes, 2003). L'organisation des gènes de résistance aux aminosides *aadE-sat4-aphA-3* indique la présence du transposon Tn5405 tronqué dans ce plasmide pRUM (Derbise *et al.*, 1996; Werner *et al.*, 2001). Les plasmides dérivés de pRUM (pRUM-like) sont des plasmides

mosaïques ayant intégré divers éléments mobiles, notamment le transposon Tn1546 porteur de l'opéron *vanA* conférant la résistance aux glycopeptides (Clewell *et al.*, 2014).

Chez *E. faecium*, on s'intéressera principalement à la famille des transposons conjugatifs de type Tn916 et au transposon non-composite Tn1546 porteur de l'opéron *vanA* (Clewell *et al.*, 2014). Le premier transposon conjugatif décrit est Tn916, et la majorité des transposons conjugatifs dérive de Tn916 et est regroupée dans la famille dite des transposons Tn916-like (Rice, 1998; Roberts et Mullany, 2011; Ciric *et al.*, 2013). Celui-ci a été décrit initialement chez *E. faecalis* (Franke et Clewell, 1981). Ce transposon a été mis en évidence avec l'observation d'un transfert de résistance à la tétracycline en absence de plasmide détectable. Les auteurs ont suggéré qu'un EGM, désigné transposon Tn916, pouvait se transférer entre deux cellules par conjugaison indépendamment d'un plasmide conjugatif. Le transposon Tn916 est composé de 18 kb contenant 38,8 % (G+C) et 24 cadres ouverts de lecture (ORF) (Flannagan *et al.*, 1994). Les ORF peuvent être séparés en plusieurs modules :

- le module « tétracycline » contenant le gène de résistance aux tétracyclines *tet(M)* (ORF11) précédé de l'ORF12, peptide leader hypothétiquement impliqué dans la régulation de ce gène (Burdett, 1990; Su *et al.*, 1992),
- le module de recombinaison avec les gènes impliqués dans le mécanisme de recombinaison avec généralement une excisionase (ORF1 *xis*) et une intégrase (ORF2 *int*) (Su et Clewell, 1993; Jaworski *et al.*, 1996),
- le module de conjugaison impliquant les ORF 13 à 24
- le module de régulation comprenant les gènes de régulation (ORF 6 à 10 et 12) (Celli et Trieu-Cuot, 1998),
- le module de gènes accessoires pouvant contenir divers gènes apportant un avantage sélectif comme les gènes impliqués dans la résistance aux antibiotiques.

Les transposons de la famille des Tn916-like ont impérativement les modules de conjugaison, de recombinaison et de régulation [Figure 10].



**Figure 10.** Disposition schématique des différents modules composant les transposons Tn916-like (Roberts et Mullany, 2009).

La diversité dans cette famille est surtout retrouvée au niveau des gènes « accessoires » pouvant notamment coder des protéines impliquées dans l'antibiorésistance (Ciric *et al.*, 2013). Les transposons conjuguatifs de cette famille sont largement répandus dans le règne bactérien (Roberts et Mullany, 2009). Le module de recombinaison est positionné à l'extrémité finale du transposon permettant l'excision par l'excisionnase codée par le gène *xis* et l'intégration de celui-ci par l'intégrase codée par le gène *int*. Chez certains transposons de la famille Tn916-like, une résolvase codée par le gène *tndX* suffit pour effectuer la recombinaison (Wang et Mullany, 2000; Wang *et al.*, 2000; Giovanetti *et al.*, 2007). Ces transposons ont la capacité de s'insérer sur diverses positions riches en A:T (Trieu-Cuot *et al.*, 1993; Scott *et al.*, 1994; Wang *et al.*, 2006). Entre ces deux étapes, le transposon se retrouve sous une forme circulaire (Scott *et al.*, 1994; Salyers *et al.*, 1995) [Figure 11.A]. C'est lors de cet état circulaire que la conjugaison peut s'effectuer (Salyers *et al.*, 1995; Ciric *et al.*, 2013) [Figure 11.B]. Les transposons de la famille Tn916-like sont largement impliqués dans la diffusion de gènes de résistance aux antibiotiques comme les

tétracyclines mais aussi les macrolides et les aminosides (Roberts et Mullany, 2009; Ciric *et al.*, 2013). En plus d'être porteur de résistance aux antibiotiques, certains antibiotiques et antiseptiques sembleraient augmenter la conjugaison de ces transposons favorisant d'autant plus leur dissémination entre différentes espèces (Seier-Petersen *et al.*, 2014; Beabout *et al.*, 2015b; Scornec *et al.*, 2017).

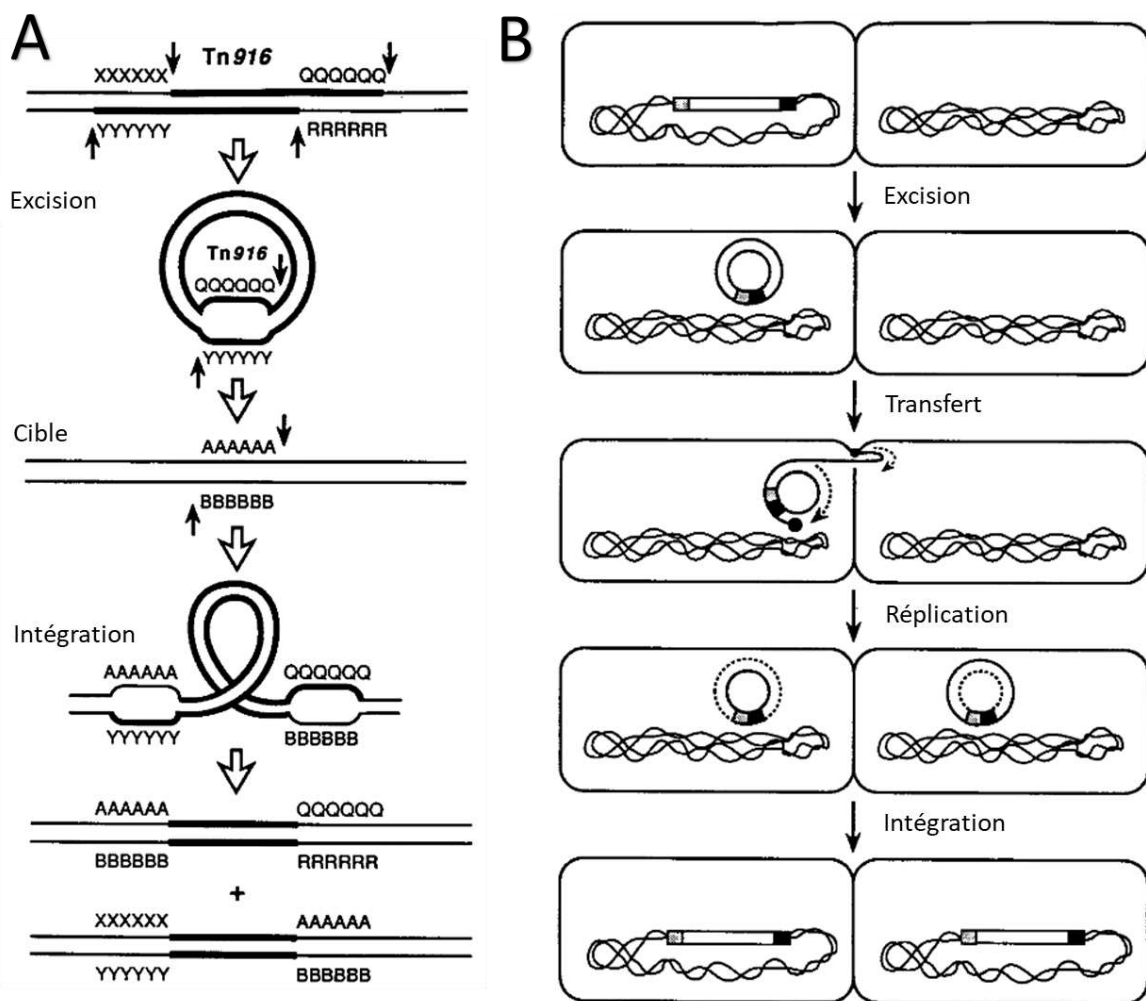


Figure 11. Représentation de la mobilisation des transposons de type Tn916 (Salysers *et al.*, 1995).

(A) Recombinaison du transposon entre 2 *loci*.

(B) Conjugaison du transposon entre 2 cellules.

Le transposon Tn1549 joue également un rôle important chez *E. faecium* car c'est un transposon conjugatif porteur de l'opéron *vanB*. Il est structurellement proche de la famille des Tn916-like avec le remplacement des modules « tétracycline » et régulation par l'opéron *vanB2* [Figure 9] (Garnier *et al.*, 2000).

Les transposons composites et non-composites peuvent également contenir des gènes de résistance aux antibiotiques, mais leur transfert horizontal génétique est médié par un plasmide conjugatif. Chez *E. faecium*, les principaux sont Tn1546 (*vanA*), Tn1547 (*vanB1*), et Tn4001 (aminosides) (Hegstad *et al.*, 2010).

Un autre mécanisme impliqué dans la résistance aux antibiotiques médié par les EGM est l'inactivation de gène par insertion notamment d'IS (Depardieu *et al.*, 2007; Siguier *et al.*, 2014).

Dans la souche d'*E. faecium* Aus0004, le gène de résistance aux tétracyclines *tet(M)* est inactivé par un transposon non-composite de type Tn3 (Lam *et al.*, 2012; León-Sampedro *et al.*, 2016). Ce mécanisme d'inactivation par insertion peut également être considéré comme un mécanisme de régulation. En effet, des souches résistantes d'*E. faecium* ont été rendues résistantes à la daptomycine en induisant des mutations au niveau de l'opéron *liaFSR*. En absence de pression de sélection, une amélioration du « fitness bactérien » peut être induite par inactivation de *liaFSR* due à l'insertion d'une IS. En cas de nouvelle exposition, une réversion de la résistance est rapidement possible par excision naturelle de l'IS (Sinéel *et al.*, 2016).

## II. Systèmes de régulation bactériens dépendant de la conformation de l'ARN

L'étude des mécanismes de régulation génétique est indispensable dans la compréhension de l'adaptation bactérienne. Des systèmes de régulation existent à chaque étape entre la séquence nucléotidique et l'effet produit par la protéine (Joanne M. Willey *et al.*, 2013). En effet, la régulation peut s'effectuer lors de l'initiation de la transcription, lors l'élongation transcriptionnelle (aussi appelé régulation post-transcriptionnelle), à l'étape de la traduction et sur la protéine produite (aussi appelé régulation post-traductionnelle). Ces mécanismes de régulation sont largement répandus et peuvent être impliqués dans l'antibiorésistance (Dersch *et al.*, 2017).

Pour la suite de ce manuscrit, nous nous intéresserons principalement aux mécanismes dits de riborégulation. La riborégulation implique généralement une régulation post-transcriptionnelle ou traductionnelle directement causée par la conformation de l'ARN en cours de synthèse. La conformation prise par l'ARN peut avoir un impact direct sur une terminaison prématurée de la transcription par formation d'une tige-boucle terminatrice (régulation post-transcriptionnelle), le recrutement/protection d'une RNase ou une séquestration/libération de la séquence SD (régulation traductionnelle). Plusieurs facteurs peuvent influencer sur la conformation de cet ARN : divers métabolites (riboswitch), le ribosome (mécanisme d'atténuation) et les ARN régulateurs (ARNreg).

## A. ARN régulateur « conventionnel »

Selon les auteurs, la définition des ARN régulateurs peut varier, nous définirons dans ce manuscrit un ARN régulateur « conventionnel » comme un ARN généralement non codant d'une taille variable de 50 à 500 nucléotides et pouvant interagir avec d'autres ARN voire des protéines, provoquant ainsi une régulation dépendant de la zone d'interaction. Ils peuvent donc avoir un rôle dans la régulation post-transcriptionnelle, traductionnelle voire post-traductionnelle (Waters et Storz, 2009). La régulation post-traductionnelle est due à l'interaction entre l'ARN et une protéine (Babitzke et Romeo, 2007). On dit d'un ARN régulateur qu'il a une action en *cis* si celui-ci est codé sur le brin opposé de sa cible. L'interaction sera forte grâce à une homologie de séquence parfaite entre l'ARNreg et l'ARN cible (Brantl, 2007). Ces ARN sont également appelés ARN antisens et sont impliqués dans les systèmes T-AT de type I (Gerdes et Wagner, 2007). Un ARNreg peut également avoir une activité dite en *trans* si celui-ci peut interagir avec un ARN situé sur un *locus* différent de celui de l'ARNreg. Généralement, un ARNreg peut avoir de multiples cibles en *trans*, mais l'interaction sera généralement modérée due à une homologie de séquence imparfaite. L'interaction entre l'ARNreg et l'ARNm cible peut alors provoquer différents effets [Figure 12] :

- Interaction sur une séquence importante
  - Au niveau du RBS : empêche le ribosome d'initier la traduction (régulation traductionnelle)
  - Au niveau d'un site de restriction : empêche une RNase de dégrader l'ARN cible (régulation post-transcriptionnelle)
- Modification de la structure secondaire de l'ARN cible
  - Conformation permettant de libérer le RBS, permettant l'initiation de traduction (régulation traductionnelle)
  - Conformation induisant ou empêchant la formation d'une tige-boucle terminatrice (régulation post-transcriptionnelle)



- Recrutement d'une RNase : provoque la dégradation de l'ARN cible (régulation post-transcriptionnelle).

Le rôle phénotypique de l'ARN régulateur dépendra de la fonction de sa cible. Ces ARN régulateurs peuvent donc jouer un rôle dans la virulence (Caldelari *et al.*, 2013; Pitman et Cho, 2015; Quereda et Cossart, 2017), l'antibiorésistance (Lalaouna *et al.*, 2014; Felden et Cattoir, 2018), la régulation métabolique (Michaux *et al.*, 2014b), le maintien d'EGM par les système T-AT de type I (Gerdes et Wagner, 2007) ou la formation de biofilm (Mika et Hengge, 2013).

Plusieurs études ont mis en évidence l'existence d'ARN régulateurs putatifs chez les entérocoques, mais peu d'entre eux ont une fonction ou un rôle défini (Fouquier d'Hérouël *et al.* 2011; Shioya *et al.* 2011; Innocenti *et al.* 2015; Sinel *et al.* 2017; Michaux *et al.* 2020). On nommera dans ce manuscrit par défaut « petit ARN » ou ARN non codant, un ARN exprimé sans rôle défini. En effet, il s'avère qu'un petit ARN puisse être finalement une région régulatrice en 5'UTR plutôt qu'un ARN régulateur « conventionnel ».

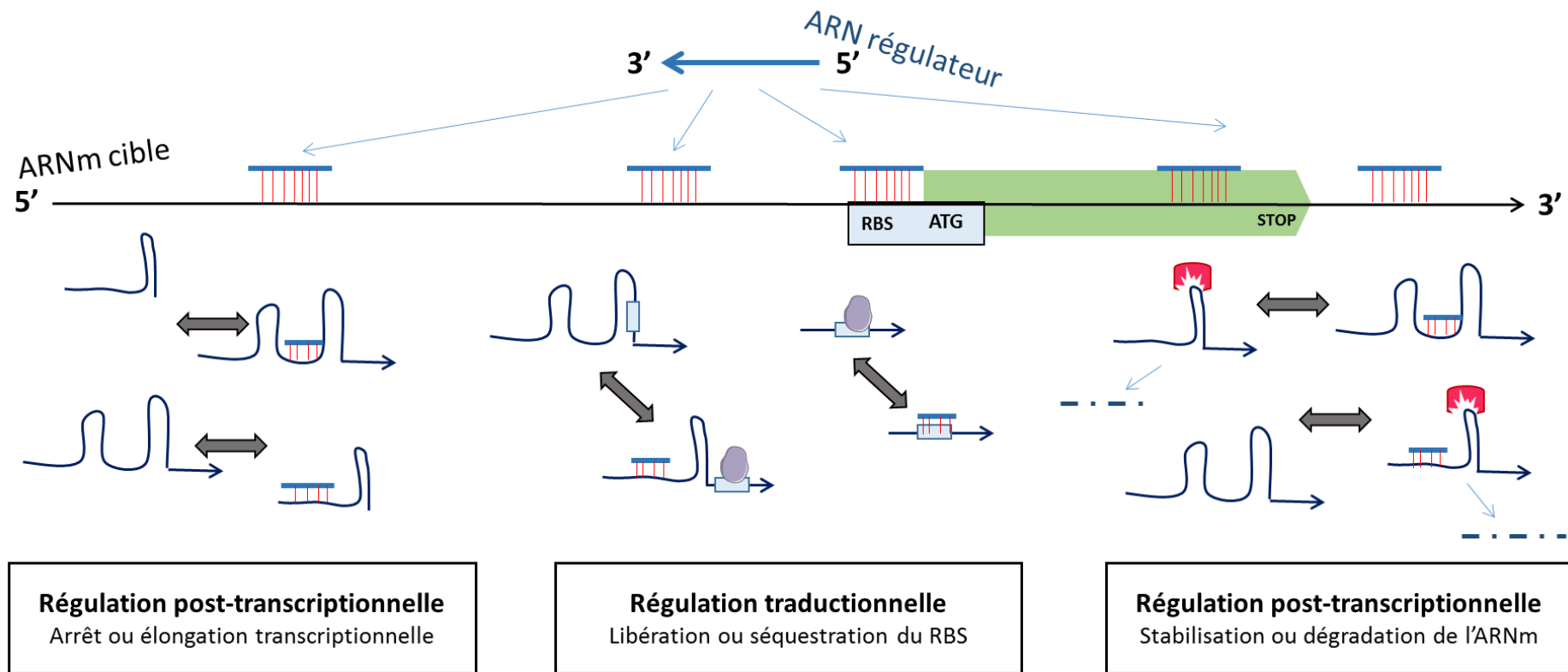
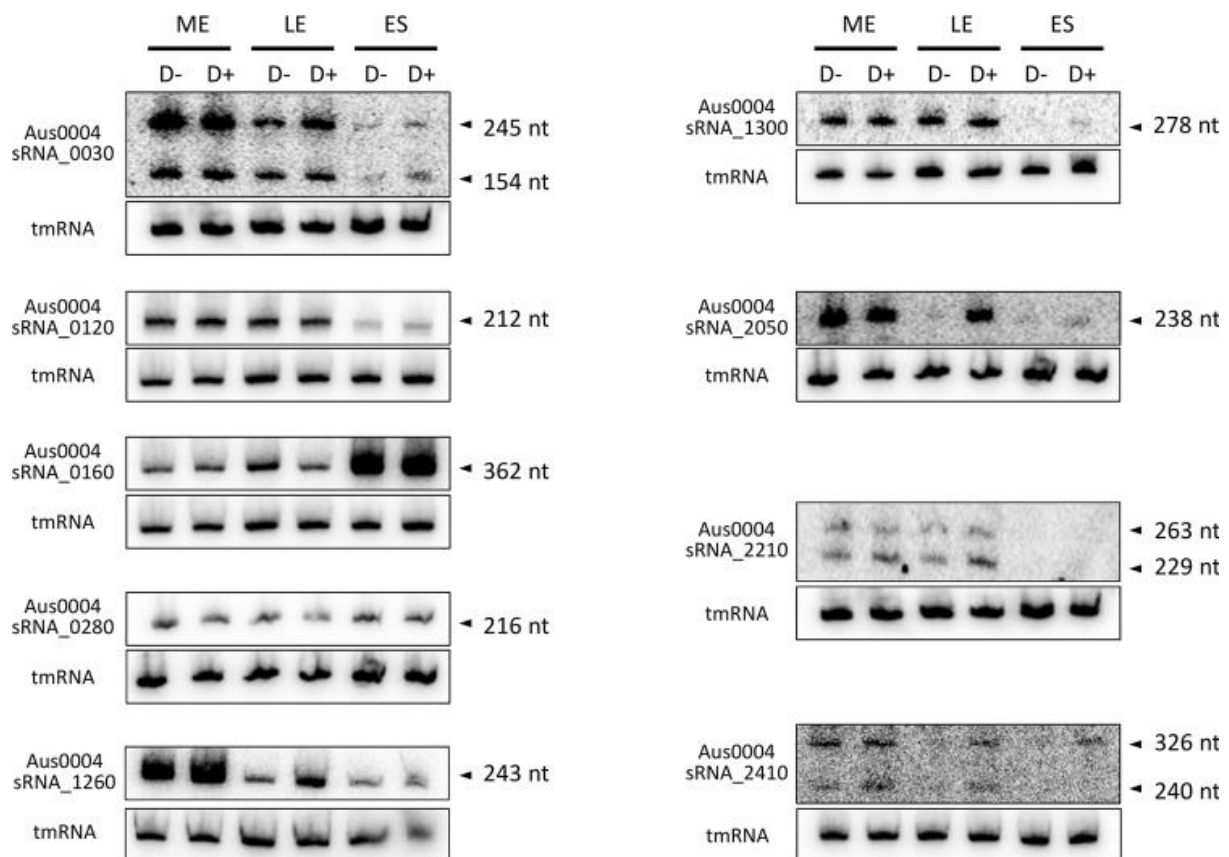


Figure 12. Représentation des mécanismes de régulation induits par l'interaction entre un ARNreg et un ARNm.

L'ARNreg est en bleu et l'ARNm est en noir. La séquence codante est indiquée en vert. Le RBS de la protéine régulée est représenté par un rectangle bleu. Le ribosome est représenté par la protéine violette. La présence de la flèche indique l'élongation transcriptionnelle de l'ARNm. La protéine rose est une RNase. Les pointillés indiquent un ARN dégradé.

Chez *E. faecalis*, initialement, une trentaine d'ARN non codants a pu être mise en évidence par 5'tagRACE dont 16 antisens et 13 validés par Northern-Blot (Fouquier d'Hérouël *et al.*, 2011). Une dizaine d'autres petits ARN a été révélée par tilling microarray et validée par Northern Blot (Shioya *et al.*, 2011). On y retrouve des homologues au tmRNA, RNA I et 4.5S (Greenfield *et al.*, 2001; Corvaisier *et al.*, 2003; Treviño *et al.*, 2010). Une fonction dans la virulence de certains de ces ARN a été déterminée par un modèle d'infection *in vivo* chez *Galleria mellonella*, mais le réseau de régulation et les cibles de ces petits ARN restent à identifier (Michaux *et al.*, 2014a). Enfin, une analyse globale par transcriptomique combinée au 5'tagRACE a été effectuée ajoutant 85 nouveaux petits ARN putatifs chez *E. faecalis* (Innocenti *et al.*, 2015).

Plus récemment, les premiers petits ARN ont été mis en évidence chez l'espèce *E. faecium*, notamment en lien avec l'antibiorésistance, problématique principale de cette espèce (Siné *et al.* 2017). Une soixantaine d'ARN putatifs a été identifiée par une approche de transcriptomique globale comparative en soumettant la bactérie à un stress antibiotique, une concentration subinhibitrice de daptomycine. Sur l'ensemble de ces candidats, les 10 ARN les plus exprimés par RNAseq ont été validés par Northern-Blot et RT-qPCR [Figure 13]. Ces données ont permis d'effectuer une première caractérisation de ces ARN avec leurs tailles, leurs profils d'expression selon la phase de croissance ou le stress antibiotique et la présence ou non de formes alternatives.



**Figure 13.** Northern blot validant 10 ARN candidats exprimés par *E. faecium* (Sinel *et al.* 2017). Les ARN totaux ont été extraits à trois phases de croissance (milieu de phase exponentielle ME, fin de phase exponentielle (LE) et début de phase stationnaire (ES) et en présence (D+) ou non (D-) de concentration subinhibitrice de daptomycine.

Les trois nouveaux petits ARN les plus exprimés par RNAseq sont sRNA\_0160, sRNA\_0280 et sRNA\_1300. Seuls 6 nouveaux ARN ont une expression altérée par la présence de daptomycine, avec 3 diminutions d'expression (sRNA\_0560, sRNA\_1420 et sRNA\_2290) et 3 augmentations d'expression (sRNA\_0160, sRNA\_0170, sRNA\_0290). La conservation de tous ces nouveaux petits ARN mis en évidence par RNAseq au sein des différents clades d'*E. faecium* a été estimée. Treize de ces ARN sont totalement absents du clade B communautaire. Sur les 10 ARN validés par Northern-Blot, la plupart semble appartenir au core-génome. Seuls sRNA\_0030 et sRNA\_1300 sont retrouvés de manière inconstante selon le clade. sRNA\_0030 reste principalement retrouvé dans les clades A1 (52%) et A2 (36%). sRNA\_1300 est retrouvé dans moins de 2 % des souches d'*E. faecium*. On peut supposer que

certaines de ces nouveaux ARN décrits chez *E. faecium* sont impliqués dans l'adaptation aux antibiotiques voir l'antibiorésistance (modulation de l'expression par la daptomycine) et dans la pathogénicité et le risque épidémique (absence dans le clade B communautaire). Mais le rôle fonctionnel de ces ARN reste à démontrer. Cette exploration sera décrite au sein de ce manuscrit pour certains de ces nouveaux ARN.

Dernièrement, de nouveaux ARN ont pu être mis en évidence par une méthode de RNAseq différentielle couplée à une prédiction de site d'initiation de la transcription dans les deux espèces *E. faecalis* et *E. faecium* (Michaux *et al.*, 2020). Ces ARN ont été catégorisés en fonction de leur localisation vis-à-vis d'autres gènes, identifiant ainsi des régulateurs de type riboswitch en 5'UTR, des ARN régulateurs antisens ou en région intergénique et d'autres types de régulateurs comme la région régulatrice médiée par la protéine de régulation PyrR. Certains de ces ARN ont été validés par Northern Blot et soumis à diverses conditions de culture et de stress (14 chez *E. faecalis* et 19 chez *E. faecium*).

## B. Régulation en 5'UTR

La région non traduite en 5' (5'UTR) d'un gène codant est la région-clé impliquée dans un mécanisme de *cis*-régulation. En effet, la transcription du reste de l'ARNm et la traduction de la séquence codante est totalement dépendante de la conformation de cette région 5'UTR. Généralement, un gène ayant un 5'UTR de plus de 100 nucléotides sera impliqué dans un mécanisme de *cis*-régulation. On peut distinguer deux grandes catégories de mécanismes de *cis*-régulation : les mécanismes dit d'atténuation et les riboswitchs.

### 1. Atténuation transcriptionnelle et traductionnelle

Selon les auteurs, les mécanismes d'atténuation peuvent comprendre tous les mécanismes induisant un arrêt transcriptionnel prématuré impliquant métabolite (riboswitch), protéine (protéine régulatrice ou ribosome) ou ARN (ARN régulateur ou ARNt) (Lovett et Rogers, 1996; Gollnick et Babitzke, 2002). Dans ce manuscrit, on appellera par défaut « atténuation », le mécanisme d'atténuation médié par le ribosome. Le mécanisme dit d'atténuation implique la présence d'une petite séquence codante (peptide leader) située dans la région 5'UTR du gène régulé. Ce peptide leader n'aura généralement pas de fonction propre et sert seulement au recrutement d'un ribosome. Chez les procaryotes, les étapes de transcription et de traduction sont concomitantes, ce qui signifie que la synthèse peptidique peut influencer l'élongation transcriptionnelle (Miller *et al.*, 1970). La synthèse du peptide leader aura donc une influence sur la conformation de l'ARN en cours de synthèse. L'ARN prendra donc différentes conformations selon la position du ribosome lors de la synthèse du peptide leader. La conformation de l'ARN est alors dépendante de la vitesse et des pauses de ce ribosome impliqué dans la synthèse du peptide leader [Figure 14].

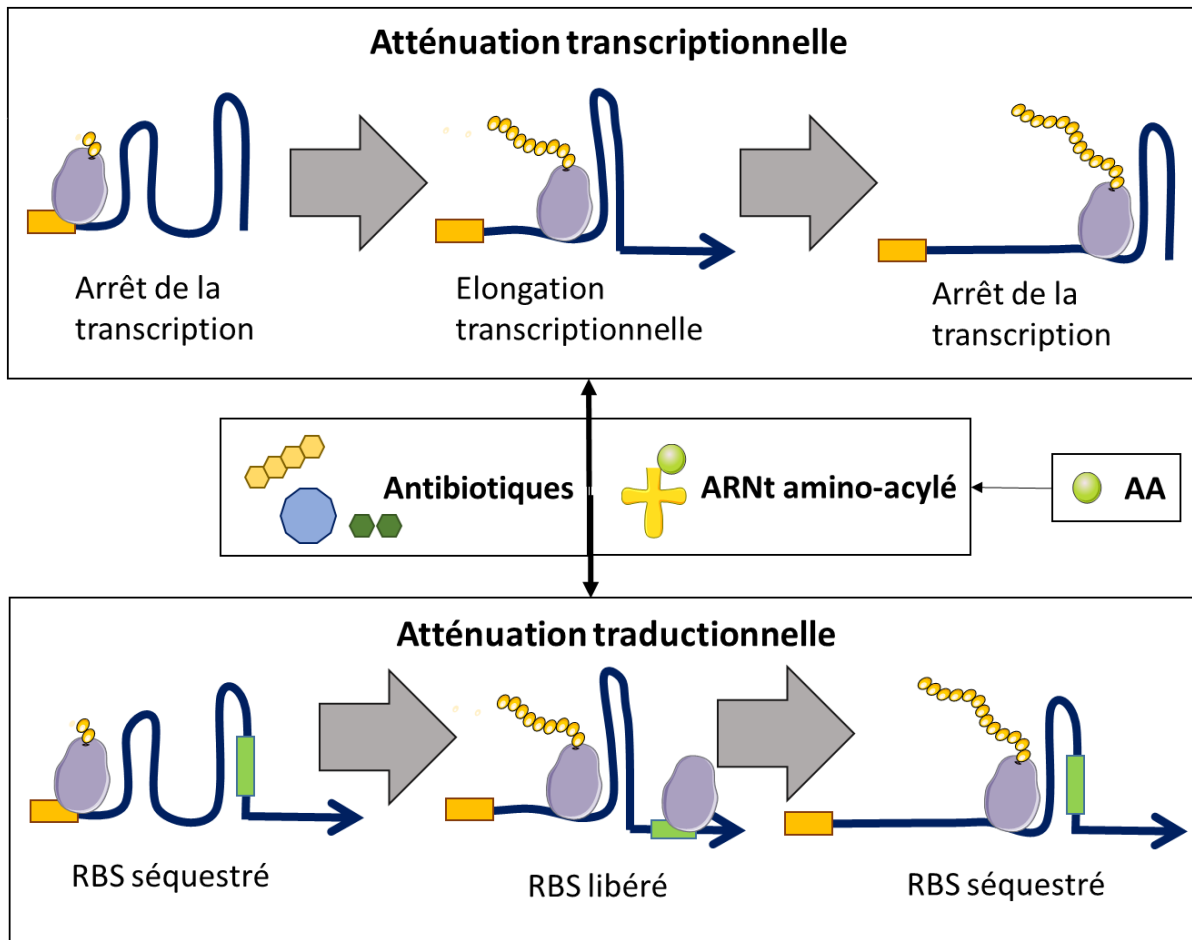


Figure 14. Représentation du principe d'atténuation transcriptionnelle et traductionnelle. Le RBS du peptide leader est représenté par un rectangle jaune, le RBS de la séquence codant la protéine régulée est représenté par un rectangle vert. Le ribosome est représenté par la protéine violette. L'ARN est représenté par la ligne bleue. La présence de la flèche bleue indique l'élongation transcriptionnelle de l'ARN.

Ce mécanisme de régulation par atténuation est donc un mécanisme largement retrouvé dans la régulation des voies métaboliques impliquant l'utilisation d'acides aminés notamment les opérons de biosynthèse d'AA (Kolter et Yanofsky, 1982). En effet, si un des AA impliqués dans la synthèse du peptide leader est en faible quantité, le ribosome sera bloqué à la position du codon correspondant. Au contraire, en excès d'AA, l'avancée du ribosome dans la synthèse du peptide leader sera favorisée. Les mécanismes d'atténuation sont également répandus dans la régulation de gène de résistance aux antibiotiques inhibiteurs de la traduction par fixation du ribosome (Dubnau, 1984; Dersch *et al.*, 2017).

Le mécanisme d'action de ces antibiotiques a indéniablement un impact sur la synthèse du peptide leader.

On distingue deux types d'atténuations : l'atténuation transcriptionnelle (Gollnick et Babitzke, 2002) et l'atténuation traductionnelle (Lovett et Rogers, 1996) [Figure 14]. Pour l'atténuation transcriptionnelle, on retrouve deux types de conformations de l'ARN : une conformation terminatrice et antiterminatrice. La conformation terminatrice implique la présence d'un terminateur rho-indépendant provoquant le décrochage de l'ARN polymérase au niveau de la tige-boucle terminatrice [Figure 14]. La conformation antiterminatrice permettra une élongation transcriptionnelle jusqu'à la séquence codante du gène régulé. La position du ribosome impliqué dans la synthèse du peptide leader va favoriser l'une ou l'autre conformation, induisant ainsi un arrêt ou une élongation transcriptionnelle. Ce mécanisme a été particulièrement étudié chez *Escherichia coli* dans la régulation de l'opéron tryptophane *trp* impliqué dans la biosynthèse du tryptophane (Yanofsky, 1981). La séquence du peptide leader contient 14 AA, dont 2 tryptophanes à la suite. Si la quantité de tryptophane est limitée, le ribosome sera bloqué au niveau des codons correspondants (UGG) lors de la synthèse du peptide leader. La position du ribosome va alors induire la conformation antiterminatrice de l'ARN, permettant la transcription de l'opéron impliqué dans la synthèse du tryptophane. Avec une quantité suffisante de tryptophane, le ribosome ne sera plus stoppé sur les codons UGG, induisant alors une conformation terminatrice [Figure 15]. Une augmentation de l'expression de l'opéron a été mise en évidence dans des souches contenant des mutations au niveau de la tige-boucle terminatrice. Au contraire, deux systèmes déficients induisent une baisse d'expression de l'opéron. L'un nommé *trpL29*, contient une substitution du codon d'initiation AUG du peptide leader par AUA ne permettant plus la traduction du peptide leader. Le second, *trpL75*, possède un SNP (Single Nucleotid Polymorphism) en position 75, remplaçant une guanine par une adénine, induisant une fragilisation de la conformation



antiterminatrice (Yanofsky, 1981). Ce mécanisme d'atténuation est souvent retrouvé dans la régulation de la biosynthèse des AA chez les bactéries à Gram négatif (Vitreschak *et al.*, 2004).

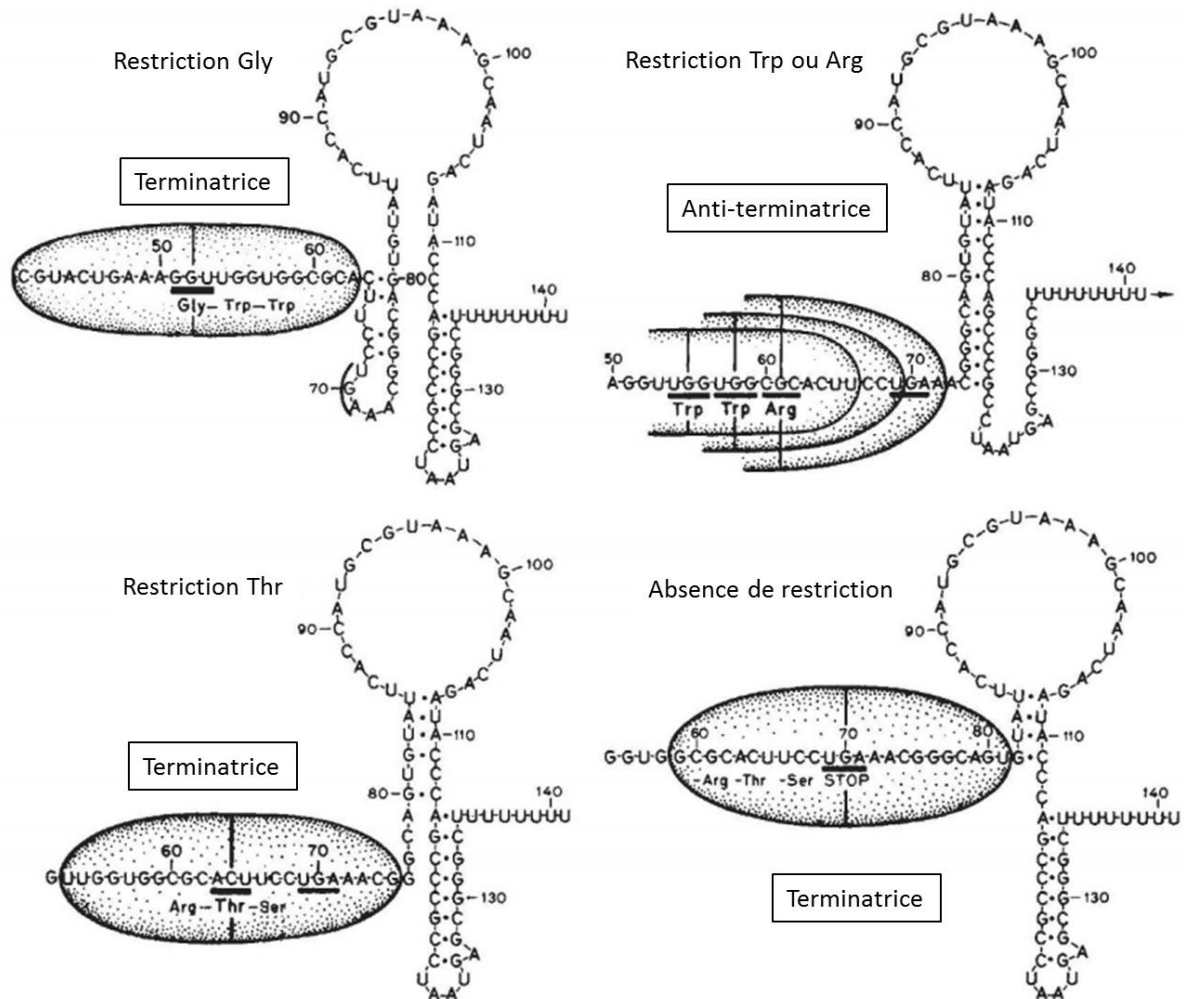


Figure 15. Représentation de la régulation par atténuation transcriptionnelle de l'opéron *trp* dépendant de la restriction ou non de différents AA (Yanofsky, 1981).

Pour l'atténuation traductionnelle, on retrouve également deux types de conformations de l'ARN : une séquestrant le site de fixation au ribosome du gène régulé et une rendant accessible le RBS permettant la synthèse de la protéine du gène régulé [Figure 14]. Comme pour l'atténuation transcriptionnelle, la position du ribosome impliqué dans la synthèse du peptide leader va favoriser l'une ou l'autre conformation, induisant ainsi une initiation ou non de la traduction de la protéine du gène régulé. Ce mécanisme de régulation est notamment impliqué dans la régulation de gènes de résistance aux antibiotiques inhibant la traduction (Dersch *et al.*, 2017). Par exemple, le mécanisme de régulation du gène *cat* (résistance au chloramphénicol) est dépendant de l'arrêt ribosomal aux alentours de 5 AA causé par l'altération de la liaison peptidique lors de l'élongation [Figure 16.A]. Dans le cas du gène *erm* (résistance phénotype MLS<sub>B</sub>), l'atténuation repose sur l'arrêt de la synthèse du peptide leader aux alentours de 8 AA assemblés causé par l'encombrement stérique de ces molécules aux niveau du tunnel de sortie peptidique du ribosome (Lovett, 1990; Lovett et Rogers, 1996; Ramu *et al.*, 2009) [Figure 16.B].

Nous allons nous intéresser à la régulation de deux gènes de résistance aux antibiotiques médiée par une mécanisme d'atténuation transcriptionnelle en lien avec les ARN étudiés dans ce manuscrit : *mef(E)* et *tet(M)*.

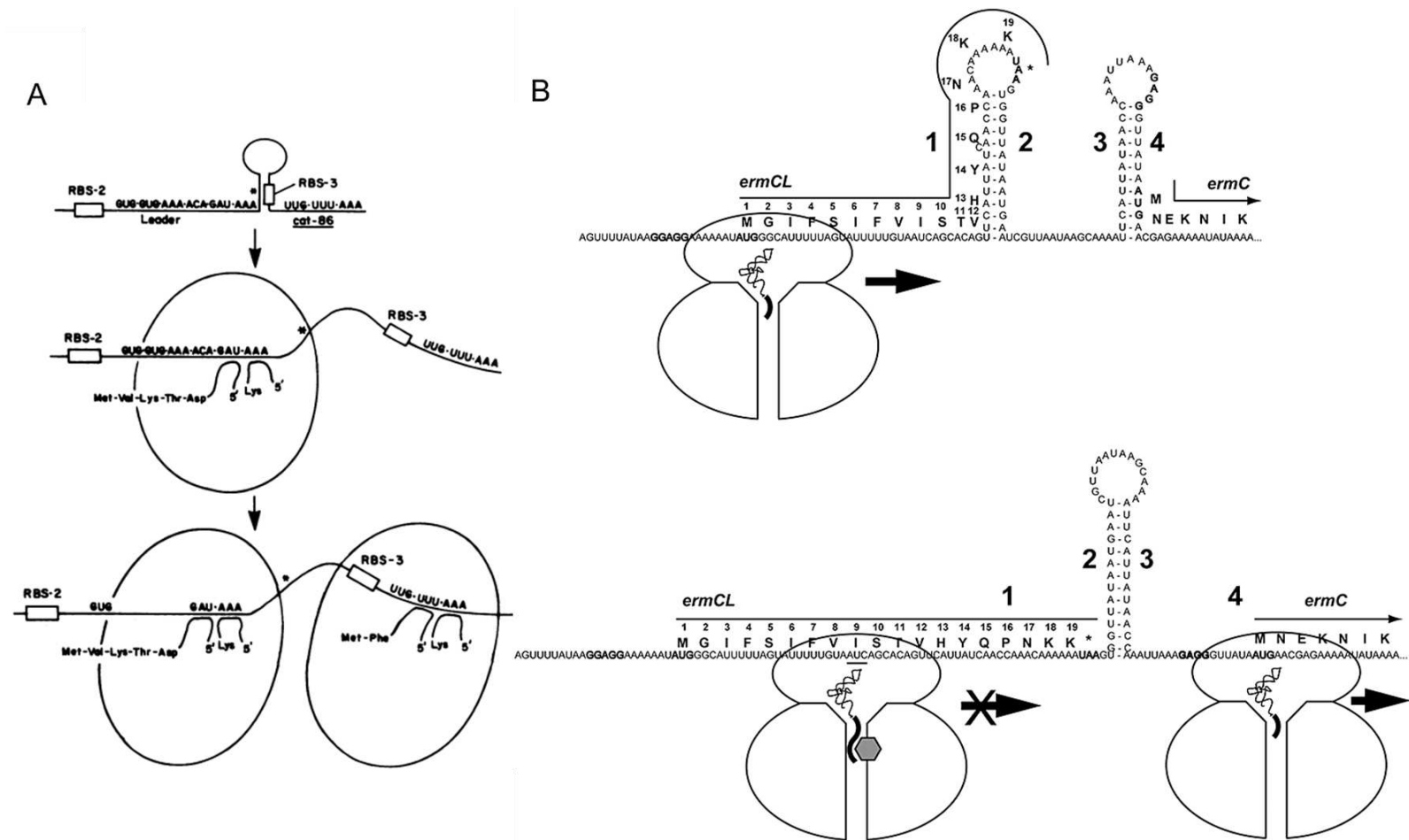


Figure 16. Représentation d'une atténuation traductionnelle impliquant des gènes de résistance aux antibiotiques (phénicolés et macrolides) (Lovett, 1990; Ramu *et al.*, 2009).

(A) Régulation du gène *cat*. RBS-2 est le RBS du peptide leader et RBS-3 est celui codant la protéine Cat (Lovett, 1990).

(B) Régulation du gène *erm(C)*. Le peptide leader est nommé *ermCL* (Ramu *et al.*, 2009).

### *a. Régulation de mef(E)*

La régulation des gènes de type *mef* (*mef(E)* et apparentés) a été décrite expérimentalement et phénotypiquement. Il existe une surexpression du gène *mef(E)* et une induction de résistance par addition de macrolides (Chancey *et al.*, 2011). Cette observation peut être en accord avec un système d'atténuation transcriptionnelle. Les auteurs ont défini la région régulatrice par 5'RACE. La région 5'UTR du gène *mef* contient 327 nucléotides précédant le codon d'initiation de *mef*. Cette région peut-être structurée en 8 domaines correspondant à 4 tige-boucles, avec un peptide leader putatif de 8 AA (A<sub>34</sub> à A<sub>57</sub>) contenu dans le domaine R1 et la tige-boucle R5-R6 impliquée dans un terminateur rho-indépendant [Figure 17]. L'élongation transcriptionnelle induite par la présence d'érythromycine (macrolide à 14 atomes) a été démontré par RNAseq. Cette induction n'est pas présente avec l'exposition au LL-37 (peptide antimicrobien) et à la spiramycine (macrolide à 16 atomes ayant un site d'interaction différent de celui de l'érythromycine). Plusieurs mutants avec diverses mutations ciblées au niveau de la région régulatrice ont été utilisés pour déterminer les régions impliquées dans le mécanisme de régulation. Cette exploration a été effectuée à l'aide d'un système rapporteur et de tests de détermination de la sensibilité à l'érythromycine. Le rôle du peptide leader a été validée avec un mutant contenant la substitution du RBS du peptide leader (AGGAGG) par une séquence « inerte » (CCTCCT). L'absence de la traduction de ce peptide leader favorisant alors la conformation terminatrice, la souche devient sensible à l'érythromycine [Figure 17]. Un mécanisme de régulation complémentaire semble exister, mis en évidence par l'importance de la structure R3-R4 qui, lorsque déléetée, induit une perte de résistance à l'érythromycine associée à une absence d'induction transcriptionnelle par l'érythromycine. Même avec une délétion des régions R6-R7 empêchant la formation du terminateur, une induction transcriptionnelle par l'érythromycine reste observable, ne pouvant pas être seulement expliquée par l'atténuation transcriptionnelle médiée par la traduction du peptide leader.

Les auteurs mettent également en évidence indirectement une potentielle protection de l'ARN d'une dégradation médiée par des RNases quand le peptide leader reste traduit. La délétion totale de la région cause une absence d'induction, mais en gardant seulement la séquence codante du peptide leader, une induction par l'érythromycine reste observable (Chancey *et al.*, 2015). Malgré un mécanisme d'atténuation transcriptionnelle bien défini dans la régulation des gène *mef*, des mécanismes de régulation complémentaires restent encore à identifier.

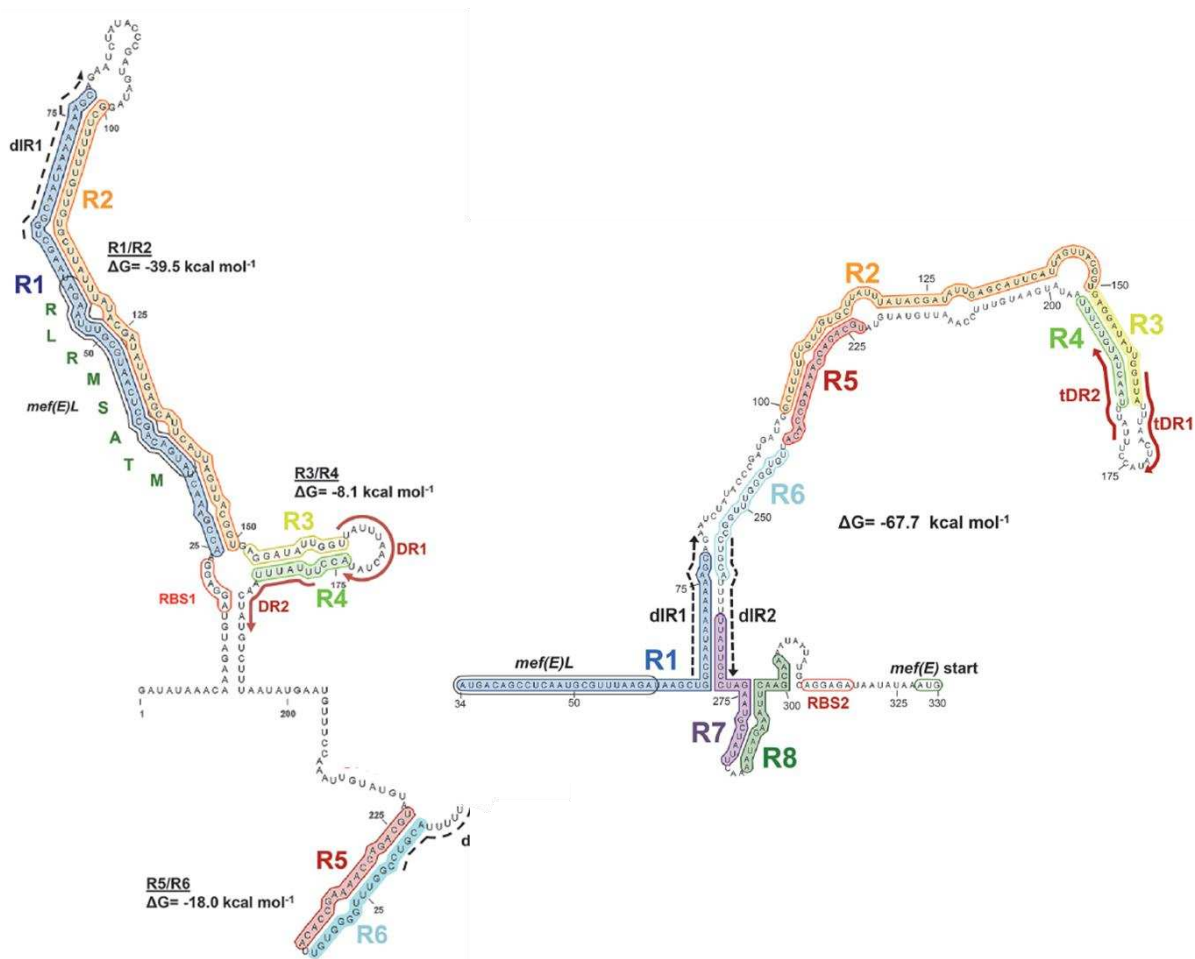


Figure 17. Représentation des deux conformations de la région 5'UTR du gène *mef(E)* (Chancey *et al.*, 2015).

La représentation à gauche est terminatrice et celle à droite est antiterminatrice. La séquence codant le peptide leader est appelée *mef(E)L*. L'équilibre thermodynamique est indiqué pour chaque structure.

### *b. Régulation de tet(M)*

Sur le même principe, nous allons nous intéresser au mécanisme de régulation du gène *tet(M)*, étant moins approfondi que la régulation de *mef(E)* dans la validation du mécanisme. La régulation de ce gène *tet(M)* est d'autant plus importante que cette régulation est impliquée dans le transfert horizontal génétique bactérien des transposons Tn916-like (Celli et Trieu-Cuot, 1998; Beabout *et al.*, 2015b). L'hypothèse que le gène *tet(M)* soit régulé par un mécanisme d'atténuation transcriptionnelle a été posée en 1992 (Su *et al.*, 1992). Après détermination de la région 5'UTR du gène *tet(M)* par 5'RACE, les auteurs ont observé un ORF pouvant potentiellement coder pour un peptide leader. Par Northern Blot, l'élongation transcriptionnelle induite par la tétracycline a pu être mise en évidence. Par contre, l'induction de la conformation antiterminatrice par la tétracycline est plus difficile à expliquer. En effet, après la fixation du ribosome sur le codon d'initiation, la tétracycline empêche le complexe aaARNt/EF-Tu de se loger au niveau du site A. En regardant les prédictions de structure, la tétracycline favoriserait donc la structure 4 [Figure 18]. Les auteurs expliquent l'induction par la tétracycline liée à une « accélération » de la traduction du peptide leader. En effet, la séquence du peptide leader contient des codons associés à des aaARNt « rares », provoquant une plus forte probabilité de pause ribosomale. Par son mécanisme d'action, l'altération générale de la traduction par la tétracycline va augmenter la quantité d'aaARNt disponible, induisant alors une « accélération » de la traduction du peptide leader. Sans pause ribosomale, la tige-boucle terminatrice ne pourra pas se former et cela provoquera une élongation transcriptionnelle. Ce premier article met en évidence seulement une élongation transcriptionnelle induite par la tétracycline, mais le mécanisme d'atténuation transcriptionnelle bien que plausible, n'est ici pas démontré (Su *et al.*, 1992). Un second article montre l'importance de cette région régulatrice annotée *Ptet* dans la transposition du transposon Tn916 (Celli et Trieu-Cuot, 1998). En effet, l'élongation transcriptionnelle

induite par la tétracycline augmente l'expression du gène *tet(M)* mais également l'expression des gènes retrouvés en aval, notamment des gènes de régulation et les gènes impliqués dans la transposition [Figure 10].

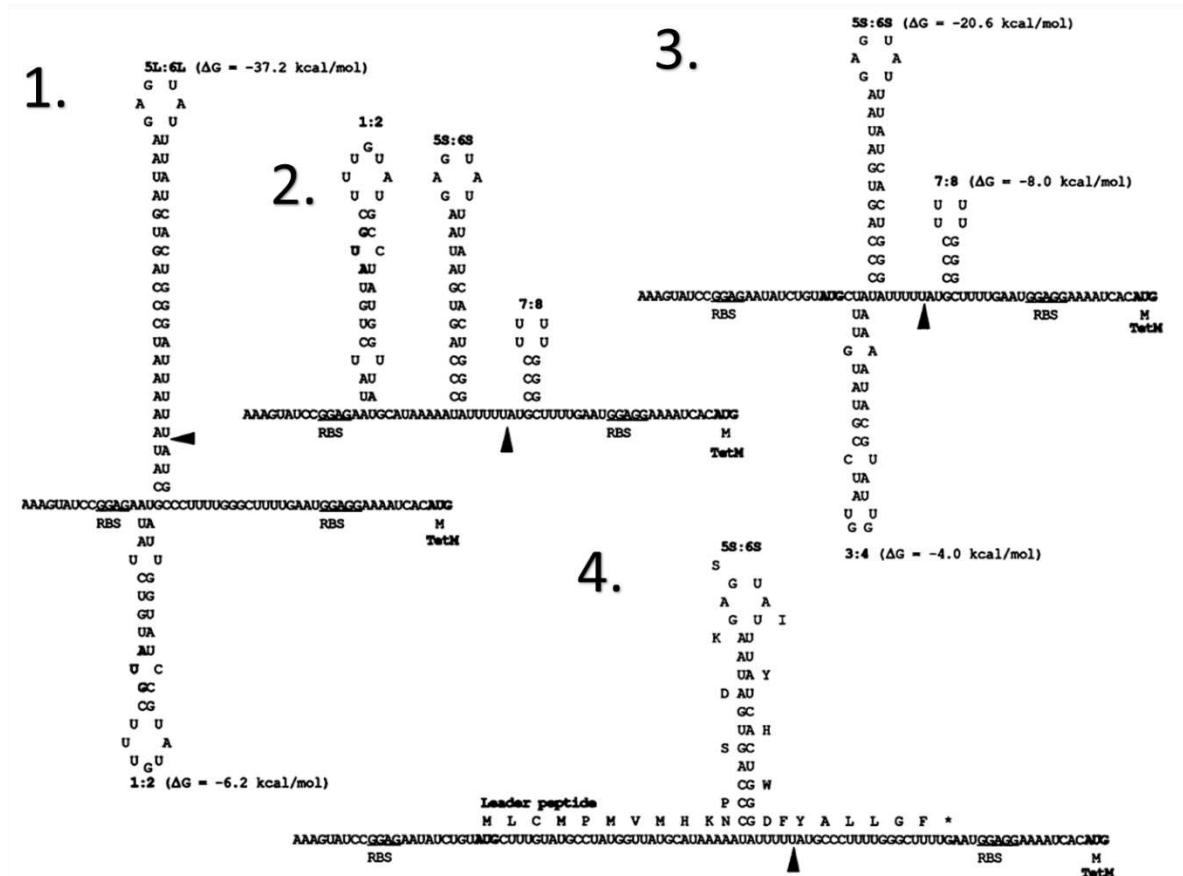


Figure 18. Structures secondaires prédites selon la stabilité des tige-boucles (Su *et al.*, 1992). Les positions du RBS du peptide leader putatif et du gène *tet(M)* sont indiquées. Le triangle noir indique la position des 5 uraciles impliqués dans le terminateur rho-indépendant.

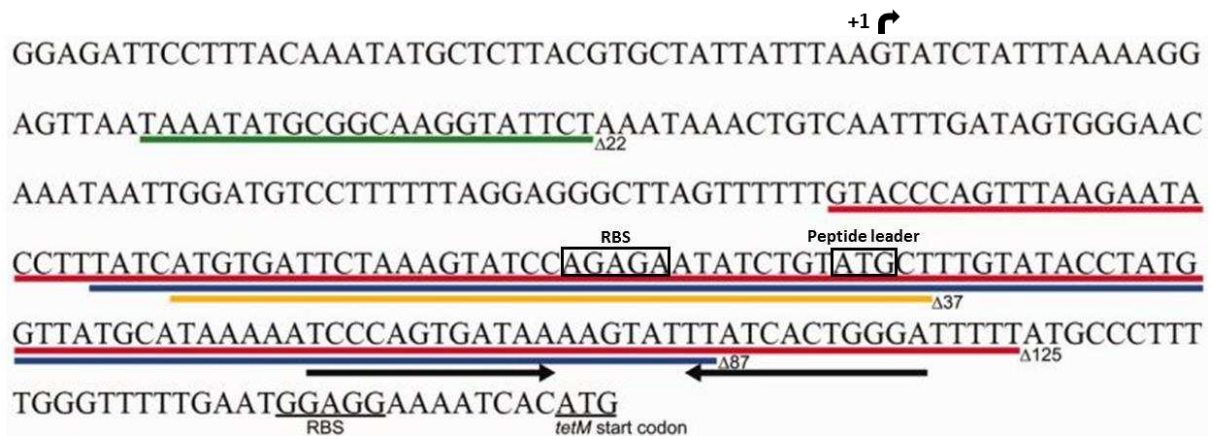
Un dernier article explorant le rôle des tétracyclines dans les échanges génétiques horizontaux a pu démontrer phénotypiquement le rôle de la région 5'UTR de *tet(M)* dans la résistance aux tétracyclines et l'hyperconjugaison du transposon Tn916 chez *E. faecalis* (Beabout *et al.*, 2015b). Cette mise en évidence a été effectuée à l'aide de souches mutantes résistantes à la tigécycline (désignées BTR) induites par une exposition répétée de concentration subinhibitrice de tigécycline. Après caractérisation par séquençage de

ces souches, les mutations retrouvées sont principalement une délétion de 12 pb au niveau du gène *rpsJ* ou des délétions au niveau de *Ptet*.

Quatre types de délétions de *Ptet* ont pu être distinguées [Figure 19]:

- BTR125 : souche avec une délétion de 125 pb contenant les délétions retrouvées dans les souches BTR87 et BTR37. La délétion contient le RBS du peptide leader jusqu'à son 22<sup>ème</sup> codon, déléant ainsi totalement la tige-boucle terminatrice. La CMI à la tigécycline de cette souche atteint 2 mg/L (contre 0,125 mg/L pour la souche initiale).
- BTR87 : souche avec une délétion de 87 pb contenant la délétion retrouvée dans la souche BTR37. La délétion contient le RBS du peptide leader jusqu'à son 17<sup>ème</sup> codon, déléant ainsi partiellement la tige-boucle terminatrice. La CMI à la tigécycline de cette souche atteint également 2 mg/L.
- BTR37 : souche avec une délétion de 37 pb incluant le RBS du peptide leader et son 2<sup>ème</sup> codon. La CMI à la tigécycline de cette souche atteint 1 mg/L.
- BTR22 : souche avec une délétion en 5' de *Ptet* située à plus de 100 pb en amont du codon d'initiation du peptide leader putatif. La CMI à la tigécycline de cette souche atteint 2 mg/L.





**Figure 19.** Position des délétions retrouvées dans les souches BTR (Beabout *et al.*, 2015b). Les séquences délétées sont soulignées en vert (BTR22), rouge (BTR125), jaune (BTR37) et bleu (BTR87).

Les souches avec la tige-boucle terminatrice altérée sont prédominantes (BTR125 et BTR87). Les souches BTR22 et BTR87 ont une surexpression constitutive de l'expression de *tet(M)*. L'excision et la *trans*-conjugaison du transposon sont augmentées avec les deux souches. Étrangement, l'expression du gène *xis* n'est augmentée qu'avec la souche BTR87, et reste inductible par la tigécycline (Beabout *et al.*, 2015b). Toutes ces observations prouvent l'importance de la région *Ptet* dans la régulation de *tet(M)* et dans la conjugaison du transposon, mais semblent indiquer l'existence de mécanismes de régulation alternatifs et/ou complémentaires à l'hypothèse d'atténuation transcriptionnelle. L'induction de *tet(M)* par des biocides autres que la tétracycline est un argument supplémentaire (Seier-Petersen *et al.*, 2014; Scornec *et al.*, 2017). Avec l'utilisation d'un système rapporteur, les auteurs ont réussi à mettre en évidence une élongation transcriptionnelle induite par des antibiotiques fixant le ribosome, et plus particulièrement, la spectinomycine et les lincosamides. Avec ces antibiotiques, l'induction est indépendante de la présence du reste du transposon, contrairement aux tétracyclines de 1<sup>ère</sup> et 2<sup>ème</sup> générations (Scornec *et al.*, 2017). L'exploration de cette région *Ptet* sera approfondie au sein de ce manuscrit (Article 1, Le Neindre *et al.*, 2021).

## 2. Riboswitch

On définira dans ce manuscrit un riboswitch comme le 5'UTR d'un gène codant dont la conformation dépendra de l'interaction « directe » avec une molécule effectrice (Serganov et Nudler, 2013). Comme pour les mécanismes d'atténuation, le changement conformationnel peut avoir un impact sur l'élongation transcriptionnelle ou l'initiation de la traduction [Figure 20]. Les riboswitchs comportent généralement des séquences consensus conservées impliquées dans l'interaction spécifique avec le ligand (séquence « specifier » ou « tige-boucle d'interaction »).

Les premiers riboswitchs mis en évidence ont pour ligand des vitamines comme la thiamine (TPP ou vitamine B1) dans la régulation de sa biosynthèse codée par l'opéron *thiC* (Miranda-Ríos *et al.*, 2001; Mironov *et al.*, 2002) ou la cobalamine/adénosylcobalamine (Cbl/AdoCbl ou vitamine B12) dans la régulation de son transporteur codé par *btuB* (Nou et Kadner, 1998; Nahvi *et al.*, 2002). Depuis, de nombreux riboswitchs ont été découverts avec une grande variété de ligands comme des petites molécules organiques, des ions, des acides aminés, des protéines ou des ARN (Serganov et Nudler, 2013) [Figure 20].

Nous allons aborder trois exemples :

- Les riboswitchs appartenant aux ribozymes, avec l'exemple du ribozyme régulant *glmS*,
- Le riboswitch impliqué dans la régulation de la biosynthèse des pyrimidines ayant pour ligand la protéine régulatrice PyrR,
- Les riboswitchs ayant pour ligand les ARNt, appelés T-box riboswitch.

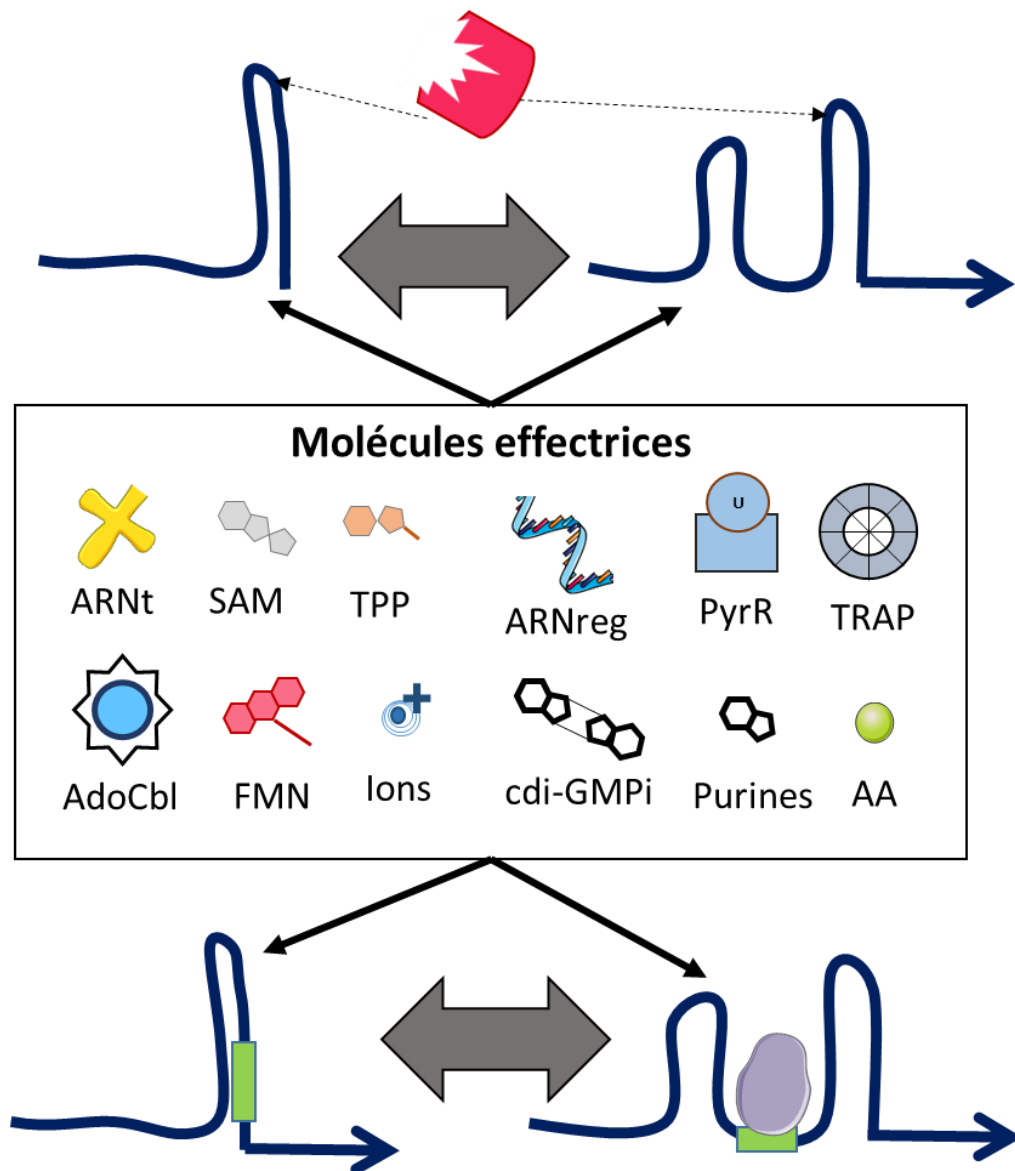


Figure 20. Représentation du principe de riboswitch.

Les molécules effectrices (liste non exhaustive) fixent directement une séquence ou une tige-boucle du riboswitch changeant ainsi sa conformation. Le RBS de la séquence codant la protéine régulée est représenté par un rectangle vert. Le ribosome est représenté par la protéine violette. L'ARN est représenté par la ligne bleue. La présence de la flèche bleue indique l'élongation transcriptionnelle de l'ARN. La RNase potentiellement recrutée est représentée par la protéine rose. Abréviations : SAM (S-adenosyl-méthionine) ; TPP (Pyrophosphate de thiamine) ; ARNreg (ARN régulateur) ; PyrR et TRAP (protéines régulatrices) ; AdoCbl (Adénosylcobalamine) ; FMN (Flavine mononucléotide) ; di-GMPc (Di-guanosine monophosphate cyclique) ; AA (Acides aminés).

### a. Régulation de *glmS* chez les Firmicutes : ribozyme

L'enzyme GlmS est impliquée dans la biosynthèse de précurseurs du peptidoglycane (Anderson et Raetz, 1987). C'est une enzyme pouvant être ciblée par certains antibiotiques (Molloy *et al.*, 1972; Kenig et Abraham, 1976).

La régulation de cette enzyme diffère selon les bactéries. Elle est médiée par un ARN régulateur, « GlmZ », chez les entérobactéries (Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008; Khan *et al.*, 2016) et par un ribozyme chez les *Firmicutes* (McCown *et al.*, 2011, 2012). On s'intéressera au second mécanisme car *E. faecium* appartient aux *Firmicutes*. Les ribozymes sont des ARN ayant une capacité de catalyse (Doherty et Doudna, 2000). Le ribozyme régulant *glmS* est localisé dans sa région 5'UTR et s'autocatalysera en présence de sa molécule effectrice qui est la glucosamine-6-phosphate (Winkler *et al.*, 2004; Savinov et Block, 2018). La catalyse induit alors une déstabilisation de l'ARNm de *glmS* en présence de RNase (Collins *et al.*, 2007) [Figure 21].

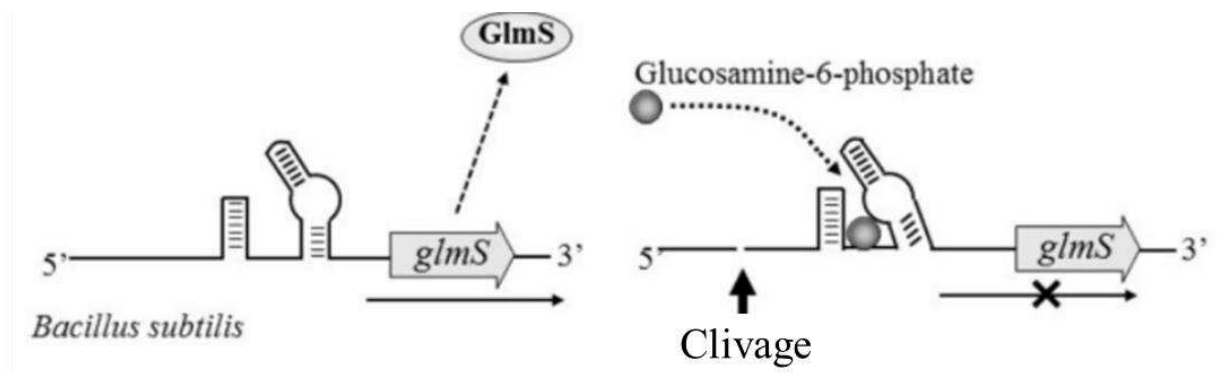


Figure 21. Représentation de la régulation de *glmS* par son ribozyme (Kawada-Matsuo *et al.*, 2017).

### *b. Régulation de la biosynthèse des pyrimidines*

Les pyrimidines sont des composés chimiques indispensables à la vie, nécessaire à la synthèse des acides nucléiques. La voie de biosynthèse des pyrimidines est similaire chez les eucaryotes et les procaryotes (Moffatt et Ashihara, 2002). Cette voie a pour objectif la synthèse d'UMP, qui pourra dériver en CTP et TTP. Elle nécessite un apport (endogène et/ou exogène) de glutamine, d'aspartate et de PRPP (5-phospho- $\alpha$ D-ribose 1-diphosphate). Cette voie de bio-synthèse comporte 6 étapes (O'Donovan et Neuhard, 1970) [Figure 22.A] :

- Synthèse de carbamyl-phosphate à partir de L-glutamine et d'ions bicarbonate : la réaction est irréversible, ATP dépendante, nécessite 2 co-facteurs ( $Mg^{2+}$  et  $K^{+}$ ) et est médiée par la carbamylphosphatase composée de 2 sous-unités codées par *carA* et *carB* (ou *pyrAA* et *pyrAB*).
- Synthèse de N-carbamyl-L-aspartate à partir de L-aspartate et carbamyl-phosphate : la réaction est médiée par une aspartate carbamyltransférase codée par *pyrB*.
- Synthèse de L-dihydroorotate à partir de carbamyl-aspartate : la réaction nécessite un co-facteur  $Zn^{2+}$  et est médiée par une dihydroorotase codée par *pyrC*.
- Synthèse d'orotate à partir de L-dihydroorotate : réaction d'oxydo-réduction couplée avec  $NAD^{+}/NADH$  médiée par une dihydroorotate déshydrogenase codée par *pyrD* (ou *pyrDI*) et *pyrK* (ou *orf2* ou *pyrDII*).
- Synthèse d'OMP (orotidine monophosphate) à partir d'orotate et de PRPP : la réaction nécessite un co-facteur  $Mg^{2+}$  et est médiée par une orotate phosphoribosyltransférase codée par *pyrE*.
- Synthèse d'UMP (uracile monophosphate) à partir d'OMP : la réaction est irréversible et est médiée par une OMP décarboxylase codée par *pyrF*.

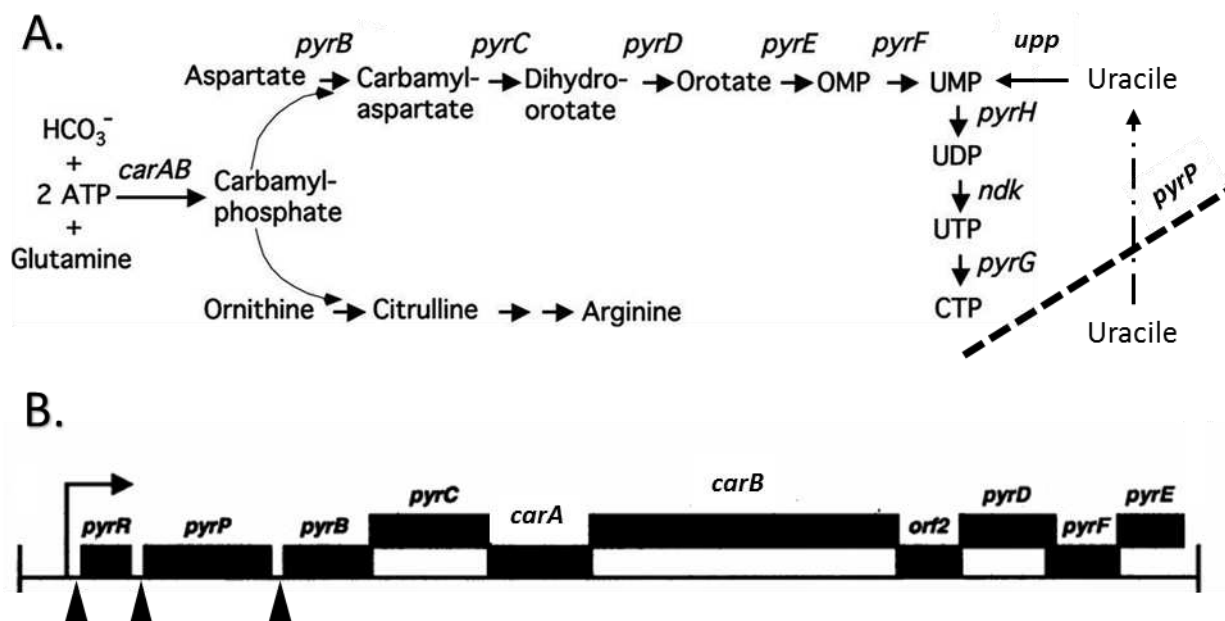


Figure 22. Représentation des gènes impliqués dans la biosynthèse des pyrimidines.

(A) Voie de biosynthèse des pyrimidines. Les flèches pleines représentent les réactions enzymatiques, les flèches en pointillé représentent un transport (Turnbough et Switzer, 2008).

(B) Organisation de l'opéron *pyr* chez *Bacillus subtilis* (Turner *et al.*, 1994). Les riboswitchs sont identifiés par des triangles noirs.

Tous les gènes impliqués dans la biosynthèse de l'UMP sont codés sur l'opéron *pyr* chez *Bacillus subtilis* (Turner *et al.*, 1994) [Figure 22.B]. Dans cet opéron, les deux premiers gènes en 5' ne sont pas impliqués dans la biosynthèse. On distingue le gène *pyrR* qui code pour une protéine de régulation et le gène *pyrP* qui code pour une uracile perméase. Chez les bactéries à Gram positif, la transcription de cet opéron dépend de la conformation d'un riboswitch localisé en 5'UTR de *pyrR*. L'expression de cet opéron est dépendante de la concentration cellulaire en UMP, avec un arrêt transcriptionnel prématuré en présence d'UMP. L'élément-clé faisant le lien entre le riboswitch et la concentration d'UMP est la protéine de régulation PyrR, codée par le gène *pyrR*. Cette protéine possède une activité uracile phosphoribosyltransférase, mais également une capacité d'interaction avec une tige-boucle spécifique de l'ARN leader [Figure 23, domaine 1-2] dans la région 5'UTR de l'opéron *pyr* qu'on appellera « tige-boucle d'interaction » (Turner *et al.*, 1998). Cette

affinité est majorée en présence d'UMP formant le complexe PyrR/UMP. La fixation du complexe PyrR/UMP sur la « tige-boucle d'interaction » va alors favoriser l'interaction des domaines 3 et 4 formant ainsi un terminateur rho-indépendant [Figure 23]. En absence de cette interaction, la conformation la plus stable est la tige-boucle antiterminatrice, favorisant l'élongation transcriptionnelle du reste de l'opéron *pyr* (Turner *et al.*, 1994, 1998; Zhang et Switzer, 2003). Cette conformation antiterminatrice est favorisée par le complexe PyrR/GTP ayant une faible capacité d'interaction avec l'ARN (Chander *et al.*, 2005).

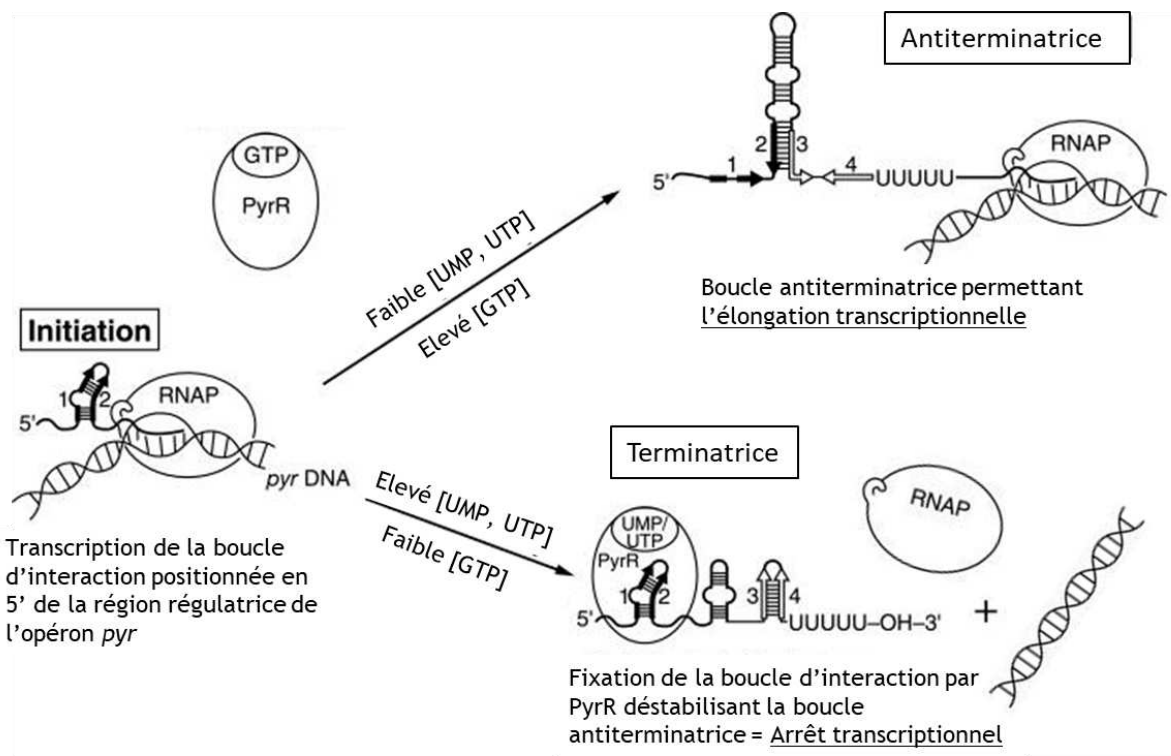


Figure 23. Mécanisme de régulation de l'opéron *pyr* médié par un riboswitch PyrR-dépendant (Zhang et Switzer, 2003).

Cette protéine PyrR est également capable de se fixer sur deux autres « tige-boucles d'interaction » positionnées au niveau d'un premier riboswitch dans la région intergénique entre *pyrR* et *pyrP* et d'un second riboswitch entre les gènes *pyrP* et *pyrB* [Figure 22.B] (Lu *et al.*, 1995; Zhang et Switzer, 2003). Il y a donc 3 points de contrôle sur la régulation de l'élongation transcriptionnelle de l'opéron *pyr*, favorisant majoritairement l'expression de *pyrR*, puis l'expression de *pyrP* et enfin l'expression des gènes de biosynthèse des pyrimidines. L'absence de PyrR favorise la transcription du premier gène *pyrR*. Malgré la présence de PyrR, en absence d'UMP, la conformation antiterminatrice sera favorisée entre *pyrR* et *pyrP*, impliquant la transcription de *pyrP*. Le gène *pyrP* code pour une uracile perméase, permettant donc un apport d'uracile pouvant provoquer une régulation de l'opéron médié par PyrR [Figure 22.A] (Ghim et Neuhard, 1994). Si l'apport d'uracile et donc d'UMP reste faible, la conformation antiterminatrice entre *pyrP* et le reste de l'opéron *pyr* sera favorisée, provoquant la transcription de tous les gènes nécessaires à la biosynthèse d'UMP. L'objectif de cet opéron est donc d'obtenir un apport exogène et endogène d'UMP nécessaire à la synthèse des acides nucléiques, et est donc autorégulé indirectement selon la concentration d'UMP. Ce mécanisme de régulation semble conservé chez les bactéries à Gram positif et a été mis en évidence dans certains d'entre eux dont *E. faecalis* (Ghim *et al.*, 1999; Switzer *et al.*, 1999).



### c. *T-box riboswitch*

L'aminacylation des ARNt est une étape indispensable pour la synthèse protéique. Cette réaction ATP-dépendante est catalysée par les aminoacyl-ARNt synthétase ou ligase (aaRS), permettant la fixation d'un AA au niveau de l'extrémité 3' de l'ARNt correspondant (Rajendran *et al.*, 2018). Chez les bactéries, divers mécanismes existent dans la régulation de ces enzymes. Cette diversité est principalement observée chez les bactéries à Gram négatif. On retrouve plusieurs mécanismes de régulation (Putzer et Laalami, 2013) :

- Régulation transcriptionnelle (inhibition de l'initiation de transcription) due à la fixation de la région promotrice par une protéine régulatrice comme la régulation du gène *lysU* par la protéine Lrp,
- Régulation post-transcriptionnelle (inhibition de l'élongation transcriptionnelle) par atténuation transcriptionnelle comme la régulation du gène *pheST*,
- Régulation traductionnelle (inhibition de l'initiation de la traduction) causée par la présence d'une protéine bloquant l'accès au RBS comme la régulation du gène *thrS* par la protéine ThrS,
- D'autres mécanismes encore mal décrits.

Contrairement aux bactéries à Gram négatif, un seul mécanisme de régulation des aaRS est décrit chez les bactéries Gram positif : T-box riboswitch (Kreuzer et Henkin, 2018).

Un T-box riboswitch est un riboswitch dont la conformation va dépendre de la fixation avec un ARNt aminoacylé (aaARNt) ou un ARNt « libre » (ARNt) [Figure 24] (Kreuzer et Henkin, 2018). Dans ce type de riboswitch, il y a deux séquences d'interaction pouvant être fixées par la molécule effectrice qui est l'ARNt. La première séquence est le codon « spécifier » fixant spécifiquement l'anticodon de l'ARNt. Ce codon « spécifier » correspondant généralement au codon de l'AA impliqué dans l'aminacylation de l'ARNt médiée par l'enzyme régulée. La seconde séquence est un triplet UGG complémentaire à l'extrémité 3' -CCA d'un ARNt « libre ». Ce triplet est localisé dans une séquence dite T-box. Cette

seconde interaction va causer un changement conformationnel du T-box riboswitch [Figure 24]. La régulation est principalement post-transcriptionnelle avec une conformation terminatrice ou antiterminatrice mais existe rarement en tant que régulateur traductionnel avec une conformation séquestrant ou non le RBS du gène régulé comme pour la régulation d'*ileS* chez les Actinobacteria (Seliverstov *et al.*, 2005; Vitreschak *et al.*, 2008).

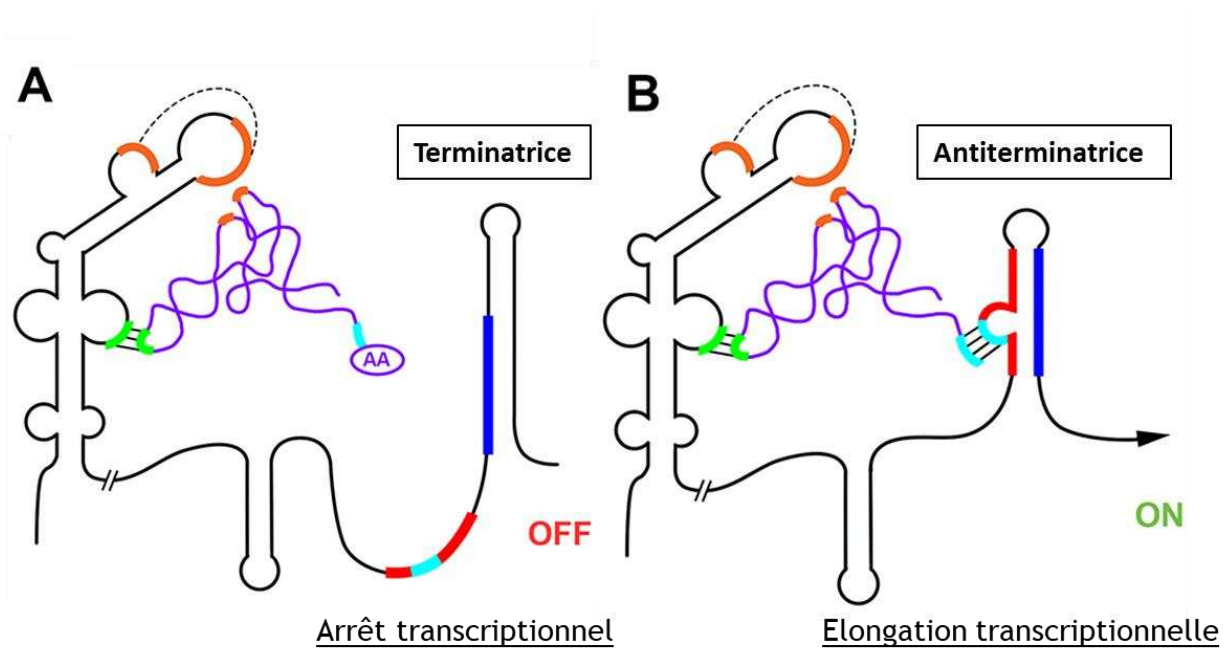


Figure 24. Représentation d'un T-box riboswitch (Kreuzer et Henkin, 2018).

Le codon « spécifier » est représenté en vert, la séquence T-box en rouge avec le triplet -UGG- en bleu cyan et les séquences conservées de la boucle « spécifier » en orange.

(A) Conformation « OFF » du T-box riboswitch induisant un arrêt transcriptionnel.

(B) Conformation « ON » du T-box riboswitch induisant l'élongation transcriptionnelle.

Ces systèmes ont été mis en évidence par l'observation de séquences consensus très conservées retrouvées dans une majorité de 5'UTR d'aaRS chez les bactéries à Gram positif (Henkin *et al.*, 1992; Grundy et Henkin, 1994; Vitreschak *et al.*, 2008). La séquence consensus généralement utilisée pour détecter un T-box riboswitch est la séquence T-box, composée de 14 bases (5'-AGGGUGGNACCGCG-3', les bases soulignées sont les plus conservées) (Rollins *et al.*, 1997; Vitreschak *et al.*, 2008). Ce système de T-box riboswitch est également présent dans la régulation de gènes impliqués dans la biosynthèse d'AA et de gènes codant des transporteurs ou des protéines de régulation (Delorme *et al.*, 1999; Vitreschak *et al.*, 2008; Schoenfelder *et al.*, 2013).

Le mécanisme de régulation a été principalement exploré dans la régulation de *tyrS* chez *B. subtilis* [Figure 25] (Rollins *et al.*, 1997). L'expression de ce gène est indirectement dépendante de la concentration de tyrosine dans le milieu (Grundy et Henkin, 1993). La région 5'UTR de *tyrS* possède une structure secondaire prédite très proche des régions 5'UTR d'autres aaRS (Grundy et Henkin, 1993; Rollins *et al.*, 1997). On retrouve strictement conservée la tige-boucle « spécifier » (Stem I) localisée en 5' du T-box riboswitch et la tige-boucle « terminatrice-antiterminatrice » positionnée en 3' du T-box riboswitch. D'autres tige-boucles peuvent être observées (Stem II, IIA/B et III) mais sont variables selon les gènes régulés et les espèces bactériennes. Une conformation simple brin est généralement observée au niveau des séquences consensus, supposant une accessibilité possible par l'ARNt (Rollins *et al.*, 1997; Vitreschak *et al.*, 2008).

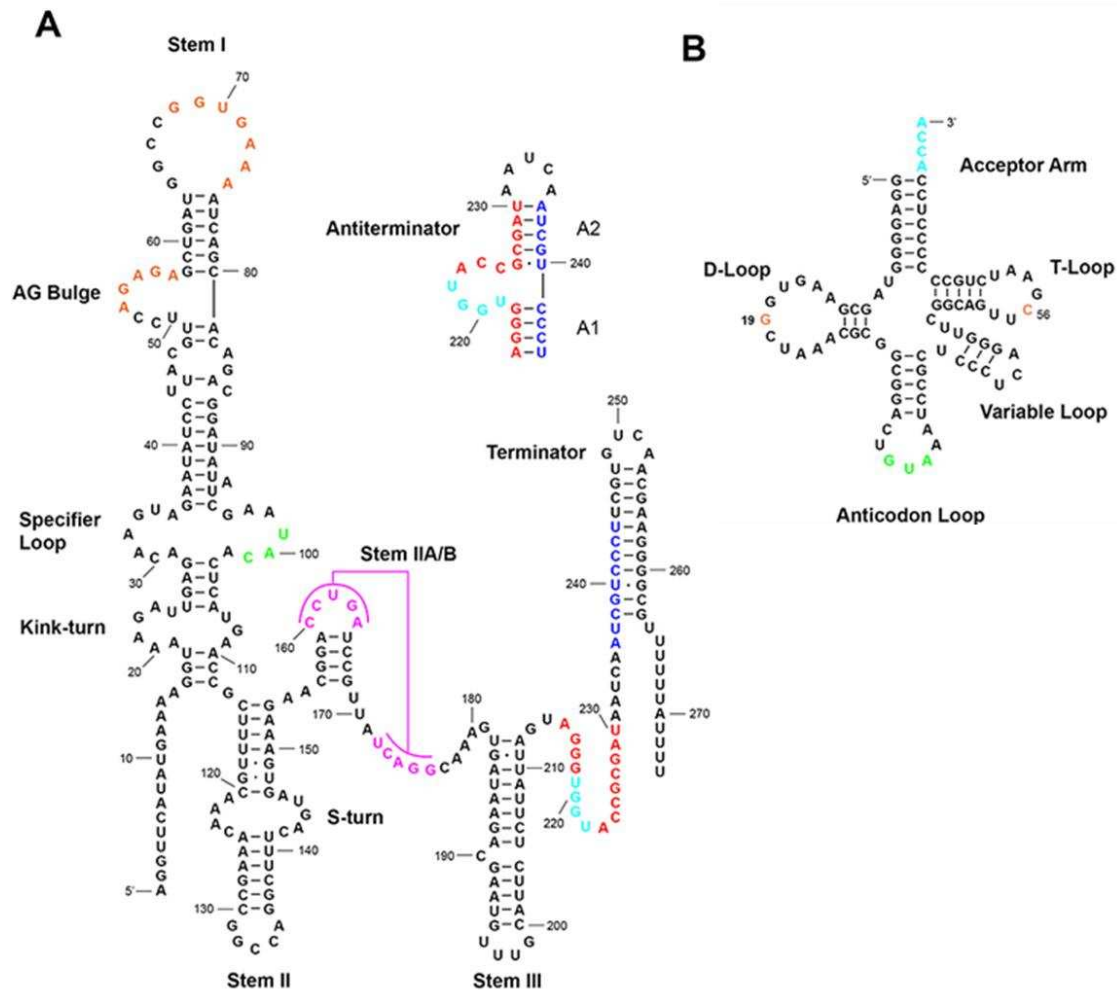


Figure 25. Structure secondaire du T-box riboswitch régulant *tyrS* et de l'ARNt<sup>Tyr</sup> chez *B. subtilis* (Kreuzer et Henkin, 2018).

(A) Représentation du 5'UTR de *tyrS*, avec les conformations terminatrices et antiterminatrices  
 (B) Représentation de l'ARNt<sup>Tyr</sup> pouvant interagir spécifiquement avec le T-box riboswitch de *tyrS*

Les séquences consensus généralement retrouvées sont des AGUA-box, AG-box, GNUG-box, F-box (5'-CCGUUA-3') et la T-box. Le codon « specifier » (non consensus au sens strict car spécifique de l'ARNt « ligand ») est généralement positionné entre la GNUG-box et la F-box. Démontré à l'aide d'un système rapporteur, les mutations au niveau de ces séquences consensus induisent la conformation terminatrice suggérant une perte d'interaction entre l'ARNt et le T-box riboswitch (Rollins *et al.*, 1997).

Une interaction directe entre l'ARNt et deux séquences spécifiques existent [Figure 24 et Figure 25] (Gerdeman *et al.*, 2002; Nelson *et al.*, 2006) :

- Entre l'anticodon de l'ARNt et le codon « spécifier » du T-box riboswitch. C'est cette interaction qui définit la spécificité de l'ARNt. Une mutation ponctuelle sur ce codon « spécifier » change le type d'ARNt qui s'y fixera, selon la séquence du nouveau codon, modifiant alors l'AA impliqué indirectement dans la régulation (Grundy et Henkin, 1993; Putzer *et al.*, 1995).
- Entre l'extrémité 3' -CCA de l'ARNt libre et la séquence T-box. La présence d'un AA sur cette extrémité 3' empêche donc l'interaction.

Le rôle des autres séquences consensus réside dans la structure secondaire du T-box riboswitch de type « plate-forme d'amarrage » favorisant le « logement » de l'ARNt au sein du T-box riboswitch (Grigg et Ke, 2013; Zhang et Ferré-D'Amaré, 2013).

La première tige-boucle (stem I) ou tige-boucle « spécifier » comporte trois domaines :

- Un motif « Kink-turn » qui comporte généralement un motif GA (5'UTR *tyrS* de *B. subtilis* : G<sub>21</sub>A<sub>22</sub>). Ce motif est localisé à la base de la tige-boucle et peut être retrouvé dans d'autres riboswitchs. Ce motif aurait un rôle de « flexibilité » de la tige-boucle, permettant d'améliorer l'affinité de l'ARNt avec la tige-boucle « spécifier » (Winkler *et al.*, 2001; Wang et Nikonowicz, 2011).
- La « spécifier loop » opposant en 5' une AGUA-box (5'UTR *tyrS* de *B. subtilis* : A<sub>32</sub>-A<sub>35</sub>) et en 3' le codon « spécifier » (5'UTR *tyrS* de *B. subtilis* : U<sub>99</sub>-C<sub>101</sub>). Cette boucle laisse le codon « spécifier » accessible pour l'interaction avec l'ARNt.
- La « plate-forme d'amarrage » incluant l'AG-box (5'UTR *tyrS* de *B. subtilis* : A<sub>53</sub>-A<sub>57</sub> ; formant l'AG Bulge) et la GNUG box (5'UTR *tyrS* de *B. subtilis* : G<sub>68</sub>-A<sub>74</sub>). En plus de la conservation des deux motifs, le nombre de nucléotides au niveau des zones simple brin est important pour garder une structure secondaire cohérente et

donc une bonne affinité avec l'ARNt ligand, notamment au niveau de sa T-loop (Rollins *et al.*, 1997; Grigg et Ke, 2013; Zhang et Ferré-D'Amaré, 2013).

Suite à cette première tige-boucle, la présence d'autres tige-boucles (stem II, IIA/B et III) est fréquente mais pas toujours conservée. Quand celles-ci sont présentes, l'orientation de ces tige-boucles semble importante pour stabiliser l'interaction avec l'ARNt (Yousef *et al.*, 2003). On y retrouve potentiellement une AGUA-box (5'UTR *tyrS* de *B. subtilis* : A<sub>143</sub>-A<sub>146</sub>) et F-box (5'UTR *tyrS* de *B. subtilis* : C<sub>166</sub>-A<sub>171</sub>) (Rollins *et al.*, 1997).

Ce mécanisme de régulation est très répandu et conservé chez les *Firmicutes*, mais également peut facilement évoluer et changer d'ARNt spécifique par simple mutation au niveau du codon « spécifier » (Vitreschak *et al.*, 2008). Par exemple, pour le cas de la régulation prédite de l'opéron *hisS-aspS* codant les enzymes HisS et AspS, on peut remarquer une variation dans le codon « spécifier » pouvant être GAC (Asp) chez la majorité des Lactobacillales, CAC (His) dans l'espèce *Lactobacillus reuteri* ou AAC (Asn) chez certains Bacillales [Figure 26]. Il existe également des T-box riboswitch plus originaux comme des T-box riboswitch en tandem (Vitreschak *et al.*, 2008) ou un exemple de T-box riboswitch spécifique de 2 ARNt (Saad *et al.*, 2013).

*E. faecium* étant un *Firmicutes*, il est normal de retrouver des T-box riboswitchs chez cette espèce (Marchand *et al.*, 2020; Michaux *et al.*, 2020). Dans ce manuscrit, nous validerons le T-box riboswitch régulant l'opéron *hisS-aspS* dans l'espèce *E. faecium*.

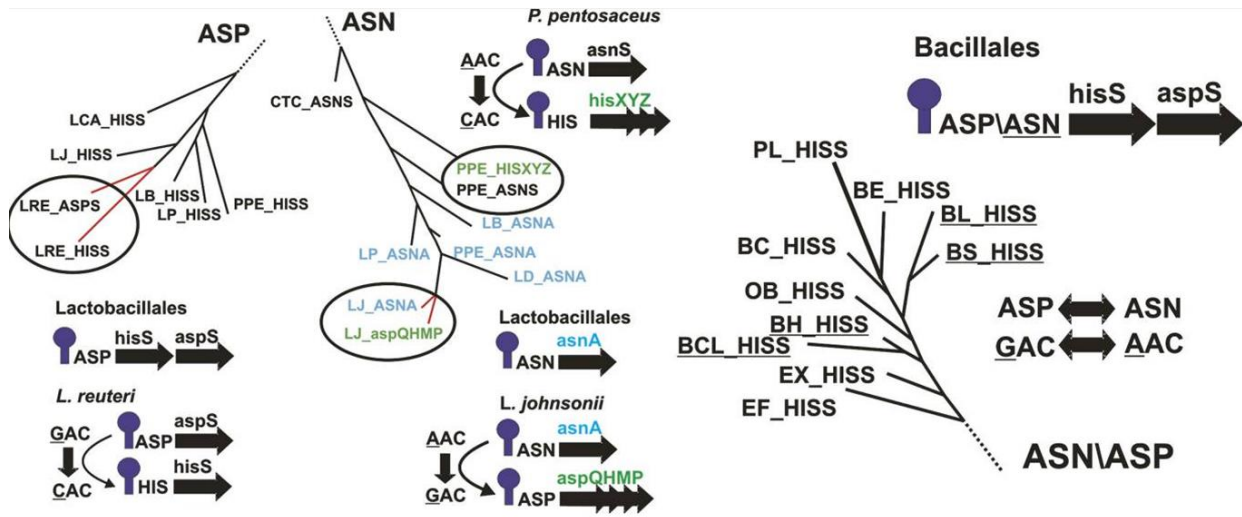


Figure 26. Arbre phylogénétique montrant l'évolution des T-box riboswitchs ayant un codon « spécifier » Asp ou Asn (Vitreschak *et al.*, 2008).

Le T-box riboswitch est représenté en violet, avec l'AA correspondant au codon. Les gènes codant des aaRS sont en noir, des protéines de biosynthèse en bleu et des transporteurs en vert. LCA = *Lactobacillus casei*, LJ = *Lactobacillus johnsonii*, LRE = *Lactobacillus reuteri*, LB = *Lactobacillus brevis*, LP = *Lactobacillus plantarum*, PPE = *Pediococcus pentosaceus*, CTC = *Clostridium tetani*, LD = *Lactobacillus delbrueckii*, EF = *Enterococcus faecalis*, EX = *Exiguobacterium*, BCL = *Bacillus clausii*, BH = *Bacillus halodurans*, OB = *Oceanobacillus iheyensis*, BC = *Bacillus cereus*, BE = *Bacillus stearothermophilus*, BL = *Bacillus licheniformis*, BS = *Bacillus subtilis*.

## **Contribution scientifique**



## Projet de recherche

Ces dernières années, *E. faecium* est devenu un pathogène majeur impliqué dans de nombreuses infections et épidémies hospitalières grâce à sa capacité d'adaptation et de dissémination. Depuis février 2017, les souches résistantes à la vancomycine (ERV/ERG) sont classées en priorité élevée par l'Organisation Mondiale de la Santé pour la recherche de nouvelles thérapies anti-infectieuses (<https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed>).

Les ARN régulateurs bactériens sont de plus en plus décrits comme acteurs majeurs dans les réponses adaptatives notamment aux stress antibiotiques et semblent impliqués dans la pathogénie et la résistance.

Ce projet de recherche fait suite à la découverte des premiers ARN régulateurs décrits chez *E. faecium* (Siné et al., 2017a). L'objectif du travail a été de déterminer le rôle fonctionnel de certains de ces ARN régulateurs pour mieux comprendre leurs liens avec l'antibiorésistance et la capacité d'adaptation d'*E. faecium*.

Ce travail est séparé en trois parties :

**1/ Une première caractérisation et analyse des ARN décrits chez *E. faecium* ayant permis :**

- De sélectionner et mettre en évidence des ARN potentiellement impliqués dans les infections humaines (clade A1), dans l'antibiorésistance (tétracyclines, macrolides, fluoroquinolones...) et dans le « fitness » bactérien (métabolisme des acides aminés, transporteurs...).

- De préparer des souches d'étude délétées par double recombinaison homologue et *trans*-complémentées de l'ARN sélectionné.

**2/ L'étude de la régulation du gène de résistance aux tétracyclines *tet(M)* par *ern0030*.** (Article soumis à « Frontiers in microbiology » et déposé sur le serveur de pré-print « BioRxiv » (Le Neindre *et al.*, 2021))

L'un des ARN candidats, *ern0030*, s'est avéré toujours être positionné dans la région 5'UTR de *tet(M)* (annoté *Ptet*) indépendamment de l'espèce bactérienne. L'augmentation transcriptionnelle du gène *tet(M)* induit par la tétracycline est connue mais seul un article a décrit le mécanisme de régulation par un mécanisme hypothétique d'atténuation transcriptionnelle (Su *et al.*, 1992). Or, ce mécanisme n'a, à ce jour, pas été validé expérimentalement dans la régulation du gène *tet(M)*. De plus, la délétion partielle en 5' de *Ptet* cause une surexpression de *tet(M)*, suggérant un mécanisme de régulation alternatif ou additionnel à celui précédemment décrit (Beabout *et al.*, 2015b).

La caractérisation génétique et l'exploration phénotypique et transcriptionnelle d'*ern0030* apportent de nouvelles connaissances dans la compréhension de ce mécanisme de régulation.

**3/ Description expérimentale de la régulation de l'opéron *hisS-aspS* par un T-box riboswitch chez *E. faecium*.** (Article en cours de rédaction pour soumission à « Nucleic Acid Research »)

Chez les bactéries à Gram positif, la régulation des aminoacyl-ARNt synthétases (aaRS) est majoritairement médiée par un système T-box, riboswitch fixant spécifiquement l'anticodon d'un ARNt. L'un des ARN candidats, *ern2050*, s'avère être positionné en amont du gène *hisS* codant une histidyl-ARNt synthétase suivi du gène *aspS* codant une aspartyl-ARNt synthétase. L'objectif ici est de valider expérimentalement *ern2050* comme un T-box riboswitch par une approche transcriptomique, phénotypique et mécanistique. Les données

obtenues et présentées dans ce manuscrit sont le fruit d'une collaboration avec les Dr F.DR. Wencker et Pr W. Ziebuhr (Université de Würzburg, Allemagne).

Les données et souches d'étude obtenues lors de la première partie ont été utilisées pour l'exploration du rôle des ARN en lien avec les antiseptiques et la colonisation murine, projets de recherche menés respectivement par les Dr L. Dejoies et Dr S. Reissier. Ces deux sujets font l'objet de 2 publications internationales en tant qu'auteur associé (Dejoies *et al.*, 2021) (Reissier *et al.*, 2021).

# Résultats et discussion

## I. Caractérisation et analyse initiale des ARN candidats

### A. Mise en évidence des ARN candidats d'intérêt

Pour la caractérisation fonctionnelle des nouveaux ARN mis en évidence chez *E. faecium*, il a été nécessaire d'obtenir des souches d'étude sur lesquelles des tests phénotypiques pouvaient être appliqués. Dans l'objectif de l'obtention de ces souches, les extrémités 5' et 3' de 9 ARN initialement validés par Northern Blot [Figure 13] (Siné et al., 2017a) ont été déterminées par 5'3'RACE mapping sur la souche de référence Aus0004, isolat clinique d'ERV porteur de l'opéron *vanB* et appartenant au CC17 (Lam et al., 2012) [Tableau 2]. Le principe de cette technique repose sur le séquençage de la jonction d'un ARN préalablement circularisé, permettant de déterminer ces extrémités 5' et 3' (Voir matériels et méthodes (Le Neindre, 2018)).

Tableau 2. Récapitulatif des coordonnées déterminées par 5'3'RACE des 9 ARN candidats.

ARN	Brin	N clones	Taille		Coordonnées Aus0004	
			NB	RACE	5'	3'
<u>ern0030</u>	+	40	245/154	233	65.144	65.376
<u>ern0120</u>	+	8	212	201	206.306	206.506
<u>ern0160</u>	-	3	362	365	231.312	231.676
<i>ern0280</i>	+	4	216	120	294.266	294.385
<u>ern1260</u>	+	9	243	218	1.532.590	1.532.807
<i>ern1300</i>	-	9	278	195	1.613.158	1.613.352
<u>ern2050</u>	-	4	238	237	2.519.553	2.519.789
<i>ern2210</i>	-	4	263/229	211	2.696.057	2.696.267
<i>ern2410</i>	+	11	326/240	257	2.907.933	2.908.189

Les ARN soulignés ont une taille déterminée par 5'3'RACE proche de celle déterminée par Northern blot (Siné et al., 2017a).

Sur les 9 ARN, 5 avaient une longueur cohérente avec celle déterminée par Northern blot [Figure 13] (Siné et al., 2017a). Des souches d'étude ont été produites pour les ARN *ern0030*, *ern0160* et *ern2050*. Pour *ern0030* et *ern2050* (Articles 1 et 2), les souches utilisées sont décrites ultérieurement. Pour *ern0160*, une première souche délétée de la séquence déterminée par 5'3'RACE a été produite par double recombinaison homologue :  $\Delta ern0160$ . Cette souche a été *trans*-complémentée par le plasmide pAT29 vide et pAT29 $\Omega ern0160$ . Le plasmide pAT29 $\Omega ern0160$  contient toute la région intergénique entre EFAU004\_00234 et EFAU004\_00235, incluant la région promotrice et le terminateur natif de *ern0160*. Ces souches ont été validées expérimentalement par Northern Blot et RT-qPCR [Figure 27]. Les souches *trans*-complémentées surexpriment l'ARN d'au moins un facteur 10 sur chaque phase de croissance.

L'exploration fonctionnelle de l'ARN *ern0160* sera traitée au sein de deux autres thèses que celle-ci (Dr L. Dejoies et Dr S. Ressler).

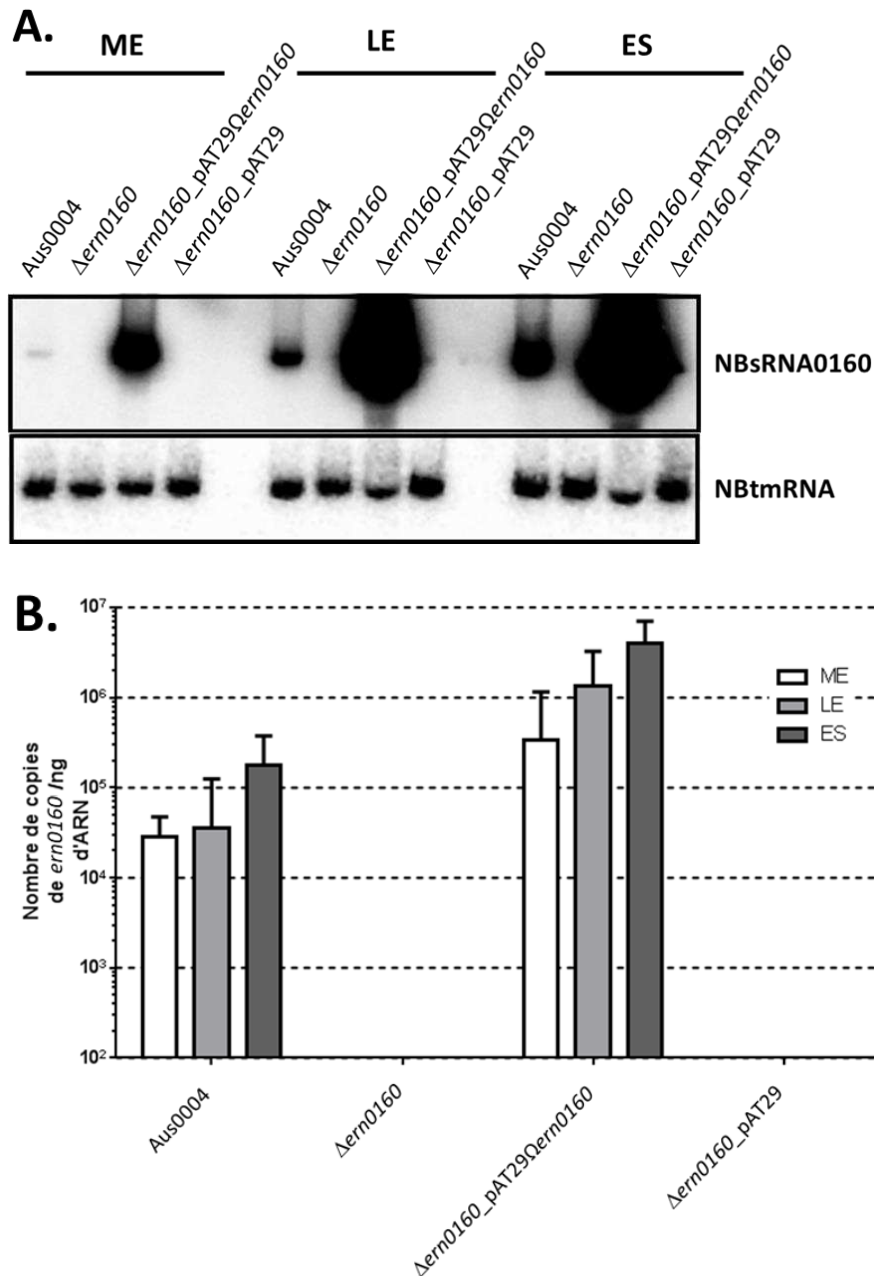


Figure 27. Validation des souches d'étude d'*ern0160*.

Les ARN ont été extraits aux 3 phases de croissance : milieu (ME) et fin (LE) de phase exponentielle et début de phase stationnaire (ES) à partir des souches sauvage (Aus0004), délétée ( $\Delta ern0160$ ), délétée *trans*-complémentée ( $\Delta ern0160\_pAT29\Omega ern0160$ ) et délétée avec un plasmide vide ( $\Delta ern0160\_pAT29$ ).

(A) Validation par Northern Blot.

La sonde NBtmRNA a été utilisée en tant que témoin de dépôt.

(B) Validation par RT-qPCR.

L'expression de l'ARN est exprimée en nombre de copies par nanogramme d'ARN totaux. Les barres représentent l'intervalle de confiance à 95 % entre 3 réplicats biologiques indépendants.

Une analyse bio-informatique a été effectuée pour compléter les informations obtenues avec les données de transcriptomique globale en lien avec la daptomycine (Sinel *et al.*, 2017a). Dans ce but, l'annotation des 61 ARN putatifs mis en évidence dans l'analyse « daptomycine » (Sinel *et al.*, 2017a) a été ajoutée dans l'analyse de transcriptomique globale en lien avec la ciprofloxacine effectuée précédemment (Sinel *et al.*, 2017b). Cette nouvelle analyse couplée aux données transcriptomiques obtenues avec le stress à la daptomycine et de conservation génomique selon les clades (Sinel *et al.*, 2017a) a permis de mettre en évidence des ARN probablement impliqués dans la tolérance aux antibiotiques ou dans la pathogénicité d'*E. faecium* [Tableau 3].

Tableau 3. Liste des 16 ARN sélectionnés.

ARN	Daptomycine			Ciprofloxacine		
	(Sinel <i>et al.</i> , 2017a)			(Sinel <i>et al.</i> , 2017b)		
	Max RPKM	FC	p-val	Max RPKM	FC	p-val
<i>ern0030*</i>	6 781	-1,91	0,5	2 690	-1,99	1
<i>ern0120*</i>	8 721	-3,06	<b>0,02</b>	3 544	-3,56	<b>&lt;1,0<sup>E-6</sup></b>
<i>ern0160*</i>	147 750	-11,19	<b>0,02</b>	42 465	-9,89	<b>&lt;1,0<sup>E-6</sup></b>
<i>ern0280*</i>	77 498	-6,64	0,05	28 132	-4,09	<b>3,8<sup>E-6</sup></b>
<u><i>ern0510</i></u>	55	-1,17	0,78	423	5,47	<b>&lt;1,0<sup>E-6</sup></b>
<u><i>ern0980</i></u>	286	2,33	<b>8,9<sup>E-6</sup></b>	152	2,34	<b>0,01</b>
<i>ern1030</i>	118	-3,70	<b>5,34<sup>E-6</sup></b>	38	-3,55	<b>1,1<sup>E-4</sup></b>
<u><i>ern1060</i></u>	57	1,05	0,9	143	3,49	<b>1,9<sup>E-5</sup></b>
<i>ern1260*</i>	2 672	2,05	0,4	1 804	2,07	1
<u><i>ern1300*</i></u>	64 411	-3,94	0,2	28 198	-4,08	<b>5,43<sup>E-6</sup></b>
<i>ern1900</i>	247	2,97	<b>&lt;1,0<sup>E-6</sup></b>	240	5,05	<b>&lt;1,0<sup>E-6</sup></b>
<i>ern1940</i>	244	2,15	<b>2,5<sup>E-4</sup></b>	125	2,93	<b>3,2<sup>E-6</sup></b>
<u><i>ern2020</i></u>	167	-4,92	0,05	262	-5,05	<b>5,61<sup>E-6</sup></b>
<i>ern2050*</i>	551	1,74	0,23	541	3,19	<b>0,03</b>
<i>ern2210*</i>	1 269	-2,64	0,3	189	-2,00	0,24
<i>ern2410*</i>	647	2,27	<b>0,03</b>	861	4,14	<b>0,001</b>

Les ARN retenus ont au moins 100 RPKM (dans une des conditions) et répondent au moins à 2 des critères suivants : variation significative en présence de daptomycine (DEseq analyse) ou de ciprofloxacine (Bonferoni analyse) ou absence du clade B (ARN souligné). Les ARN initialement validés par Northern Blot et RT-qPCR (Sinel *et al.*, 2017a) sont également disposés dans ce tableau\*.

La colonne « daptomycine » reprend les données obtenues lors de l'analyse de transcriptomique globale en présence ou non de daptomycine. La colonne « ciprofloxacine » présente la nouvelle analyse à partir de « raw data » obtenues lors du RNAseq en présence ou non de ciprofloxacine.

Sur le même principe que précédemment décrit [Figure 13] (Siné et al., 2017a), l'expression de 6 de ces ARN dont 2 non validés initialement a été évaluée par Northern Blot en présence d'une concentration subinhibitrice de ciprofloxacine (même condition utilisée lors de l'étude transcriptomique (Siné et al., 2017b), soit 1/8<sup>e</sup> de la CMI) [Figure 28]. Les profils d'expression selon la phase de croissance des 4 ARN initialement validés restent cohérents avec les profils précédemment observés [Figure 13]. La différence majeure réside dans l'observation d'une seconde forme pour l'ARN *ern1300* s'exprimant lorsque la forme à 278 nt ne s'exprime plus [Figure 28].

Les deux nouveaux ARN décrits, *ern0980* et *ern2020*, sont des ARN exclusivement retrouvés dans les clades A1 et A2 (Siné et al., 2017a). L'ARN *ern0980* présente 3 formes en milieu de phase exponentielle (ME), dont la plus exprimée est la plus grande. Les deux autres formes ne sont plus distinguables aux deux autres points de croissance. L'autre ARN, *ern2020*, est principalement exprimé en milieu de phase exponentielle (ME) [Figure 28].

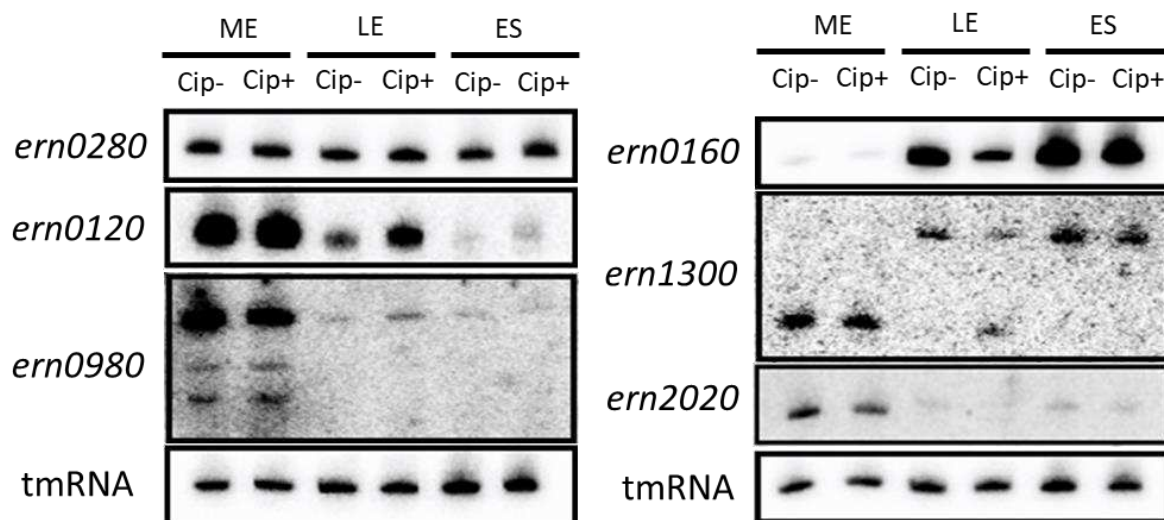


Figure 28. Expression d'ARN altérée par la présence de ciprofloxacine.

Les ARN ont été extraits aux 3 phases de croissance : milieu (ME) et fin (LE) de phase exponentielle et début de phase stationnaire (ES) à partir de la souche Aus0004 en présence (Cip+) ou non (Cip-) de concentration subinhibitrice de ciprofloxacine. La sonde NBtmRNA a été utilisée en tant que témoin de dépôt.



Avant l'investigation fonctionnelle de ces ARN, l'environnement génétique de ces ARN a été exploré pour avoir une base de réflexion sur le mécanisme d'action de l'ARN [Figure 29]. On distingue alors deux grands groupes parmi les ARN sélectionnés [Tableau 3] : des régulateurs en 5'UTR et des ARN régulateurs « conventionnels ».

Un régulateur en 5'UTR est suspecté si l'ARN est exprimé en 5' d'un gène, sur le même brin et ayant une distance de moins de 100 nucléotides entre le 3' de l'ARN et le codon d'initiation du gène adjacent. En l'absence de ces critères, une fonction en tant qu'ARN régulateur « conventionnel » est envisagée. On peut distinguer 2 types d'ARN régulateurs conventionnels : ayant une action potentiellement en *cis* ou en *trans*.

Une action en *cis* est suspectée pour :

- Les couples d'ARN sens-antisens, c'est-à-dire deux ARN exprimés sur le même *locus* mais sur les brins opposés.
- Les ARN exprimés sur le brin opposé d'un gène, de la position -50 nt avant le codon d'initiation à +50 nt après le codon stop.

En l'absence de ces critères, seule une action en *trans* peut être envisagée.

Les coordonnées 5'3' ont été ajustées (pour les ARN mappés) ou prédites à l'aide des logiciels BDGP Neural Network Promoter Prediction (Reese, 2001) et ARNold finding terminator (Naville *et al.*, 2011) [Tableau 4 & Tableau 5].

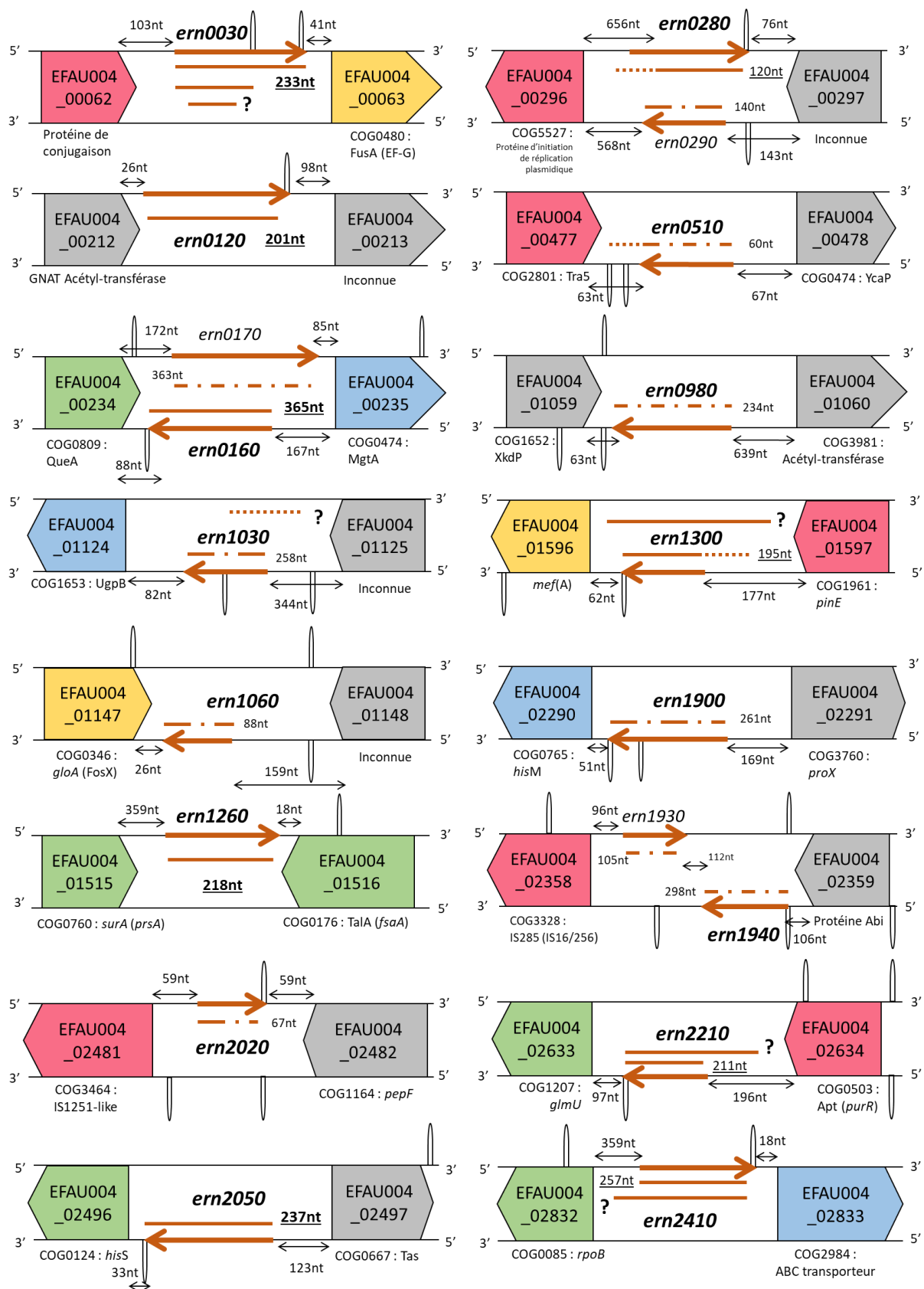


Figure 29. Représentation de l'environnement de chacun des ARN candidats.

La flèche marron représente la séquence transcrite de l'ARN. Les terminateurs prédits par ARNold finding terminator (Naville *et al.*, 2011) sont indiqués par les tige-boucles sur les brins correspondants. Les espaces entre gènes adjacents et ARN sont indiqués par les double-flèches avec le nombre de nucléotides correspondant. Les transcrits ayant un RACE mapping sont indiqués par une barre pleine et un nombre de nucléotides. Ce nombre est en gras si le RACE mapping est proche de la taille déterminée initialement par Northern blot. Les autres transcrits sont indiqués par une barre en semi-pointillé et la taille correspondant à l'estimation sur RNAseq est indiquée. Des ajustements de la taille de transcrits suggérés sont indiqués par une barre en pointillés. Les fonctions potentielles des protéines codées par les gènes adjacents de l'ARN sont indiquées par un code couleur : Gris : hypothétique ; Rose : éléments mobiles génétiques ; Vert : enzymes et métabolisme ; Bleu : transporteurs ; Jaune : résistance aux antibiotiques.

Pour évaluer les cibles ARNm en *trans* par complémentarité, des prédictions ont été effectuées par Target RNA2 (Kery *et al.*, 2014) et copraRNA (Wright *et al.*, 2014). Les séquences ajustées [Tableau 4 & Tableau 5] ont été utilisées. Pour targetRNA2, les cibles recherchées ont été fixées à -150 nt à +50 nt après l'ATG. Pour copraRNA, les critères étaient -200 nt à +100 nt après l'ATG des gènes cibles prédits. Les différents génomes utilisés dans l'analyse copraRNA sont indiqués pour chaque ARN en annexes (comportant une homologie de 100 % avec la séquence de l'ARN chez Aus0004). Cette prédiction n'est pas suffisante pour discriminer les ARN régulateurs agissant en *trans* des autres types de régulateurs investigués [Tableau 6]. Le détail des cibles prédites sont décrites en annexes pour les ARN *ern0120* et *ern0160*.

Toutes ces informations ont permis d'obtenir les premières pistes pour une investigation mécanistique et fonctionnelle de ces différents ARN.

Nous allons d'abord présenter les caractéristiques de certains de ces ARN permettant de déterminer leurs fonctions potentielles, puis approfondir ceux pour lesquels nous avons obtenu des données expérimentales.

Tableau 4. Ajustement des coordonnées 5' et 3' des ARN mappés par 5'3'RACE.

ARN	5'3'RACE		BDGP prédiction			ARNold finding terminator			Taille (nt)		
	5'	3'	5'	Nt	Score	3'	Nt	$\Delta G$	RACE	Ajustée	NB
<i>ern0030</i>	T <sub>65.144</sub>	A <sub>65.376</sub>	65.144	T <sub>+1</sub>	0,95	65376	[T <sub>205</sub> -A <sub>234</sub> ]	-12,2	233	233	245
			65.176	C <sub>+33</sub>	0,98	65246	[G <sub>85</sub> -C <sub>102</sub> ]	-7,0		70	~100
<i>ern0120</i>	T <sub>206.306</sub>	A <sub>206.506</sub>	206.308	T <sub>+1</sub>	0,35	206.500	[T <sub>164</sub> -A <sub>193</sub> ]	-8,10	201	193	212
<i>ern0160</i>	G <sub>231.676</sub>	A <sub>231.312</sub>	231.677	T <sub>+1</sub>	1,00	231.292	[G <sub>365</sub> -G <sub>386</sub> ]	-14,3	366	386	362
<i>ern0280</i>	T <sub>294.266</sub>	C <sub>294.385</sub>	294.196	T <sub>+1</sub>	0,63	294.400	[G <sub>178</sub> -G <sub>204</sub> ]	-14,7	120	204	216
<i>ern1260*</i>	A <sub>1.532.590</sub>	T <sub>1.532.807</sub>	1.532.589	A <sub>+1</sub>	0,96	1.532.800	[G <sub>182</sub> -G <sub>213</sub> ]	mfold	218	213	243
<i>ern1300*</i>	A <sub>1.613.352</sub>	A <sub>1.613.158</sub>	1.613.439	T <sub>+1</sub>	0,86	1.613.165	[T <sub>234</sub> -A <sub>270</sub> ]	mfold	195	270	278
<i>ern2050#</i>	G <sub>2.519.789</sub>	A <sub>2.519.553</sub>	2.519.789	T <sub>+7</sub> <sup>#</sup>	0,98	2.519.553	[T <sub>205</sub> -A <sub>233</sub> ]	-12,0	237	233	238
<i>ern2210</i>	T <sub>2.696.267</sub>	G <sub>2.696.057</sub>	2.696.358	T <sub>+1</sub>	0,97	2.696.072	[G <sub>230</sub> -C <sub>251</sub> ]	-10,6	211	287	263
			2.696.299	T <sub>+59</sub>	1,00					228	229
<i>ern2410</i>	C <sub>2.907.933</sub>	T <sub>2.908.189</sub>	2.907.847	A <sub>+1</sub>	0,96	2.908.177	[C <sub>309</sub> -G <sub>333</sub> ]	-7,1	257	333	326
			2.907.925	A <sub>+79</sub>	0,99					254	240

Les promoteurs sélectionnés l'ont été sur les critères suivants : selon les prédictions par BDGP (Reese, 2001), le plus cohérent avec la taille NB et l'extrémité 5' déterminée par RACE. Le +1 de transcription est annoté +1 selon le promoteur sélectionné (#excepté pour *ern2050* où l'extrémité 5'RACE a été gardée en tant que +1).

Les terminateurs ont été prédits par le logiciel ARNold finding terminator (Naville *et al.*, 2011) avec l'équilibre thermodynamique indiqué (\*deux exceptions car aucun terminateur prédit, la tige-boucle a donc été détectée par le logiciel mfold (Zuker, 2003)).

Tableau 5. Prédiction des coordonnées 5' et 3' des ARN non mappés par 5'3'RACE.

ARN	Estimation RNAseq		BDGP prédiction			ARNold finding terminator			Taille (nt)	
	5'	3'	5'	Nt	Score	3'	Nt	$\Delta G$	RNAseq	Ajustée
<i>ern0510</i>	493.357	493.298	493.385	T <sub>+1</sub>	0,91	493.263	[G <sub>104</sub> -C <sub>123</sub> ]	-14,6	60	123
						493.216	[G <sub>104</sub> -C <sub>123</sub> ]	-5,1		142
<i>ern0980</i>	1.081.114	1.080.881	1.081.132	T <sub>+1</sub>	0,90	1.080.861	[G <sub>240</sub> -C <sub>272</sub> ]	-9,8	234	272
<i>ern1030</i>	1.154.749	1.154.492	1.154.881	A <sub>+1</sub>	0,95	1.154.642	[G <sub>215</sub> -C <sub>240</sub> ]	-5,2	258	240
<i>ern1060</i>	1.175.385	1.175.299	1.175.385	A <sub>+1</sub>	0,95	Pas de prédiction			87	/
<i>ern1900</i>	2.328.162	2.327.902	2.328.241	C <sub>+1</sub>	0,90	2.327.991	[G <sub>227</sub> -C <sub>251</sub> ]	-10,3	261	251
						2.327.882	[C <sub>309</sub> -G <sub>360</sub> ]	-13,4		360
<i>ern1940</i>	2.399.705	2.399.408	2.399.720	T <sub>+1</sub>	0,97	2.399.287	[G <sub>396</sub> -T <sub>434</sub> ]	NA	298	434
<i>ern2020</i>	2.497.479	2.497.545	2.497.392	T <sub>+1</sub>	0,88	2.497.564	[A <sub>138</sub> -T <sub>173</sub> ]	- 7,0	67	173

Les promoteurs sélectionnés l'ont été sur les critères suivants : selon les prédictions par BDGP (Reese, 2001) et l'extrémité 5' estimée par les données transcriptomiques. Le +1 de transcription est annoté +1 selon le promoteur sélectionné. Les terminateurs ont été prédits par le logiciel ARNold finding terminator (Naville *et al.*, 2011) avec l'équilibre thermodynamique indiqué. NA : Non applicable.

**Tableau 6. Bilan des cibles prédites par TargetRNA2 (p-value <0,05) selon la catégorie (COG) de la cible et les différents ARN candidats.**

	ern0030	ern0120	ern0160	ern0280	ern0510	ern0980	ern1030	ern1060	ern1260	ern1300	ern1900	ern1940	ern2020	ern2050	ern2210	ern2410
<b>STOCKAGE ET TRAITEMENT D'INFORMATIONS</b>	7	12	12	15	5	13	17	7	12	12	8	11	0	13	15	0
[J] Traduction, structure ribosomale et biogenèse	0	2	4	5	2	5	7	0	1	4	1	0		6	4	
[K] Transcription	2	4	2	3	0	3	4	1	5	4	3	4		3	4	
[L] Réplication, recombinaison et réparation	5	6	6	7	3	5	6	6	6	4	4	7		4	7	
<b>TRAITEMENT CELLULAIRE ET SIGNALISATION</b>	0	10	9	8	3	5	2	4	3	3	6	10	0	11	11	0
[D] Contrôle du cycle cellulaire, division cellulaire et séparation chromosomique	0	1	0	0	0	1	0	0	0	0	1	2		0	0	
[V] Mécanismes de défense	0	4	3	0	0	1	1	0	0	0	0	1		3	1	
[T] Mécanismes de transduction des signaux	0	1	1	3	0	1	0	1	1	1	0	1		1	2	
[M] Biogenèse de la paroi cellulaire, de la membrane et de l'enveloppe	0	4	4	2	1	0	0	3	2	2	2	2		6	5	
[U] Trafic intracellulaire, sécrétion et transport vésiculaire	0	0	1	1	1	0	1	0	0	0	0	2		1	0	
[O] Modification posttraductionnelle, renouvellement protéique et protéines chaperonnes	0	0	0	2	1	2	0	0	0	0	3	2		0	3	
<b>METABOLISME</b>	4	20	14	20	9	15	10	11	20	7	11	6	0	20	21	0
[C] Production et conversion d'énergie	0	1	3	1	0	1	0	1	1	1	0	0		4	0	
[G] Transport et métabolisme des sucres	3	6	7	5	3	4	3	4	9	3	3	4		10	10	
[E] Transport et métabolisme des acides aminés	1	3	0	3	0	4	4	2	6	1	5	0		3	1	
[F] Transport et métabolisme des nucléotides	0	3	1	4	3	0	1	0	3	1	2	1		0	3	
[H] Transport et métabolisme des coenzymes	0	3	0	1	1	0	1	2	0	0	0	0		2	1	
[I] Transport et métabolisme des lipides	0	2	0	2	2	0	0	1	0	0	0	0		0	2	
[P] Transport et métabolisme des ions inorganiques	0	2	3	4	0	6	1	1	1	1	1	1		1	4	
<b>FAIBLEMENT CARACTERISEE</b>	9	23	28	44	20	28	17	20	22	22	13	22	0	25	33	0
[R] Fonction générale	2	7	6	9	4	5	5	4	6	5	6	7		5	8	
[S] Fonction inconnue	3	7	8	14	7	4	2	3	6	5	5	3		4	8	
[-] Protéine hypothétique	4	9	14	21	9	19	10	13	10	12	2	12		16	17	
<b>MIXTE</b>	1	5	1	5	1	3	4	1	4	1	4	1	0	1	3	0

## B. Régulateurs en 5'UTR

### 1. Antibiotique-dépendant

#### a. Tétracyclines et *ern0030*

L'ARN *ern0030* (initialement nommé sRNA\_0030) est le 5<sup>ème</sup> ARN le plus exprimé des 16 sélectionnés [Tableau 3]. L'extrémité 3' déterminée par 5'3'RACE mapping, correspondant au terminateur prédit, est située à moins de 50 nucléotides du codon d'initiation du gène EFAU004\_00063 annoté comme un facteur d'élongation EF-G, suggérant fortement une *cis*-régulation en 5' de ce gène [Figure 29]. Cet ARN est largement retrouvé dans les transposons de la famille Tn916-like répandus dans diverses espèces bactériennes (Siné et al., 2017a). De plus, il est localisé dans la région 5'UTR du gène de résistance à la tétracycline *tet(M)* au sein du transposon Tn916. L'alignement de séquence du gène *tet(M)* chez Aus0004 a permis d'identifier la première partie du gène *tet(M)* correspondant au gène EFAU004\_00063 et la seconde partie du gène *tet(M)* correspondant au gène EFAU004\_00067. L'annotation de ces deux gènes comme EF-G est sûrement due à l'homologie de la séquence peptidique existante entre les protéines Tet(M) et le facteur d'élongation EF-G (Leipe et al., 2002). Chez Aus0004, *ern0030* est donc localisé dans la région 5'UTR d'un gène *tet(M)* tronqué, au sein d'un transposon de la famille des Tn5801-like de type A5, apparenté au transposon Tn916 (León-Sampedro et al., 2016).

Dans le but de déterminer la fonction de *ern0030* dans la résistance aux tétracyclines, une autre souche a été utilisée, la souche E1162 résistante aux tétracyclines et portant l'opéron *tet(M)-tet(L)* (van Schaik et al., 2010). L'ARN *ern0030* est localisé sur le contig00130 tandis que les gènes *tet(M)* et *tet(L)* sont localisés sur le contig00022. Par séquençage de Sanger, la localisation d'*ern0030* dans la région 5'UTR de *tet(M)* a été confirmée dans la souche E1162. L'expression d'*ern0030* a également été validée dans cette seconde souche [Figure 30]. Une troisième forme d'*ern0030* a ainsi été mise en

évidence, le transcrit étant plus petit que les deux autres et principalement exprimé en fin de croissance.

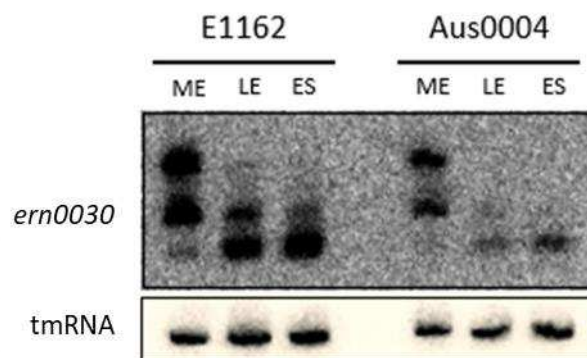


Figure 30. Validation de l'ARN *ern0030* dans les souches E1162 et Aus0004.

Les ARN ont été extraits aux 3 phases de croissance : milieu (ME) et fin (LE) de phase exponentielle et début de phase stationnaire (ES). L'ARN *tmRNA* a été utilisé en tant que témoin de dépôt.

Des essais préliminaires de détermination de demi-vie ont également été effectués dans cette souche [Figure 31]. N'ayant effectué qu'un seul réplicat biologique, les demi-vies de ces ARN n'ont pas été calculées. De plus, le délai d'exécution semble non négligeable (différence entre T et 0), limitant la détermination précise du temps de demi-vie. Néanmoins, en comparant avec le délai de dégradation de l'ARN *tmRNA*, nos essais mettent en évidence :

- Deux ARN avec une stabilité comparable au *tmRNA* : *ern0120* et *ern0160*.
- Un ARN avec une stabilité limitée : *ern0030*.
- Deux ARN dégradés dès l'ajout de rifampicine : *ern1260* et *ern2050*.



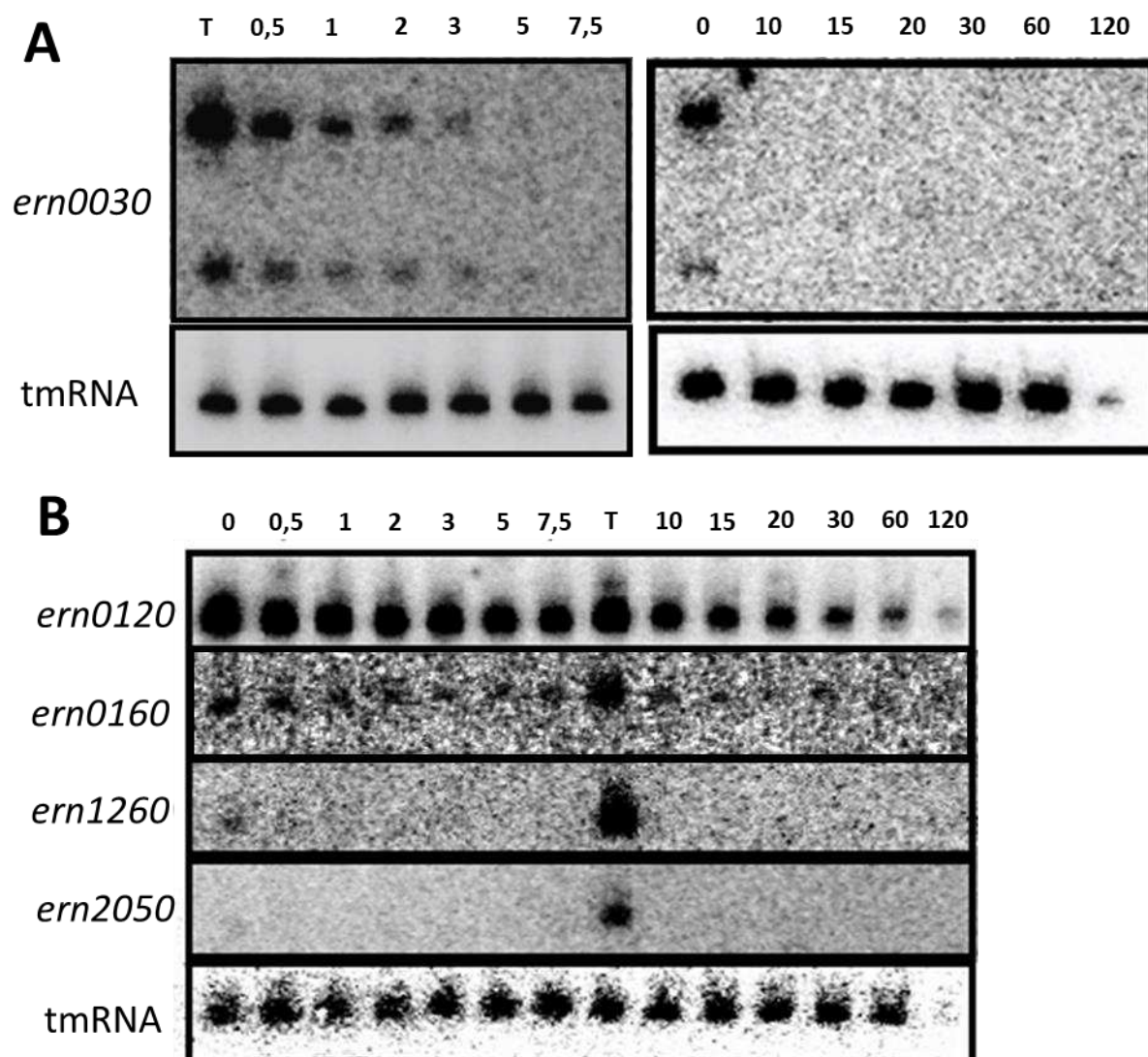


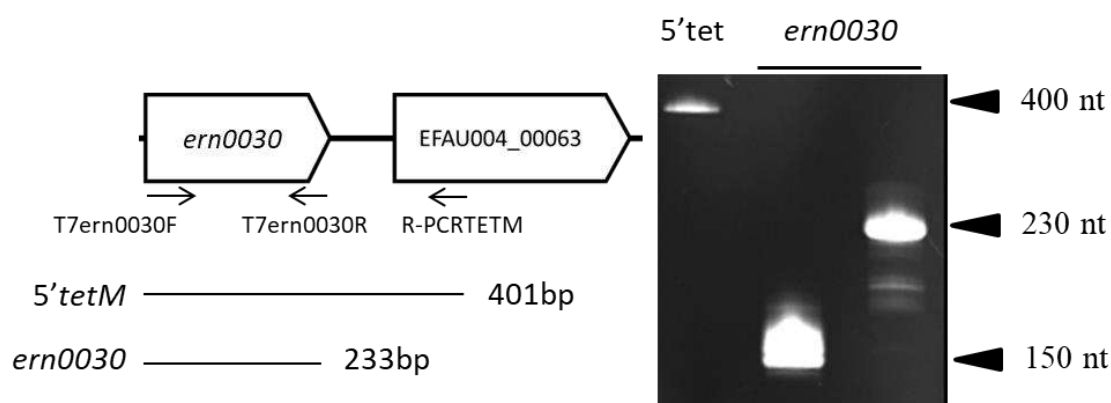
Figure 31. Essai de détermination de demi-vie dans la souche E1162 d'*E. faecium*.

La souche E1162 a été mise en culture en milieu BHI à 37°C, jusqu'à atteindre le milieu de phase exponentielle (ME) ( $DO_{600nm} = 1$ ). Un premier culot bactérien a été effectué, mis dans la glace pour obtenir l'extrait ARN T (témoin utilisé pour estimer le délai de prise en charge des culots bactériens obtenus lors de l'essai). De la rifampicine (inhibiteur de la transcription) a ensuite été ajoutée jusqu'à l'obtention d'une concentration  $C = 200$  mg/L. Des culots successifs ont ensuite été effectués à différents temps : 0 ; 0,5 ; 1 ; 2 ; 3 ; 5 ; 7,5 ; 10 ; 15 ; 20 ; 30 ; 60 ; 120 minutes. La dégradation des ARN a été révélée par Northern Blot avec les sondes correspondantes. L'ARN tmRNA a été utilisé comme témoin de dépôt et contrôle positif d'ARN à longue demi-vie.

(A) Estimation de la demi-vie de l'ARN *ern0030*.

(B) Estimation des demi-vies des ARN *ern0120*, *ern0160*, *ern1260* et *ern2050*.

Malheureusement, la délétion par double recombinaison d'*ern0030* dans la souche E1162 n'a pas fonctionné, probablement due à une localisation plasmidique de la séquence étudiée. Des souches d'étude exprimant *tet(M)* dérivant d'*E. faecium* Aus0004 ont alors été produites. Dans l'objectif d'effectuer des retards sur gels, des transcriptions *in vitro* d'*ern0030* et de la région 5'UTR jusqu'au 100 premiers nucléotides de *tet(M)* ont été effectués. Une originalité a pu être observée lors de cette transcription, avec l'observation de 2 formes de 230 et 100 nucléotides pour la transcription correspondant au 5' et 3' d'*ern0030* mais qu'une seule forme d'environ 350 nucléotides correspondant au 5' d'*ern0030* jusqu'au début de la séquence codante de *tet(M)* [Figure 32]. L'investigation fonctionnelle d'*ern0030* dans la résistance aux tétracyclines est décrite ultérieurement et fait l'objet d'un article soumis à « Frontiers in microbiology » (Article 1) ainsi qu'un dépôt sur le serveur de pré-print « BioRxiv » (Le Neindre *et al.*, 2021).



**Figure 32.** Contrôle des transcrits obtenus par transcription *in vitro* après élution.  
A gauche, la taille des transcrits théoriques est indiquée selon la position des oligonucléotides utilisées pour produire la matrice utilisée lors de la transcription *in vitro*.  
A droite, visualisation des transcrits obtenus sur gel acryl-urée 8% révélé par BET.

### *b. Macrolides et ern1300*

L'ARN *ern1300* (initialement nommé sRNA\_1300) est le 3<sup>ème</sup> ARN le plus exprimé des 16 sélectionnés [Tableau 3]. L'extrémité 3' déterminée par 5'3'RACE mapping est située à environ une soixantaine de nucléotides du codon d'initiation du gène EFAU004\_01596 annoté comme le gène de résistance aux macrolides *mef*(A) [Figure 29].

Cet ARN est retrouvé dans moins de 2 % des souches d'*E. faecium*, mais est très répandu dans les souches de streptocoques et est également localisé dans la région 5'UTR des gènes *mef*. Le mécanisme de régulation des gènes *mef* a été bien décrit dans l'espèce *Streptococcus pneumoniae* (Chancey *et al.*, 2011, 2015). L'ARN *ern1300* semble donc correspondre à la région régulatrice des gènes *mef* impliquant un mécanisme d'atténuation transcriptionnelle [Figure 17]. L'alignement d'*ern1300* de la souche Aus0004 d'*E. faecium* avec la région régulatrice de *mef*(E) de la souche GA17457 de *S. pneumoniae* montre une conservation importante des différents domaines [Figure 33]. Cette conservation implique une importance de la structure secondaire de cette région régulatrice. De plus, la séquence impliquée dans la traduction du peptide leader est également conservée. Les différences majeures sont localisées au niveau de la région promotrice, comme cela a déjà été décrit (Chancey *et al.*, 2015). Il y a également quelques mutations au niveau du domaine R3 et du domaine R6. Comme décrit précédemment (partie bibliographique), malgré un mécanisme d'atténuation transcriptionnelle démontré, il existerait un mécanisme de régulation complémentaire nécessitant de plus amples investigations. Cependant, les gènes *mef* étant peu retrouvés chez *E. faecium*, l'intérêt de l'exploration de la régulation de ces gènes reste limité chez cette espèce.

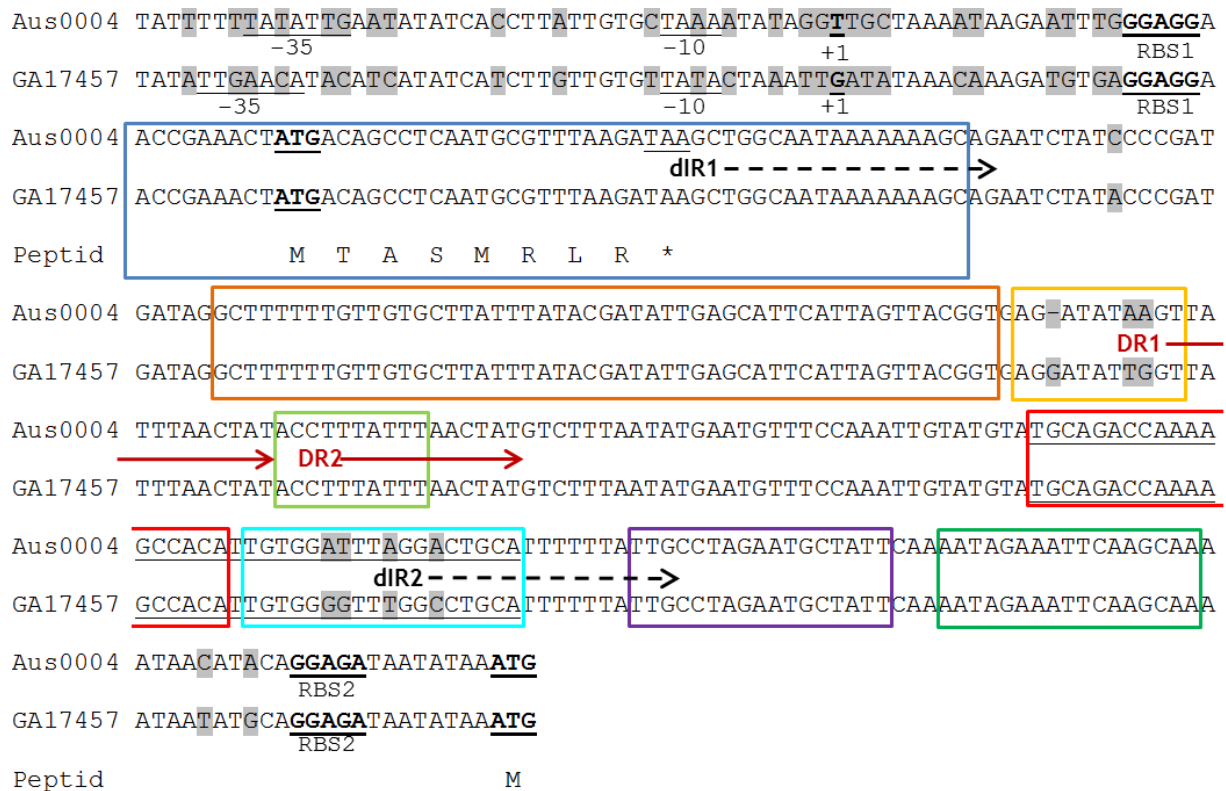


Figure 33. Alignement des séquences nucléotidiques des séquences 5'UTR du gène *mef(A)* pour la souche Aus0004 et *mef(E)* pour la souche GA17457.

Les séquences promotrices sont soulignées et annotées par la numérotation correspondante (-35 ; -10 ; +1). Les RBS et codon d'initiation sont soulignés et en gras. RBS1 correspond au peptide leader et RBS2 correspond à la protéine de résistance. Les séquences soulignées sans annotation correspondent au terminateur. Les différents domaines présentés par Chancey *et al.* (Chancey *et al.*, 2015) sont indiqués par les rectangles de couleur avec : R1 en bleu foncé, R2 en orange, R3 en jaune, R4 en vert clair, R5 en rouge, R6 en bleu cyan, R7 en violet et R8 en vert foncé. Les séquences répétées directes (DR) sont indiquées par les flèches rouges et les séquences répétées inversées (dIR) sont indiquées par la flèche en pointillés noirs. La séquence du peptide leader est indiquée sur la ligne « Peptid ».

## 2. T-box riboswitch

### a. L'opéron *hisS-aspS* et *ern2050*

Initialement, l'ARN *ern2050* (initialement nommé sRNA\_2050) semblait être un bon candidat dans la régulation de la résistance à la daptomycine [Figure 13] (Siné et al., 2017a). Malheureusement, la CMI à la daptomycine reste inchangée (1,5 mg/L par bandelette E-test®) entre les souches d'étude concernant *ern2050* (Aus0004 WT,  $\Delta$ *ern2050*-Aus0004,  $\Delta$ *ern2050*-Aus0004\_pAT29,  $\Delta$ *ern2050*-Aus0004\_pAT29 $\Omega$ *ern2050*).

Les coordonnées déterminées par 5'3'RACE d'*ern2050* indiquent que cet ARN est probablement dans la région 5'UTR du gène EFAU004\_02496 annoté *hisS*, gène codant l'histidyl-synthétase [Figure 29]. La stabilité très faible de cet ARN limite une potentielle action en *trans* en tant qu'ARNreg [Figure 31] et renforce l'hypothèse d'une régulation en 5'UTR. Les aminoacyl-ARNt synthétases (aaRS) sont généralement régulées en 5'UTR par un T-box riboswitch chez les *Firmicutes* (Kreuzer et Henkin, 2018). La séquence d'*ern2050* est bien retrouvée (100% d'homologie avec l'ID TBDB [F1TD5WKQ](#)) sur la base de données TBDB (Marchand et al., 2020) en tant que T-box riboswitch régulateur du gène *hisS*. De plus, il s'agirait de la régulation d'un opéron *hisS-aspS* comme décrit précédemment [Figure 26] (Vitreschak et al., 2008). La validation expérimentale d'*ern2050* comme T-box riboswitch est décrite ultérieurement et fait l'objet d'un article en cours de rédaction (Article 2, collaboration avec le Dr F.DR. Wencker et le Pr W. Ziebuhr, Université de Würzburg, Allemagne).

## b. Opéron de gènes codant des protéines putatives d'import de l'histidine et *ern1900*

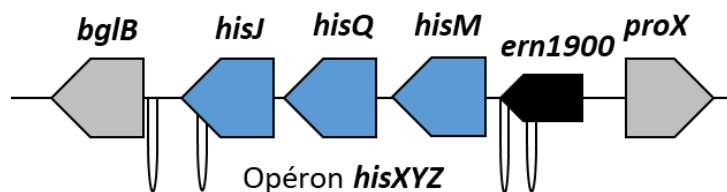
L'ARN *ern1900* est un ARN moins exprimé que les 10 ARN initialement validés par Northern Blot (Siné et al., 2017a) mais semble être impliqué dans le stress antibiotique [Tableau 3]. La position de cet ARN estimée par RNAseq le localise dans la région 5'UTR du gène EFAU004\_02290 annoté comme un ABC transporteur impliqué dans le transport des acides aminés (COG 0765 HisM). Chez les *Firmicutes*, la régulation de ces transporteurs peut être médiée par des T-box riboswitchs (Gutiérrez-Preciado et al., 2009). On retrouve effectivement les différentes séquences consensus des T-box riboswitchs dans la séquence d'*ern1900* [Figure 34].

```

2328280    AATGAAAATAATCATGTAAATTACACGTAAAGTTAAAACAAATGACATCGAAAAGAA
              -35                    -10          +1
2328222    GAGTAGCCGTTAGGTGTTTTTAGAGAGTCTCTGGTTGGTGGAAAGAGGCAGCAGTTAA
              AGUAbbox                AGbox                GNUGbox
2328164    CCGTGAAACACATCTTGGAGTCGATTTTCTGAAGCGTATAGTAGGAAAATCCGGGGAG
              Codon
2328106    CCCGTTACAGCTGCGGTCGTTGCTAGATGACCAGATAAGGCTGGCTTTTCCGTGAGGA
              F-box
2328048    GTCGGTAAAAAAAGGTGAACCACGTGATTAACGTCCTTTGATCTTCGGATCAAAGGA
              T-box                                -10,3 kcal/mol
2327990    CTTTTTGTTTTTCATCTGAACGGCAACGACTGTGTGAGGTTTTTCTGCAAGGAAAAAA
2327932    TAAACAGAGGGTGAACCACGTATAATACGTCCTCTTGGCATTGCCAAGAGGATTTT
              T-box                                -13,4 kcal/mol
2327874    TTATTTTAAAGGAGGAAAAGAAGATGGAGTATTTGATTGAAGTTATGCCGCAATTATT
              RBS                                M E Y L I E V M P Q L L
  
```

**Figure 34.** Région non traduite en 5' (5'UTR) du gène codant la protéine hypothétique HisM. Les séquences consensus T-box sont annotées en orange, vert (Codon « spécifier ») et bleu. Le +1 de transcription prédit par BDGP est indiqué. Les terminateurs rho-indépendants prédits par ARNold finding terminator (Naville et al., 2011) sont soulignés. La séquence surlignée en gris correspond à la séquence prédite par RNAseq. Les séquences importantes dans les initiations de transcription et traduction sont en gras (Promoteur transcriptionnel, RBS, codon d'initiation...).

La séquence contient deux terminateurs prédits précédés chacun d'un motif T-box, indiquant qu'*ern1900* serait un T-box riboswitch en tandem (Vitreschak *et al.*, 2008). En effet, Vitreschak *et al.* avaient détecté un tel riboswitch régulant l'opéron *hisXYZ* codant pour des transporteurs putatifs chez les *Lactobacillales* dont font partie les entérocoques (Vitreschak *et al.*, 2008). De plus, la séquence d'*ern1900* est bien retrouvée (100 % d'homologie avec l'ID TBDB [LVNDX2SZ](#)) sur la base de données TBDB (Marchand *et al.*, 2020) en tant que T-box riboswitch spécifique de l'histidine et régulateur d'un ABC transporteur d'AA. On peut donc supposer qu'*ern1900* soit un T-box riboswitch régulant des transporteurs impliqués dans l'import de l'histidine [Figure 35]. D'autres arguments en faveur de cette hypothèse sont décrits ultérieurement (Article 2).



**Figure 35.** Représentation de l'opéron *hisXYZ* et des gènes adjacents. Les gènes de l'opéron sont en bleu, le T-box riboswitch en noir et les gènes adjacents en gris. Les terminateurs prédits par ARNold finding terminator (Naville *et al.*, 2011) sont indiqués par les boucles.

### 3. Riboswitchs putatifs avec effecteur inconnu

#### a. Régulation du gène *glmU* par *ern2210* ?

L'ARN *ern2210* (anciennement sRNA\_2210) est un ARN faiblement exprimé, comportant deux formes [Figure 13] (Siné et al., 2017a). Il est localisé approximativement à une centaine de nucléotides en amont du gène *glmU* [Figure 36]. Une régulation en 5'UTR de type riboswitch peut être envisagée.

```
2696400      ACATTTTTTCATAAATTTGGCTTAATACCGAGAAACTGGCTTTTGAACACCGTTTAT
                                         +1
2696343      CTAGAGGTTGTGAAATCAGCGTCCTGTCCTGAGTATTAGCTCAATAAACGGAAAAAT
                                         +1
2696286      TGCTCCTCATATTTTGCCGTTTATTGAGCTAATACCGAAGGTCAGCTGATTGAAAAA
                                         5'
2696229      CGTTTAATGAAGAGGTTGTAAGCCTACCGTTTAACTCTAAGTCACAAGAAACGATTT
2696172      CAAAAACAGCTTTTCATGTTTTATGAGTAGTGTGACTTGTGACCGAGGAGTTGGCA
2696115      GGAGACCACCGGTTACCAAATAGATAGGCCCGTTATGGTCTATCTATTTTTTTATCC
                                         -10,6 kcal/mol
2696058      AGAAAACGCTGGGATAAAAGAAAATCACTTGTTTTGTGAAAAAATGGGTATACTGG
                                         3'
2695972      GATAAACTGAAATCAGCAAACTATATTTAGGAGTGACGCAGTTGGACGCACGTTATG
                                         RBS
```

Figure 36. Séquence nucléotidique de la potentielle région non traduite en 5' (5'UTR) du gène *glmU*.

Les +1 de transcription prédit par BDGP (Reese, 2001) sont indiqués. Le terminateur rho-indépendant prédit par ARNold finding terminator (Naville et al., 2011) est souligné. Les données 5'3' RACE sont également annotées. La séquence alignée sur plus de la moitié des *loci* des génomes est surlignée en gris.

L'enzyme GlmU est impliquée dans la synthèse d'un précurseur du peptidoglycane, l'UDP-N-acétylglucosamine (GlcNAc) (Anderson et Raetz, 1987; Mengin-Lecreux et van Heijenoort, 1994; Olsen et al., 2007). La transcription du gène *glmU* est régulé par les facteurs de transcription NagC (*E. coli*) (Plumbridge, 1995) ou NagR (*B. subtilis*) (Gaugué et al., 2014). Ces protéines reconnaissent des séquences ADN consensus favorisant alors l'initiation de la transcription du gène régulé. Sur la base de données RegPrecise (Novichkov et al., 2013), des sites de fixation à NagR ont été identifiés à -194 nucléotides





Sur le même modèle que la régulation de *glmS*, *ern2210* pourrait être un ribozyme régulant *glmU* et spécifique de GlcNAc-1P et/ou UDP-GlcNAc. Généralement, le site de catalyse du ribozyme se situe aux alentours de 200 à 300 nucléotides avant le gène *glmS* (Kawada-Matsuo *et al.*, 2017). Pour le gène *glmU*, un ribozyme pourrait expliquer l'observation des deux formes d'*ern2210* [Figure 13]. La présence d'un terminateur rho-indépendant prédit peut suggérer un second mécanisme de régulation, mais ne correspondant pas à l'extrémité 3' déterminé par 5'3'RACE [Tableau 4 ; Figure 36]. On peut également remarquer que la séquence *ern2210* s'aligne (étendue de 53 à 257 nucléotides et avec de 80 à 90 % d'homologie avec BLAST (Altschul *et al.*, 1990)) sur une vingtaine de positions dans le génome d'Aus0004 ainsi que d'autres souches d'*E. faecium* [Tableau 7].

**Tableau 7. Description des résultats obtenus par l'alignement de la région 5'UTR de *glmU*.**

% identité	Etendue (nt)	SNP (nt)	Gap (nt)	5'	3'	Gène en 3'	Même brin ?
100	414	0	0	2696358	2695945	<i>glmU</i>	Oui
86,77	257	31	3	22138	21884	Transposase	Oui
86,72	241	31	1	2882787	2882548	<i>rplM</i>	Oui
83,73	252	40	1	2711382	2711131	<i>rodZ</i>	Oui
84,29	210	30	3	386643	386436	<i>pgsA</i>	Non
81,06	264	34	15	17130	16875	<i>dnaB</i>	Non
85,64	181	25	1	638255	638434	<i>mdoB</i>	Oui
88,32	137	16	0	428659	428795	Hypoth.	Oui
79,52	249	37	13	2159012	2159254	Hypoth.	Non
86,58	149	19	1	386888	386740	<i>pgsA</i>	Non
88,15	135	16	0	682182	682048	<i>rimL</i>	Non
90	120	12	0	2585057	2584938	<i>mdlB</i>	Non
85,40	137	18	2	725693	725827	Hypoth.	Oui
85,40	137	19	1	2363062	2363197	<i>mdlB</i>	Non
85,50	131	17	2	1956836	1956708	Hypoth.	Oui
85,27	129	19	0	2709336	2709208	<i>cinA</i>	Oui
88,07	109	13	0	521689	521797	Hypoth.	Oui
83,82	136	22	0	1280126	1279991	<i>cls</i>	Non
86,33	117	16	0	2128223	2128339	<i>fbp2</i>	Non
81,34	134	21	4	203342	203472	Transposase	Oui
87,14	70	9	0	2762517	2762448	Hypoth.	Oui
88,68	53	4	1	144279	144229	Hypoth.	Oui

### *b. Régulation d'un opéron de transporteur ABC putatif par *ern2410* ?*

L'ARN *ern2410* (anciennement sRNA\_2410) s'exprime également sous deux formes [Figure 13] (Sinel *et al.*, 2017a). Cet ARN est localisé à moins de vingt nucléotides d'un opéron d'ABC-transporteur putatif, suggérant une régulation en 5'UTR de type riboswitch [Figure 38]. On peut supposer un régulateur de type riboswitch ayant comme effecteur la molécule impliquée dans le transport médié par le transporteur ABC régulé.

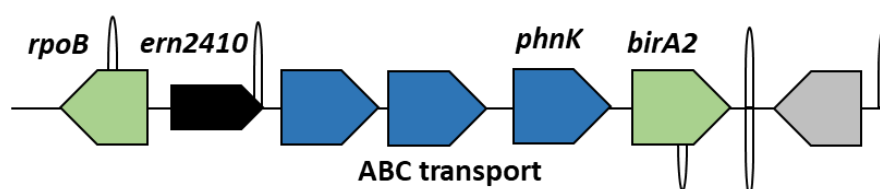


Figure 38. Représentation de l'opéron potentiellement régulé par *ern2410* et des gènes adjacents.

L'opéron comprend également le gène *birA2* codant pour une enzyme impliquée dans le métabolisme de la biotine et également pour un répresseur transcriptionnel (Rodionov *et al.*, 2002). De plus, les transporteurs de la biotine ne sont pas annotés chez *E. faecium*. Il existe des riboswitchs ayant comme molécules effectrices des vitamines, mais pas à la biotine à ce jour. L'ARN *ern2410* peut donc être potentiellement un riboswitch médié par la biotine régulant l'opéron des transporteurs ABC jusqu'au gène *birA2*. La protéine BirA2 pourrait alors inhiber l'initiation de transcription d'*ern2410*. Cet opéron aurait donc potentiellement une double régulation (transcriptionnelle et post-transcriptionnelle).

## C. ARN régulateurs

### 1. Action en *cis*

#### a. Le couple *ern0280-ern0290* comme toxine-antitoxine de type 1

L'ARN *ern0280* (anciennement sRNA\_0280) est le second ARN le plus exprimé des 16 ARN sélectionnés [Tableau 3]. Celui-ci est sur le brin opposé des deux gènes adjacents et à plus de 50 nucléotides du codon stop du gène le plus proche. L'ARN *ern0280* pourrait donc être un ARN régulateur « conventionnel » [Figure 29]. Sur le même *locus* mais brin opposé, il existe un ARN antisens, *ern0290*, très faiblement exprimé par RNAseq (38 RPKM) (Siné et al., 2017a). Ce système d'ARN sens-antisens peut suggérer un système T-AT de type I (Gerdes et Wagner, 2007). Une recherche de peptide a été effectuée par ORFfinder (ncbi) révélant un peptide putatif de 28 AA dans la séquence d'*ern0280* et un autre de 39 AA dans la séquence d'*ern0290* [Figure 39]. Par alignement de séquences peptidiques, le peptide putatif codé par *ern0280* semble appartenir à la famille des Fst, toxines impliquées dans les systèmes T-AT, notamment dans le maintien de plasmides comme pAD1 (Weaver et al., 1996, 2003). Le second peptide, codé par *ern0290*, s'aligne partiellement sur une protéine hypothétique (deux souches d'*E. hirae* et une d'*E. faecium*). L'alignement nucléotidique de la séquence étudiée [Figure 39] montre une grande conservation (100 % d'homologie) au sein de plasmides d'*E. faecium*. Le système *ern0280-ern0290* semble donc être un système T-AT de type I avec *ern0280* codant une toxine Fst-like et *ern0290*, l'ARN antisens complémentaire de la région d'initiation de la traduction de la toxine d'*ern0280* [Figure 39].



**Figure 39. Séquence nucléotidique et peptidique contenant les 2 ARN *ern0280-ern0290*.**  
 L'ARN *ern0280* est transcrit sur le brin sens (5'-3') et l'ARN *ern0290* est transcrit sur le brin antisens (3'-5') comme indiqué à gauche de la figure. Le peptide codé par *ern0280* est indiqué en rouge et celui codé par *ern0290* est indiqué en noir. Les séquences d'initiation de transcription (-35, -10 et +1 prédit par BDGP) et de traduction (RBS et ATG prédit par ORFinder) sont annotées. Le terminateur rho-indépendant prédit par ARNold finding terminator est souligné. Les triangles noir et blanc indiquent respectivement les positions 5' et 3' d'*ern0280* déterminées par 5'3'RACE. La séquence prédite par RNAseq d'*ern0290* est surlignée en gris.

### b. Régulation d'une aldolase (*fsaA*-like) par *ern1260* ?

L'ARN *ern1260* (anciennement sRNA\_1260) est le 6<sup>ème</sup> ARN le plus exprimé de 16 sélectionnés [Tableau 3]. Cet ARN est proche (moins de 20 nt) de la séquence codante de EFAU004\_01516 (codant pour l'enzyme fructose-6-phosphate aldolase FsaA) mais sur le brin opposé [Figure 29]. Une action en *cis* de l'ARN *ern1260* sur la région 3'UTR du gène peut être envisagée. Une action en *trans* semble peu probable à cause de sa stabilité limitée [Figure 31]. Chez *E. coli*, le gène *fsaA* peut être régulé en 3'UTR par le recrutement de RNase médié par des séquences REP (Repetitive Extrinsic Palindrome) (Liang *et al.*, 2015; Liang et Deutscher, 2016). Un système analogue pourrait exister chez *E. faecium*, et l'interaction entre l'ARN régulateur et la région 3'UTR empêcherait la formation de cette structure REP, stabilisant ainsi l'ARNm cible [Figure 40].

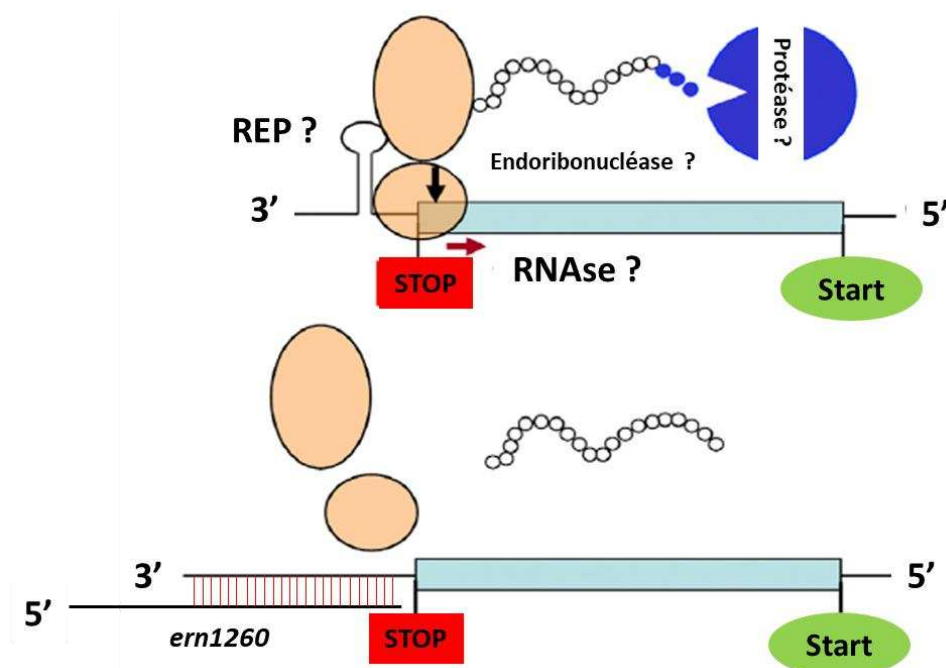


Figure 40. Représentation du mécanisme de régulation hypothétique de *ern1260*. Schéma reprenant le mécanisme de régulation par les domaines REP chez *E. coli* (Liang *et al.*, 2015).

Une autre hypothèse peut être émise impliquant une interaction plus importante entre l'ARN régulateur *ern1260* et l'ARNm de *fsaA*. En effet, l'extrémité 3' déterminée par 5'3'RACE ne correspond pas à un terminateur prédit par ARNold finding terminator (Naville *et al.*, 2011). Cependant, il existe un terminateur prédit sur le brin de l'ARN régulateur, localisé vers le début de la séquence codante de *fsaA*. L'existence d'une seconde forme d'*ern1260* peut être suggérée, qui semblerait être induite par la présence de daptomycine [Figure 41]. La forme plus grande pourrait donc interagir plus fortement avec la séquence de l'ARNm de *fsaA* sur le brin opposé. L'ARN *ern1260* pourrait être un ARN régulateur agissant en *cis* sur le gène *fsaA*.

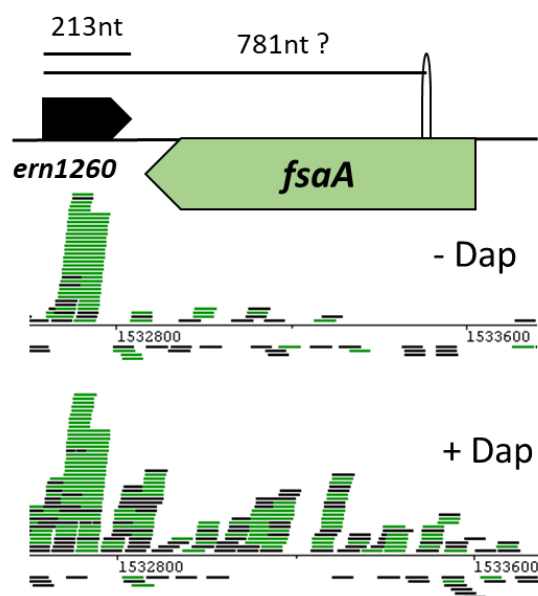


Figure 41. Représentation de l'élongation de *ern1260* induite par la daptomycine vis-à-vis de la position du gène *fsaA*. L'expression (RNAseq) est illustrée en absence (- Dap) ou présence (+ Dap) de daptomycine par Artemis (Carver *et al.*, 2012). Les données proviennent des « raw data » obtenues par C. Sinel *et al.* (Sinel *et al.*, 2017a).

## 2. Action en *trans* d'*ern0120* ?

L'ARN *ern0120* (anciennement sRNA\_0120) est le 4<sup>ème</sup> ARN le plus exprimé [Tableau 3]. Cet ARN est proche du codon stop du gène précédent EFAU004\_00212 (codant une acétyltransférase putative). Les promoteurs prédits par BDGP (Reese, 2001) aux alentours de l'extrémité 5' déterminée par 5'3'RACE ont des scores de confiance inférieurs à 0,5 [Tableau 4]. Ces informations semblent suggérer qu'*ern0120* puisse être inclus dans la région 3'UTR du gène EFAU004\_00212 [Figure 29]. La grande stabilité d'*ern0120* est cohérente avec une potentielle activité en *trans* [Figure 31]. Les prédictions de cibles semblent peu contributives, excepté la régulation de la ribonucléase HIII [Tableau 8]. On peut supposer une « activation » d'*ern0120* après clivage de la région 3'UTR du gène EFAU004\_00212 par la ribonucléase HIII [Figure 42]. L'ARN *ern0120*, alors clivé, pourrait agir sur différentes cibles dont l'ARNm de *rnhC*. L'interaction se situant à environ une centaine de nucléotides en amont de *rnhC*, la régulation devrait être post-transcriptionnelle.

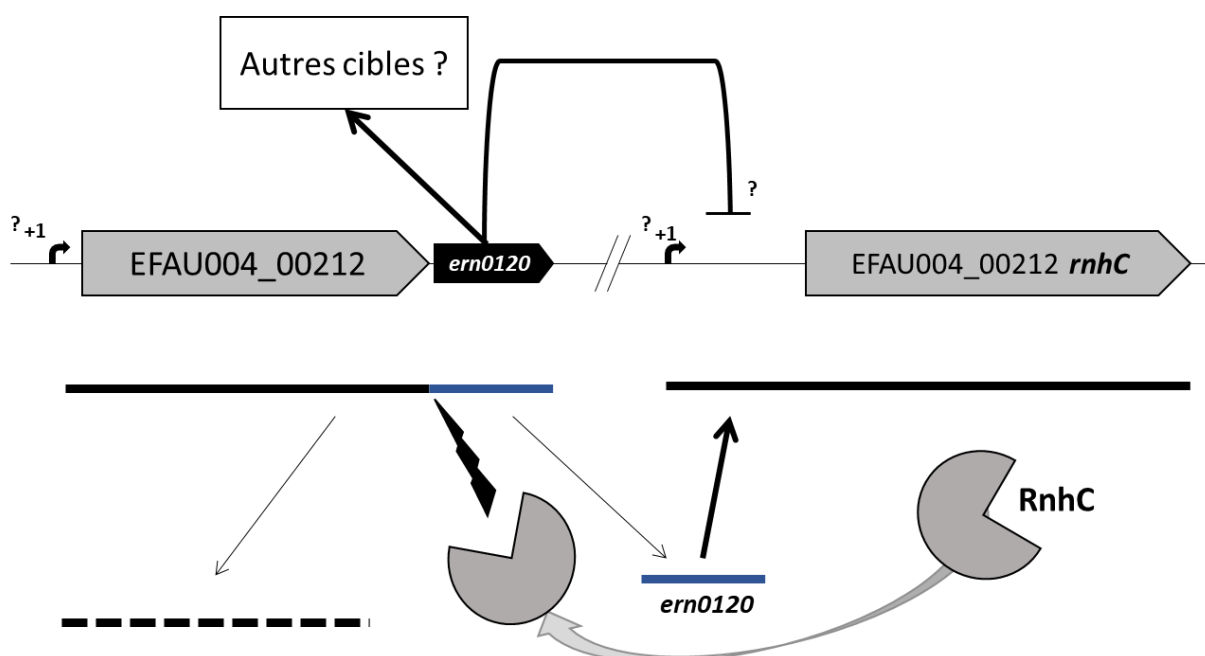


Figure 42. Représentation de l'autorégulation hypothétique entre *ern0120* et *rnhC*.



**Tableau 8. Sélection de cibles prédites par Target RNA2 et copraRNA de *ern0120*.**

La sélection est basée sur les cibles avec une p-value  $\leq 0,001$  (ou les 5 avec la plus petite p-value) ou des cibles en communs (p-value  $\leq 0,05$ ).

Numéro	Gène	COG	Annotation	Target RNA2		CoprRNA		
				Rang	p-value	Rang	p-value	FDR
EFAU004_02111	<i>rnhC</i>	L	ribonuclease HIII	1	<b>&lt;0,001</b>	28	<b>0,007</b>	0,60
EFAU004_01530	<i>tra8</i>	L	Integrase	2	<b>&lt;0,001</b>	-	-	-
EFAU004_01822	-	-	hypothetical protein	3	<b>0,001</b>	-	-	-
EFAU004_01750	<i>yqgQ</i>	S	hypothetical protein	4	<b>0,001</b>	-	-	-
EFAU004_01028	<i>mscS</i>	M	transporter, small conductance mechanosensitive ion channel family protein	5	<b>0,001</b>	-	-	-
EFAU004_00163	<i>lrgA</i>	R	LrgA family protein	6	<b>0,001</b>	-	-	-
EFAU004_00720	<i>nrdG</i>	O	anaerobic ribonucleotide reductase small subunit	-	-	2	<b>&lt;0,0001</b>	0,02
EFAU004_00162	<i>glpF</i>	G	glycerol MIP family major intrinsic protein channel protein	-	-	3	<b>&lt;0,0001</b>	0,04
EFAU004_00213	-	-	hypothetical protein	-	-	4	<b>0,0002</b>	0,11
-	<i>rex2</i>	-	redox-sensing transcriptional regulator	-	-	5	<b>0,0004</b>	0,24
EFAU004_00628	<i>panE</i>	H	2-dehydropantoate 2-reductase	9	<b>0,002</b>	53	0,015	0,72
EFAU004_02474	<i>mgtE</i>	P	MgtE family magnesium (Mg <sup>2+</sup> )/cobalt (Co <sup>2+</sup> ) transporter-E	47	0,028	40	0,011	0,69
EFAU004_01757	<i>folA/dfrD</i>	H	dihydrofolate reductase	61	0,045	99	0,028	0,73

### 3. Action mixte en *cis* et *trans* d'*ern0160* ?

L'ARN *ern0160* (anciennement sRNA\_0160) est l'ARN le plus exprimé et ayant la plus grande variation d'expression en stress antibiotique et antiseptique (Siné et al., 2017a; Dejoies et al., 2021) [Tableau 3]. Cet ARN est sur le brin opposé des deux gènes adjacents et à plus de 50 nucléotides du codon stop du gène le plus proche. Sur le brin opposé, un autre ARN, *ern0170*, est également exprimé (95 RPKM) (Siné et al., 2017a) [Figure 29]. L'absence de peptide prédit dans les deux séquences d'ARN écarte l'hypothèse d'un système T-AT. L'ARN *ern0170* est à moins de 100 nucléotides du codon d'initiation du gène EFAU004\_00235 (*mgtA*) pouvant suggérer une potentielle régulation en 5'UTR. On peut donc imaginer une action en *cis* d'*ern0160* sur *ern0170*, qui aura une conséquence directe sur la transcription du gène *mgtA* comme suggère les prédictions de cibles [Tableau 9].

Une action en *trans* d'*ern0160* est également envisagée due à la relative stabilité de cet ARN [Figure 31].

**Tableau 9. Sélection de cibles prédites par Target RNA2 et copraRNA d'*ern0160*.**

La sélection est basée sur les cibles avec une p-value  $\leq 0,001$  (ou les 5 mieux classées) ou des cibles en communs (p-value  $\leq 0,05$ ).

Numéro	Gène	COG	Annotation	Target RNA2		CoprRNA		
				Rang	p-value	Rang	p-value	FDR
EFAU004_01131	<i>ygaE</i>	S	membrane protein	1	<b>&lt;0,001</b>	-	-	-
EFAU004_01533	<i>eis</i>	R	GNAT family acetyltransferase	2	<b>0,002</b>	-	-	-
EFAU004_01585	<i>galA</i>	G	alpha-galactosidase	3	<b>0,004</b>	-	-	-
EFAU004_00236	<i>ybfA</i>	S	integral membrane protein	4	<b>0,004</b>	-	-	-
EFAU004_02429	<i>tra8</i>	L	Integrase	5	<b>0,004</b>	-	-	-
EFAU004_00235	<i>mgtA</i>	P	P-ATPase superfamily P-type ATPase cation transporter	-	-	1	<b>&lt;0,0001</b>	<b>&lt;0,0001</b>
EFAU004_02084	<i>lplA</i>	H	lipoate--protein ligase A	-	-	2	<b>&lt;0,0001</b>	0,096
EFAU004_02667	<i>padR</i>	K	PadR family transcriptional regulator	-	-	3	<b>0,0004</b>	0,27
EFAU004_02554	<i>rluA/rluD</i>	J	pseudouridylate synthase RluD	-	-	4	<b>0,0005</b>	0,27
EFAU004_01812	<i>srtA/srtC3</i>	M	sortase	-	-	5	<b>0,0005</b>	0,27
EFAU004_01059	<i>lysM</i>	S	LysM family protein	-	-	6	<b>0,0007</b>	0,27
EFAU004_00344	<i>lysR</i>	K	LysR family transcriptional regulator	35	0,025	22	<b>0,0055</b>	0,62
EFAU004_02124	<i>mleA/sfcA</i>	C	malic enzyme, NAD binding domain protein	44	0,032	148	0,042	0,74

## II. Implication de l'ARN *ern0030* dans la régulation de la résistance aux tétracyclines induit par le gène *tet(M)*

### A. Résumé de l'article 1

#### 1. Présentation

Ce travail présente l'investigation fonctionnelle d'un des premiers petits ARN mis en évidence chez *E. faecium* : *ern0030* (Siné et al., 2017a). Cet ARN est exprimé sous trois formes de tailles variables et son expression diminue en fin de croissance [Figure 30]. Après les premières caractérisations et l'analyse génétique, on a remarqué que l'ARN *ern0030* correspondait à la région régulatrice du gène de résistance aux tétracyclines *tet(M)* (5'UTR d'environ 230 nucléotides) décrit par Su et al., 1992. L'expression de *tet(M)* et des gènes en aval impliqués dans le transfert génétique horizontal des transposons conjugatifs de la famille Tn916-like est régie par la présence de tétracycline et de cette région régulatrice (Beabout et al., 2015; Celli et Trieu et Cuot, 1998; Su et al., 1992). Néanmoins, le mécanisme de régulation repose sur une hypothèse d'atténuation transcriptionnelle non entièrement démontrée et nécessite plus d'investigations.

Les différents objectifs de ce travail ont été de :

- Décrire le rôle fonctionnel de l'ARN *ern0030*,
- Caractériser plus précisément *ern0030* (expression et mécanisme de régulation),
- Investiguer le mécanisme d'atténuation transcriptionnelle initialement décrit.

Pour cela, trois approches ont été utilisées : transcriptomique, phénotypique et mécanistique. Des souches d'étude exprimant le gène *tet(M)* et dépourvues d'*ern0030* ont été produites [Article 1, Figure 1A et S1]. Des plasmides permettant de tester le rôle d'*ern0030* vis-à-vis de *tet(M)* et de la résistance aux tétracyclines ont été construits et introduits dans les souches d'étude. L'expression et la demi-vie d'*ern0030* et de l'ARNm de

*tet(M)* ont été déterminées. L'expression génique a également été analysée selon les différentes souches d'étude et la présence de tétracycline. Ces données ont été comparées avec les résultats de tests de sensibilité à la tétracycline et la tigécycline. Enfin, le mécanisme de régulation a été exploré par diverses techniques comme des essais de TOE print pour comprendre le lien existant entre les tétracyclines, *tet(M)* et *ern0030*.

## 2. Principaux résultats

Les principaux résultats de ce travail sont les suivants :

- Description de 3 transcrits (230, 150 et 100 nucléotides) dérivés d'*ern0030*. Les transcrits de plus petite taille sont positionnés en 5' de la région régulatrice de *tet(M)* [Article 1, Figure 1B et 2].
- Exclusion de l'hypothèse qu'*ern0030* régule *tet(M)* par *trans*-activité. Les demi-vies déterminées pour *ern0030* et l'ARNm de *tet(M)* sont respectivement de 4 minutes et 18 secondes et 1 minute et 48 secondes [Article 1, Figure 3]. La *trans*-complémentation d'*ern0030* n'a aucun impact sur la sensibilité aux tétracyclines [Article 1, Table 1].
- Observations phénotypiques inattendues : la délétion totale d'*ern0030* a induit une augmentation de la sensibilité de la souche aux tétracyclines (CMI allant de 32 mg/L à 4 mg/L) [Article 1, Table 1]. L'expression de *tet(M)* dans la souche délétée reste au même niveau que l'expression basale (absence de tétracycline) dans la souche non délétée. L'induction par la tétracycline est perdue dans la souche délétée d'*ern0030* [Article 1, Figure 4].
- Confirmation par Northern-blot et RT-PCR de la co-transcription entre *ern0030* et *tet(M)* induit par la présence de tétracycline [Article 1, Figure 5 et S2].

- Par TOE print, la fixation du ribosome sur le RBS du peptide leader n'a pas pu être détectée [Article 1, Figure 6B], remettant en cause le mécanisme d'atténuation transcriptionnelle.

- Aucun impact majeur sur la sensibilité aux tétracyclines des délétions/mutations des séquences présumées impliquées dans le mécanisme de régulation dans un système plasmidique [Article 1, Table 1]. Ces résultats suggèrent qu'un (ou plusieurs) acteur(s) extérieur(s) au système *ern0030-tet(M)* serait également impliqué dans la régulation.

## B. Article 1

### ***cis*-RNA-mediated expression regulation of the *tet(M)* resistance gene in *Enterococcus faecium***

**Killian Le Neindre<sup>1</sup>, Loren Dejoies<sup>1,2</sup>, Sophie Reissier<sup>1</sup>, François Guerin<sup>2,3</sup>, Brice Felden<sup>1,†</sup>, Vincent Cattoir<sup>1,2,3\*</sup>**

<sup>1</sup> Unité Inserm U1230, Université de Rennes 1, Rennes, France

<sup>2</sup> Service de Bactériologie-Hygiène hospitalière, CHU de Rennes, Rennes, France

<sup>3</sup> CNR de la Résistance aux Antibiotiques (laboratoire associé 'Entérocoques'), Rennes, France

Article soumis au journal « Frontiers in microbiology »

2 ***cis*-RNA-mediated regulation of the *tet*(M) resistance gene**  
3 **expression in *Enterococcus faecium***

4  
5 Running title: *cis*-RNA-mediated *tet*(M) regulation

6 **Killian Le Neindre<sup>1</sup>, Loren Dejoies<sup>1,2</sup>, Sophie Reissier<sup>1</sup>, François Guérin<sup>2,3</sup>, Brice**  
7 **Felden<sup>1†</sup>, Vincent Cattoir<sup>1,2,3\*</sup>**

8 <sup>1</sup>Unité Inserm U1230, Université de Rennes 1, Rennes, France

9 <sup>2</sup>Service de Bactériologie-Hygiène hospitalière, CHU de Rennes, Rennes, France

10 <sup>3</sup>CNR de la Résistance aux Antibiotiques (laboratoire associé 'Entérocoques'), Rennes,  
11 France

12 †Deceased author

13  
14 **\*Correspondence:** Prof. Vincent Cattoir, CHU de Rennes, Service de Bactériologie-Hygiène  
15 hospitalière, 2 rue Henri Le Guilloux, 35033 Rennes Cedex, France. +33-2-99-28-98-28, Fax:  
16 +33-2-99-28-41-59, E-mail: [vincent.cattoir@chu-rennes.fr](mailto:vincent.cattoir@chu-rennes.fr).

17  
18 **Keywords:** *E. faecium*, tetracycline, tigecycline, Tet(M), Tn916, 5'-untranslated region

19  
20 **Word count:** Abstract: 334; Text: 4,033 words; 1 Table; 6 Figures; 46 References; 6  
21 Supplementary data.

## ABSTRACT

A set of putative novel small RNAs was recently identified as expressed in *Enterococcus faecium*, but no biological function has been elucidated so far. Here, we investigated the role of one of them (named *ern0030*) in the regulation of a resistance gene. We revealed that *ern0030* was encoded within the 5' untranslated region of *tet(M)*, a gene conferring tetracycline resistance through ribosomal protection. Although previously described as transcriptional attenuation, the mechanism of regulation of *tet(M)* expression remains poorly characterized and hypothetically relies on the upstream sequence of *tet(M)* (designated as *Ptet*), that comprises predicted secondary structures (especially a rho-independent transcriptional terminator) and a putative peptide leader sequence. By RACE mapping, we accurately determined the boundaries of *ern0030*, which corresponded to *Ptet*. Surprisingly, Northern blot analyses revealed three transcripts of different lengths (ca. 230, 150 and 100 nucleotides) expressed from *Ptet*. Phenotypically, the total deletion of *ern0030* conferred an unexpected decrease in tetracycline MICs that was consistent with transcriptomic data showing no significant *tet(M)* induction under tetracycline SIC in *ern0030*-deleted mutant as opposed to a 10-fold increase of *tet(M)* expression in the wild-type strain. Even though we confirmed that there was a transcriptional read-through of *ern0030* and *tet(M)* promoted by tetracycline, we showed that the deletion of the rho-independent terminator did not significantly affect tetracycline susceptibility. To determine the role of the putative leader peptide, we investigated the ribosomal binding to the predicted ribosome binding sites (RBS) by toeprint assays. Whereas three toeprints were detected for *tet(M)* RBS (including the expected one), none toeprint was observed with the primer targeting the RBS of the putative leader peptide, suggesting that the latter does not exist or is not involved in ribosomal recruitment and regulation of *tet(M)* expression. This conclusion is also supported by the lack of significant effect on tetracycline MICs in the absence of leader peptide translation. Here, we demonstrate that *ern0030* has a role in regulation of *tet(M)* expression. Altogether, these findings support



50 alternative or complementary mechanism(s) of *tet(M)* regulation other than transcriptional  
51 attenuation.

## INTRODUCTION

The worldwide spread of multidrug-resistant (MDR) bacterial pathogens has become a major public health issue and there is an urgent need to find new tools and approaches to control it. The most critical MDR microorganisms have been grouped as “ESKAPE” pathogens that comprise *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* spp. (Rice, 2008; Santajit and Indrawattana, 2016). Among them, *E. faecium* is a leading cause of nosocomial infections and hospital outbreaks, especially vancomycin-resistant *E. faecium* (VREF) isolates for which limited therapeutic options are available (Arias and Murray, 2012; Lebreton et al., 2013; Ayobami et al., 2020; Zhou et al., 2020). The epidemiological success of VREF mainly results from the worldwide dissemination of a specific clonal complex (CC17) comprising clinical isolates that share a common MDR phenotype with high-level resistance to both ampicillin and fluoroquinolones (Willems et al., 2005; Lee et al., 2019). CC17 is part of a human hospital-adapted lineage (clade A1) that emerged from the animal-associated lineage (clade A2) after the introduction of antibiotics, and which differs genetically from the human community-associated lineage (clade B) (Lebreton et al., 2013). Thanks to a remarkable genomic plasticity, *E. faecium* is particularly adept at acquiring new genes through horizontal gene transfer of a large panel of mobile genetic elements (Hegstad et al., 2010; Zhong et al., 2019) while it also exhibits a versatile metabolism allowing it to cope with many environmental stresses (Arias and Murray, 2012; Gaca and Lemos, 2019).

Post-transcriptional control by non-coding RNAs (ncRNAs) enables a rapid and tight regulation to environmental cues, with gene expression directly influenced by metabolic variations (Repoila and Darfeuille, 2009). ncRNAs can be divided into small RNAs (sRNAs) and RNA leaders, the latter including riboswitches and attenuators (Dersch et al., 2017). Regulation of the expression of antibiotic resistance genes by attenuators is widespread,

78 particularly for ribosome-targeting antibiotics such as tetracyclines, aminoglycosides or  
 79 MLS (macrolides-lincosamides-streptogramins) antibiotics (Su et al., 1992; Lovett and  
 80 Rogers, 1996; Vazquez-Laslop et al., 2008; Chancey et al., 2015; Dersch et al., 2017). The  
 81 attenuation mechanism is based on the presence of a short open reading frame (ORF)  
 82 coding for a leader peptide immediately upstream of the resistance gene, which recruits  
 83 translating ribosomes to the regulatory 5'-untranslated region (5'-UTR) element. In the  
 84 absence of antibiotics, there is an inhibition of the resistance gene expression through a  
 85 premature transcriptional terminator (transcriptional attenuation) or the ribosome binding  
 86 site (RBS) masking (translational attenuation) (Lovett and Rogers, 1996; Gollnick and  
 87 Babitzke, 2002). When the antibiotic is present in the cell, there is a RNA leader  
 88 conformational switch due to the ribosome stalling resulting in alternative base-pairing  
 89 patterns that promote an anti-termination conformation or a RBS release, which  
 90 consequently activate the expression of the regulated gene (Lovett and Rogers, 1996;  
 91 Gollnick and Babitzke, 2002).  
 92 Recently, we identified the first set of transcripts expressed from non-annotated regions in  
 93 *E. faecium* using the reference strain Aus0004 (Lam et al., 2012) and provided a list of 61  
 94 sRNA candidates (Siné et al., 2017). Out of them, 10 were experimentally validated by  
 95 Northern blots and by qPCR experiments as expressed along bacterial growth (Siné et al.,  
 96 2017). However, their biological functions as well as their molecular targets and associated  
 97 gene networks remain unknown. Among them, a putative sRNA (initially designated  
 98 sRNA\_0030 and renamed here *ern0030* for *Enterococcus* ribonucleotide acid) was located  
 99 upstream of the *tet(M)* gene, which codes for a ribosomal protection protein involved in  
 100 tetracycline resistance (Burdett, 1991, 1996; Dantley et al., 1998; Brodersen et al., 2000;  
 101 Dönhöfer et al., 2012; Grossman, 2016). This gene is carried by the widespread Tn916-like  
 102 conjugative transposons (Burdett, 1990; Rice, 1998; Roberts and Mullany, 2011; Ciric et  
 103 al., 2013), and its expression is inducible by low concentrations of tetracycline (Su et al.,  
 104 1992; Beabout et al., 2015). The regulatory region upstream of *tet(M)*, known as *Ptet*, has

105 been described as RNA leader sequence encoding a putative peptide leader (Orf12), and it  
106 has been suggested by structure prediction that *tet(M)* regulation might be mediated by  
107 transcriptional attenuation with the hypothetical implication of the availability of rare  
108 aminoacyl-tRNAs (Su et al., 1992; Celli and Trieu-Cuot, 1998). Interestingly, it has also  
109 been demonstrated in *Enterococcus faecalis* that partial deletions of *Ptet* could be  
110 responsible for an upregulation of *tet(M)* expression associated with resistance to  
111 tigecycline, antibiotic usually effective despite Tet(M) production (Beabout et al., 2015).  
112 Since the regulation mechanism of *tet(M)* expression has been poorly studied, the aim of  
113 this study was to molecularly characterize a small *cis*-regulatory RNA (*ern0030*) expressed  
114 within the 5' UTR of *tet(M)* and to mechanistically investigate its role in regulation of  
115 *tet(M)* expression and associated tetracycline resistance.

## MATERIALS AND METHODS

### Bacterial strains, plasmids and growth conditions

The reference strain *E. faecium* Aus0004 (GenBank accession number CP003351.1) and isogenic mutants (Supplementary Table S1) were grown under ambient air at 37°C in brain heart infusion (BHI; Oxoid) broth or on BHI agar (Oxoid). Strains of *Escherichia coli* (Table S1) were grown under ambient air and agitation (160 rpm) at 37°C in Luria-Bertani (LB; ThermoFischer Scientific) medium or on LB agar (ThermoFischer Scientific). For some experiments, these media were supplemented with spectinomycin (Sigma) and/or tetracycline (Sigma) (see below). The plasmids pAT29 and pWS3 (Table S2) were used for cloning and gene disruption, respectively.

### Antimicrobial susceptibility testing

Minimal inhibitory concentration (MICs) of tetracycline and tigecycline were determined by the microdilution reference method in Mueller-Hinton (MH; ThermoFischer Scientific) broth in accordance with EUCAST guidelines ([www.eucast.org](http://www.eucast.org)). The sub-inhibitory concentration (SIC) of tetracycline was determined in BHI broth and was defined as the highest antibiotic concentration that did not alter significantly the bacterial growth kinetics.

### Construction of deleted mutants and *trans*-complemented strains

The restoration of a functional *tet(M)* gene (named *E. faecium* Aus0004<sup>TR</sup>) was obtained from *E. faecium* Aus0004 (as represented in Figure 1A) by homologous recombination and excision using the pWS3 suicide vector and specific primers, as previously described (Zhang et al., 2011) (Supplementary Tables S1-S3). Using the same approach, the entire *ern0030* gene (from 5' to 3' position; see Figure 1B) was deleted in *E. faecium* Aus0004<sup>TR</sup> and the corresponding deleted mutant was named  $\Delta$ *ern0030*-Aus0004<sup>TR</sup> (Supplementary Tables S1-S3). Both deletions were confirmed by PCR and sequencing using specific primers (Supplementary Table S3).

*Trans*-complemented strains were constructed by cloning the *ern0030* gene (with its native promoter and rho-independent terminator) in the pAT29 shuttle vector (Trieu-Cuot et al.,

1990) using specific primers (Supplementary Tables S1-S3). The recombinant plasmid was transformed by electroporation into *E. coli* EC1000 then into both *E. faecium* Aus0004<sup>TR</sup> and  $\Delta ern0030$ -Aus0004<sup>TR</sup> strains (Supplementary Table S1). For investigation of the *ern0030* function, different recombinant plasmids were generated using the pAT29 vector in which was cloned the functional *tet(M)* gene comprising the wild-type sequence of *ern0030* or several mutated *ern0030* sequences, and were introduced into *E. coli* EC1000 strain and then into *E. faecium* Aus0004 (Supplementary Tables S1-S3). Transformants were selected on agar plates containing spectinomycin at the concentration of 100 mg/L (*E. coli*) or 300 mg/L (*E. faecium*). Inserts of recombinant plasmids were sequenced using specific primers (Supplementary Table S3).

### 163 RNA extraction

164 Total RNAs were isolated at three different time points during growth: mid exponential (ME, OD 1), late exponential (LE, OD 2.2) and early stationary (ES, OD 3) phase. In some experiments, tetracycline was added in the growth medium at the SIC (equivalent to 1/32 MIC). Extractions were performed as previously described (Bronsard et al., 2017). Briefly, cell pellets were dissolved into 500  $\mu$ L of lysis buffer, and cells were broken by acid-treated glass beads and phenol (pH 4) using a FastPrep-24 cell disruptor (MP biomedical). Total RNAs were extracted with phenol/chloroform and ethanol precipitated overnight at -20°C with sodium acetate. Residual chromosomal DNA was removed by treating all total RNA samples with the Turbo DNA-free kit (Life Technologies).

### 163 RACE mapping

164 RACE mapping was performed using bacterial cells grown in BHI broth and collected at ME growth phase, as previously described (Pinel-Marie et al., 2014). Briefly, Total RNAs (5  $\mu$ g) from *E. faecium* Aus0004 were circularized using T4 RNA Ligase (Promega). After precipitation, the reaction was reverse transcribed by the M-MLV RT (Promega) using primer sRNA0030R1 and then two PCR reactions were performed with Taq polymerase (Invitrogen) using sRNA0030R2-F1 and sRNA0030R3-F2 primers (Supplementary Table S3).

The PCR products were cloned in pGEM-T vector (Promega) and recombinant plasmids were transformed into *E. coli* XL1blue (Supplementary Table S1), and sequenced with M13 reverse vector primer (Supplementary Table S3).

### **Northern Blot**

Northern blots were done using total RNAs extracted at the three different time points during growth (see above). RNA samples (5 to 20 µg) were loaded on denaturing 8% polyacrylamide/8M urea gels for separation of smaller RNA transcripts (100-600 nt) or denaturing 0.8% agarose gel for separation of larger RNA transcripts (0.2-3 kb), and transferred onto membrane Hybond N+ (GE Healthcare). Membranes were hybridized with specific  $\gamma^{32}\text{P}$  5' end-labeled probes (Supplementary Table S3) in ExpressHyb solution (Clontech), washed, exposed, and scanned with a Typhoon FLA 9500 scanner (GE Healthcare). tmRNA levels were used as internal loading controls.

### **RT-PCR and RT-qPCR**

For RT-PCR, DNase-treated RNAs isolated from *E. faecium* Aus0004 grown at the ME phase (OD 1) in BHI broth supplemented with tetracycline SIC were used. cDNA was synthesized using the High Capacity cDNA reverse transcription kit (ThermoFischer Scientific) with qPCRTETMR or qPCRsRNA0030R primers (Supplementary Table S3). PCRs were then carried out under standard conditions using specific primers (Supplementary Table S3). Each PCR amplification was performed on cDNA and chromosomal DNA (used as positive control).

From 40 ng of DNase-treated total RNA samples, RT-qPCR was performed using the Power SYBR Green RNA-to-CT 1-Step kit (ThermoFischer Scientific) and a StepOnePlus Real-Time PCR system (Applied Biosystems). Transcript levels (expressed as copy numbers per ng of RNA) in each sample were determined by absolute quantification using serial dilutions of PCR products. Expression of the *adk* gene as a housekeeping control gene was also measured. Each experiment was performed in triplicate. For statistical analysis, two-way ANOVA and Tukey tests were employed using the GraphPad Prism 7.0 software.

### **Half-life determination**

*E. faecium* Aus0004<sup>TR</sup> cells were cultured overnight, diluted 1:50, then grown to ME phase (OD 1) at 37°C in BHI broth supplemented or not with tetracycline SIC, and then incubated with 150 mg/L of rifampicin. Total RNAs were extracted at different time intervals and transcript levels of *ern0030* and *tet(M)* were quantified by RT-qPCR with specific primers (Supplementary Table S3). Half-lives of both RNAs were calculated by non-linear regression curve fitting (one phase decay) using GraphPad Prism 7.0 software as previously described (Ratnadiwakara and Änkö, 2018).

#### **Toeprints assays**

*In vitro* transcription was conducted from PCR-amplified templates using Aus0004 genomic DNA and the MEGAscript T7 kit (Ambion) with specific primers (Supplementary Table S3). RNAs were gel-purified, eluted and precipitated with ethanol in the presence of 0.3M sodium acetate. Toeprint assays were performed as previously described (Riffaud et al., 2019) with some minor modifications. Briefly, annealing mixtures containing 5 pmol 5'UTR-*tet(M)* RNA with 1 pmol <sup>γ</sup>32P 5' end-labeled primers (Supplementary Table S3) were incubated in a buffer (20 mM Tris-HCl, pH7.5, 60 mM NH<sub>4</sub>Cl) for 1 min at 90°C. Renaturation was done in the presence of 10 mM MgCl<sub>2</sub> for 20 min at 20°C. 70S ribosomes were reactivated for 15 min at 37°C, diluted in the reaction buffer and incubated again for 15 min at 37°C. Various concentrations (0 to 1.9 pmol) of purified 70S ribosomes were added to each sample, and these were incubated for 5 min at 37°C, after which MgCl<sub>2</sub> was adjusted to 10 mM. After 15 min at 37°C, an amount of 10 pmol of uncharged tRNA<sup>fMet</sup> was added and incubated for 5 min at 37°C. The cDNA were synthesized with 10 U of AMV RT (New England Biolabs) for 20 min at 42°C. Reactions were stopped by the addition of 15 µl of loading buffer II (Ambion), and cDNAs were loaded and separated onto 8% polyacrylamide/8M urea gels. Sequencing ladders were generated with the same 5' end-labeled primer for toeprint.



## RESULTS AND DISCUSSION

### Characterization of *ern0030* transcripts in *E. faecium* Aus0004

In a previous study, we identified a set of novel transcripts in the ESKAPE pathogen *E. faecium* (Sinel et al., 2017). Among them, *ern0030* was found to be expressed upstream from the initiation codon of the *tet(M)* tetracycline resistance gene. In *E. faecium* Aus0004, *tet(M)* was chromosomally encoded but interrupted by a large insertion sequence (from EFAU004\_00064 to EFAU004\_00066) containing the *dfrG* trimethoprim resistance gene (Figure 1A) related to Tn5801 ‘type A5’ transposons (León-Sampedro et al., 2016). Due to an overlap between *ern0030* and *Ptet*, we hypothesized a functional relationship between *ern0030* and *tet(M)*. Therefore, we repaired the *tet(M)* gene (named *E. faecium* Aus0004<sup>TR</sup>) in *E. faecium* Aus0004 and confirmed its functionality by Northern blot (Supplementary Figure S1) and determination of tetracycline MICs (Table 1).

By RACE mapping, the 5’ end of *ern0030* was determined to be located at the genomic position 65144 on *E. faecium* Aus0004 chromosome (T<sub>1</sub>; Figure 1B). This position corresponds to the transcription initiation site (TSS) previously reported for *tet(M)* (Su et al., 1992). Its 3’ end corresponded to a predicted rho-independent terminator (A<sub>233</sub>; Figure 1B). This revealed that *ern0030* corresponded to *Ptet*, the RNA leader sequence upstream *tet(M)* previously reported (Su et al., 1992; Celli and Trieu-Cuot, 1998). Surprisingly, by using the same probe previously published to reveal *ern0030* (Sinel et al., 2017), we identified three transcripts of different lengths (ca. 230, 150 and 100 nucleotides; Figure 2) expressed from *Ptet*, which has not been previously described (Su et al., 1992; Sinel et al., 2017). Using other *ern0030*-specific probes, we demonstrated that the two shorter forms were located in the 5’ end of the *Ptet* sequence between positions G<sub>29</sub> and C<sub>133</sub> (Figures 1B and 2), suggesting RNA processing or premature transcription termination. Note that Su et al. were not able to detect these shorter transcripts since they used a probe targeting the 3’ moiety of the *Ptet* sequence (Su et al., 1992). Interestingly, the transcript level of these three different transcripts significantly changed during to the bacterial

growth, with high-level expression of 230- and 150-nt transcripts and weak expression of the 100-nt transcript at the ME phase whereas an opposite pattern was observed at the ES phase (Figure 2). The high-level expression of the 100-nt transcript at the ES phase might suggest an accumulation due to stability change or a processing from a longer transcript. To complete *ern0030* characterization, the half-life of this RNA was measured in *E. faecium* Aus0004<sup>TR</sup> in the absence of tetracycline. The half-life of *ern0030* RNA (mean = 4 minutes and 18 seconds) was slightly higher than that of *tet(M)* (mean = 1 minute and 48 seconds) (Figure 3). The *ern0030* half-life is shorter than those of conventional *trans*-acting sRNAs while the very short half-life of *tet(M)* mRNA is consistent with a dynamic transcriptional adaptation. Half-life measurement with a SIC of tetracycline was not calculable since transcript levels did not differ between different time points. This may be related to the fact that tetracycline SIC may result in a global mRNA stabilization, as previously described in *Bacillus subtilis* (Wei and Bechhofer, 2002).

#### **Role of *ern0030* in *tet(M)* expression**

We firstly investigated the functional link between *ern0030*, *tet(M)* and tetracycline resistance by using the *ern0030*-deleted Aus0004<sup>TR</sup> mutant. Phenotypically, the *ern0030* deletion caused an eight-fold decrease in MICs of tetracycline (from 32 to 4 mg/L) while no change was observed for tigecycline (Table 1). This observation was consistent with variations of *tet(M)* expression between the two strains with a ~10-fold increase of *tet(M)* expression in Aus0004<sup>TR</sup> and no transcriptional response in  $\Delta$ *ern0030*-Aus0004<sup>TR</sup> (Figure 4A). Noteworthy, tetracycline did not affect *ern0030* expression in Aus0004<sup>TR</sup> (Figure 4B). This decrease in tetracycline MICs is quite surprising since bioreactor-adapted *E. faecalis* mutants with partial deletions in *Ptet* were higher levels of tetracycline resistance, especially towards tigecycline (Beabout et al., 2015). In the latter study, two types of chromosomal mutants were obtained, one with a partial deletion of 3' end of *Ptet* (named BTR87b) and another with a partial deletion of 5' end of *Ptet* (named BTR22). It was also demonstrated that both BTR87b and BTR22 strains had a greater than 10-fold increase in

*tet(M)* expression in the presence of tigecycline, which is contradictory with our results. Then, the complete deletion of *ern0030* seems to confer a phenotype different from those obtained by partial deletions, suggesting a complex mechanism of regulation.

*Trans*-complementation experiments were performed by expressing pAT29Ω*ern0030* in Aus0004<sup>TR</sup> and Δ*ern0030*-Aus0004<sup>TR</sup> strains (Supplementary Tables S1 and S2). No difference in MICs of tetracycline and tigecycline was observed between *trans*-complemented strains and those harboring the pAT29 empty vector (Table 1), demonstrating that *ern0030* did not act as a *trans*-acting sRNA.

### Investigation of transcriptional attenuation mechanism

The upstream sequence of the *tet(M)* gene, *Ptet*, was defined as a transcriptional attenuator only based on predicted secondary structures and the presence a putative peptide leader ORF (*orf*<sub>12</sub>) (Su et al., 1992; Celli and Trieu-Cuot, 1998). This mechanism hypothetically relies on peptide leader translation and ribosome stalling that may cause an alternative structure of the rho-independent transcriptional terminator located at 3' end of *ern0030* (Figure 1B). The role of these two elements (i.e. rho-independent transcriptional terminator and putative peptide leader ORF) was then investigated.

First, we verified that there was a transcriptional read-through of *ern0030* and *tet(M)* promoted by tetracycline as previously described (Su et al., 1992; Celli and Trieu-Cuot, 1998). For this, we used Aus0004<sup>TR</sup> strain and its isogenic mutant Δ*ern0030*-Aus0004<sup>TR</sup> and we confirmed the read-through after tetracycline addition (Figure 5). The co-transcription of *ern0030* and *tet(M)* was also observed by RT-PCR (Supplementary Figure S2). The role of the rho-independent transcriptional terminator was investigated by studying the Aus0004\_pAT29Ω*tet(M)*-Δ<sub>204-238</sub>*ern0030* strain that carried pAT29Ω*tet(M)* (that contained both *ern0030* and *tet(M)* genes) with a 35-nt deletion in *ern0030* comprising the rho-independent terminator (Supplementary Tables S1 and S2). MICs of tetracycline and tigecycline were similar (only a one-dilution difference) between the two strains (Table 1), suggesting a minor effect of the rho-independent transcriptional terminator.

304 The second important element in transcriptional attenuation mechanism is the ribosomal  
305 binding to the 5' UTR and the peptide leader translation. Based on structural predictions,  
306 it has been suggested that the RNA leader termination conformation was induced by the  
307 ribosome stalling during peptide leader translation (Su et al., 1992). According to this  
308 hypothesis, under non-induced conditions, the translation of the leader peptide is delayed  
309 since it is composed of codons corresponding to rare aminoacyl-tRNAs, and then the  
310 expression of Tet(M) is limited only to what can be generated as a result of transcriptional  
311 read-through of the terminator. When bacterial cells are exposed to tetracycline SIC,  
312 which delays translation, there is an increased availability of those rare aminoacyl-tRNAs,  
313 which allows leader peptide translation to more easily keep up with transcription, disrupt  
314 the terminator stem-loop base pairing and enable continuing *tet(M)* transcription. To  
315 determine whether the peptide leader had a function in regulation of *tet(M)* expression,  
316 we investigated the ribosomal loading onto the predicted peptide leader RBS by toeprint  
317 assays. This *in vitro* assay relies on the inhibition of reverse transcription due to stalled  
318 ribosome onto the initiation codon. Recognition of a RBS sequence by ribosomes induces a  
319 toeprint during primer extension that usually occurs 16-19 nucleotides downstream of the  
320 initiation codon (Nilsen, 2013). We used the two different radiolabeled primers TETMPRINT  
321 and RBSTETM-R targeting the RBS of *tet(M)* and that of the putative leader peptide (Su et  
322 al., 1992), respectively (Figure 1B). With the TETMPRINT primer, three different toeprints  
323 (A<sub>300</sub>, G<sub>292</sub> and U<sub>260</sub>) were detected (Figure 6A). The position G<sub>292</sub> corresponds to the  
324 expected *tet(M)* RBS (Figure 1B). Being close to G<sub>292</sub>, A<sub>300</sub> may be the result of structural  
325 changes due to ribosome binding. For U<sub>260</sub>, it is located ~15-20 nucleotides downstream of  
326 an AUG that has a putative RBS (GGGA) suggesting that another leader peptide might exist  
327 (Figure 1B). Note that the stop at the position A<sub>233</sub> was independently observed regardless  
328 of the ribosome concentration and corresponded to the end of the stem-loop of the  
329 *ern0030* rho-independent transcription terminator, which is undoubtedly a very robust  
330 structure (Figure 6A).

331 The research of the putative peptide leader RBS was also performed with RBSTETM-R  
332 primer. According to Su et al. (Su et al., 1992), the expected theoretical ribosome toeprint  
333 should have been detected between positions U<sub>191</sub> and U<sub>194</sub> (Figure 1B). However, none  
334 toeprint was observed suggesting that this putative leader peptide does not exist or is not  
335 involved in ribosomal recruitment and regulation of *tet(M)* expression (Figure 6B). The role  
336 of this putative peptide leader was also challenged by studying the  
337 Aus0004\_pAT29Ω*tet(M)*-TGA<sub>172-174</sub> strain that carried pAT29Ω*tet(M)* (that contained both  
338 *ern0030* and *tet(M)* genes) with an unique mutation (ATG → TGA) in the translation  
339 initiation codon of the putative peptide leader (Supplementary Tables S1 and S2). Despite  
340 the absence of leader peptide translation, tetracycline MICs were similar (only one-dilution  
341 difference) between Aus0004\_pAT29Ω*tet(M)* and Aus0004\_pAT29Ω*tet(M)*-TGA<sub>172-174</sub> strains  
342 (Table 1), indicating that this putative leader peptide (if it does exist) has no major role in  
343 tetracycline-mediated regulation of *tet(M)* regulation. Finally, it has been demonstrated  
344 that tetracycline could interact with double-stranded RNAs (Chukwudi and Good, 2016) and  
345 hypothetically might interfere with ribosomal interaction. This ability to alter ribosome  
346 binding onto the mRNA was also evaluated by toeprint assays, but we did not detect  
347 differences regardless of tetracycline concentration (Supplementary Figure S3). Another  
348 argument suggesting another regulation mechanism is that it has been previously  
349 demonstrated that a partial deletion of 5' *Ptet* (mutant BTR22), which does not include  
350 the peptide leader *orf12*, is more resistant to tetracyclines with upregulation of *tet(M)*  
351 (Beabout et al., 2015).

## CONCLUSION

Here, we have for the first time investigated the role of a regulatory RNA (*ern0030*) expressed in *E. faecium*. We have demonstrated that *ern0030* had a role in regulation of *tet(M)* expression. Even though further investigations are needed, results obtained here indicate the existence of alternative or complementary mechanism(s) of *tet(M)* regulation other than transcriptional attenuation. The complexity of regulation *tet(M)* expression has also been pointed out since it has been demonstrated that the transcription of *tet(M)* from Tn916 was induced by SICs of other antibiotics (such as spectinomycin and MLS) or biocides (such as ethanol) (Scornec et al., 2017; Seier-Petersen et al., 2014). Besides its role in the regulation of the *tet(M)* resistance gene, *ern0030* may also be involved in the control of the transfer of Tn916 and then the dissemination of antibiotic resistance genes (Roberts and Mullany, 2011; Beabout et al., 2015).

## 365 **AUTHOR CONTRIBUTIONS**

366 KLN, BF and VC conceived and designed the experiments; KLN, LD, SR and FG performed  
367 the experiments; KLN, FG, BF, and VC analyzed the data; FG, BF and VC contributed with  
368 reagents/materials/analysis tools; KLN, BF and VC wrote the manuscript. All the authors  
369 discussed the results and commented on the manuscript.

370

## 371 **FUNDING**

372 This work was supported by the the “Institut National de la Santé et de la Recherche  
373 Médicale” and Rennes 1 University.

374

## 375 **ACKNOWLEDGMENTS**

376 The authors thank Yoann Augagneur for his critical reading of the manuscript.

## REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* 215, 403-410. doi:10.1016/S0022-2836(05)80360-2.
- Arias, C. A., and Murray, B. E. (2012). The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat Rev Microbiol* 10, 266-278. doi:10.1038/nrmicro2761.
- Ayobami, O., Willrich, N., Reuss, A., Eckmanns, T., and Markwart, R. (2020). The ongoing challenge of vancomycin-resistant *Enterococcus faecium* and *Enterococcus faecalis* in Europe: an epidemiological analysis of bloodstream infections. *Emerg Microbes Infect* 9, 1180-1193. doi:10.1080/22221751.2020.1769500.
- Beabout, K., Hammerstrom, T. G., Wang, T. T., Bhatt, M., Christie, P. J., Saxer, G., et al. (2015). Rampant parasexuality evolves in a hospital pathogen during antibiotic selection. *Molecular Biology and Evolution* 32, 2585. doi:10.1093/molbev/msv133.
- Brodersen, D. E., Clemons, W. M., Carter, A. P., Morgan-Warren, R. J., Wimberly, B. T., and Ramakrishnan, V. (2000). The Structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S Ribosomal Subunit. *Cell* 103, 1143-1154. doi:10.1016/S0092-8674(00)00216-6.
- Bronsard, J., Pascreau, G., Sassi, M., Mauro, T., Augagneur, Y., and Felden, B. (2017). sRNA and cis-antisense sRNA identification in *Staphylococcus aureus* highlights an unusual sRNA gene cluster with one encoding a secreted peptide. *Sci Rep* 7, 4565. doi:10.1038/s41598-017-04786-3.
- Burdett, V. (1990). Nucleotide sequence of the *tet(M)* gene of Tn916. *Nucleic Acids Res* 18, 6137.
- Burdett, V. (1991). Purification and characterization of Tet(M), a protein that renders ribosomes resistant to tetracycline. *J. Biol. Chem.* 266, 2872-2877.
- Burdett, V. (1996). Tet(M)-promoted release of tetracycline from ribosomes is GTP dependent. *J Bacteriol* 178, 3246-3251.
- Celli, J., and Trieu-Cuot, P. (1998). Circularization of Tn916 is required for expression of the transposon-encoded transfer functions: characterization of long tetracycline-inducible transcripts reading through the attachment site. *Molecular Microbiology* 28, 103-117. doi:10.1046/j.1365-2958.1998.00778.x.
- Chancey, S. T., Bai, X., Kumar, N., Drabek, E. F., Daugherty, S. C., Colon, T., et al. (2015). Transcriptional attenuation controls macrolide inducible efflux and resistance in *Streptococcus pneumoniae* and in other gram-positive bacteria containing *mef/mel(msr(D))* elements. *PLoS ONE* 10. doi:10.1371/journal.pone.0116254.



412 Chukwudi, C. U., and Good, L. (2016). Interaction of the tetracyclines with double-  
 413 stranded RNAs of random base sequence: new perspectives on the target and  
 414 mechanism of action. *J. Antibiot.* 69, 622-630. doi:10.1038/ja.2015.145.

415 Ciric, L., Jasni, A., Vries, L. E. de, Agersø, Y., Mullany, P., and Roberts, A. P. (2013). *The*  
 416 *Tn916/Tn1545 Family of Conjugative Transposons*. Landes Bioscience.

417 Dantley, K. A., Dannelly, H. K., and Burdett, V. (1998). Binding interaction between Tet(M)  
 418 and the ribosome: requirements for binding. *J Bacteriol* 180, 4089-4092.

419 Dersch, P., Khan, M. A., Mühlen, S., and Görke, B. (2017). Roles of regulatory RNAs for  
 420 antibiotic resistance in bacteria and their potential value as novel drug targets.  
 421 *Frontiers in Microbiology* 8. doi:10.3389/fmicb.2017.00803.

422 Dönhöfer, A., Franckenberg, S., Wickles, S., Berninghausen, O., Beckmann, R., and Wilson,  
 423 D. N. (2012). Structural basis for TetM-mediated tetracycline resistance. *Proc Natl*  
 424 *Acad Sci U S A* 109, 16900-16905. doi:10.1073/pnas.1208037109.

425 Gaca, A. O., and Lemos, J. A. (2019). Adaptation to adversity: the intermingling of stress  
 426 tolerance and pathogenesis in Enterococci. *Microbiol Mol Biol Rev* 83.  
 427 doi:10.1128/MMBR.00008-19.

428 Gollnick, P., and Babitzke, P. (2002). Transcription attenuation. *Biochimica et Biophysica*  
 429 *Acta (BBA) - Gene Structure and Expression* 1577, 240-250. doi:10.1016/S0167-  
 430 4781(02)00455-4.

431 Grossman, T. H. (2016). Tetracycline antibiotics and resistance. *Cold Spring Harbor*  
 432 *Perspectives in Medicine* 6. doi:10.1101/cshperspect.a025387.

433 Hegstad, K., Mikalsen, T., Coque, T. M., Werner, G., and Sundsfjord, A. (2010). Mobile  
 434 genetic elements and their contribution to the emergence of antimicrobial resistant  
 435 *Enterococcus faecalis* and *Enterococcus faecium*. *Clinical Microbiology and Infection*  
 436 16, 541-554. doi:10.1111/j.1469-0691.2010.03226.x.

437 Lam, M. M. C., Seemann, T., Bulach, D. M., Gladman, S. L., Chen, H., Haring, V., et al.  
 438 (2012). Comparative analysis of the first complete *Enterococcus faecium* genome. *J.*  
 439 *Bacteriol.* 194, 2334-2341. doi:10.1128/JB.00259-12.

440 Lebreton, F., Schaik, W. van, McGuire, A. M., Godfrey, P., Griggs, A., Mazumdar, V., et al.  
 441 (2013). Emergence of epidemic multidrug-resistant *Enterococcus faecium* from animal  
 442 and commensal strains. *mBio* 4, e00534-13. doi:10.1128/mBio.00534-13.

443 Lee, T., Pang, S., Abraham, S., and Coombs, G. W. (2019). Antimicrobial-resistant CC17  
 444 *Enterococcus faecium*: The past, the present and the future. *J Glob Antimicrob Resist*  
 445 16. doi:10.1016/j.jgar.2018.08.016.

446 León-Sampedro, R., Novais, C., Peixe, L., Baquero, F., and Coque, T. M. (2016). Diversity  
 447 and evolution of the Tn5801-tet(M)-Like Integrative and Conjugative Elements among

*Enterococcus*, *Streptococcus*, and *Staphylococcus*. *Antimicrob. Agents Chemother.* 60, 1736-1746. doi:10.1128/AAC.01864-15.

Lovett, P. S., and Rogers, E. J. (1996). Ribosome regulation by the nascent peptide. *Microbiol Rev* 60, 366-385.

Naville, M., Ghuillot-Gaudeffroy, A., Marchais, A., and Gautheret, D. (2011). ARNold: a web tool for the prediction of Rho-independent transcription terminators. *RNA Biol* 8, 11-13. doi:10.4161/rna.8.1.13346.

Nielsen, T. W. (2013). Toeprinting. *Cold Spring Harb Protoc* 2013, 896-899.

Pinel-Marie, M.-L., Brielle, R., and Felden, B. (2014). Dual toxic-peptide-coding *Staphylococcus aureus* RNA under antisense regulation targets host cells and bacterial rivals unequally. *Cell Reports* 7, 424-435. doi:10.1016/j.celrep.2014.03.012.

Ratnadiwakara, M., and Änkö, M.-L. (2018). mRNA stability assay using transcription inhibition by Actinomycin D in mouse pluripotent stem cells. *Bio-protocol* 8, e3072-e3072.

Repoila, F., and Darfeuille, F. (2009). Small regulatory non-coding RNAs in bacteria: physiology and mechanistic aspects. *Biology of the Cell* 101, 117-131. doi:https://doi.org/10.1042/BC20070137.

Rice, L. B. (1998). Tn916 Family Conjugative Transposons and Dissemination of Antimicrobial Resistance Determinants. *Antimicrob Agents Chemother* 42, 1871-1877.

Rice, L. B. (2008). Federal funding for the study of antimicrobial resistance in nosocomial pathogens: No ESKAPE. *J Infect Dis.* 197, 1079-1081. doi:10.1086/533452.

Riffaud, C., Pinel-Marie, M.-L., Pascreau, G., and Felden, B. (2019). Functionality and cross-regulation of the four SprG/SprF type I toxin-antitoxin systems in *Staphylococcus aureus*. *Nucleic Acids Res* 47, 1740-1758. doi:10.1093/nar/gky1256.

Roberts, A. P., and Mullany, P. (2011). Tn916-like genetic elements: a diverse group of modular mobile elements conferring antibiotic resistance. *FEMS Microbiol Rev* 35, 856-871. doi:10.1111/j.1574-6976.2011.00283.x.

Santajit, S., and Indrawattana, N. (2016). Mechanisms of antimicrobial resistance in ESKAPE pathogens. *BioMed Research International* 2016. doi:10.1155/2016/2475067.

Scornec, H., Bellanger, X., Guilloteau, H., Groshenry, G., and Merlin, C. (2017). Inducibility of Tn916 conjugative transfer in *Enterococcus faecalis* by subinhibitory concentrations of ribosome-targeting antibiotics. *J Antimicrob Chemother* 72, 2722-2728. doi:10.1093/jac/dkx202.

Seier-Petersen, M., Jasni, A., Aarestrup, F. M., Vigre, H., Mullany, P., Roberts, A. P. et al. (2014). Effects of subinhibitory concentrations of four commonly used biocides on the conjugative transfer of Tn916 in *Bacillus subtilis*. *J Antimicrob Chemother* 69, 343-348.

- Sinel, C., Augagneur, Y., Sassi, M., Bronsard, J., Cacaci, M., Guérin, F., et al. (2017). Small RNAs in vancomycin-resistant *Enterococcus faecium* involved in daptomycin response and resistance. *Scientific Reports* 7. doi:10.1038/s41598-017-11265-2.
- Su, Y. A., He, P., and Clewell, D. B. (1992). Characterization of the *tet*(M) determinant of Tn916: evidence for regulation by transcription attenuation. *Antimicrobial Agents and Chemotherapy* 36, 769.
- Trieu-Cuot, P., Carlier, C., Poyart-Salmeron, C., and Courvalin, P. (1990). A pair of mobilizable shuttle vectors conferring resistance to spectinomycin for molecular cloning in *Escherichia coli* and in gram-positive bacteria. *Nucleic Acids Research* 18, 4296.
- Vazquez-Laslop, N., Thum, C., and Mankin, A. S. (2008). Molecular mechanism of drug-dependent ribosome stalling. *Molecular Cell* 30, 190-202. doi:10.1016/j.molcel.2008.02.026.
- Wei, Y., and Bechhofer, D. H. (2002). Tetracycline induces stabilization of mRNA in *Bacillus subtilis*. *Journal of Bacteriology* 184, 889-894. doi:10.1128/jb.184.4.889-894.2002.
- Willems, R. J. L., Top, J., Santen, M. van, Robinson, D. A., Coque, T. M., Baquero, F., et al. (2005). Global Spread of Vancomycin-resistant *Enterococcus faecium* from distinct nosocomial genetic complex. *Emerging Infectious Diseases* 11, 821. doi:10.3201/eid1106.041204.
- Zhang, X., Vrijenhoek, J. E. P., Bonten, M. J. M., Willems, R. J. L., and van Schaik, W. (2011). A genetic element present on megaplasms allows *Enterococcus faecium* to use raffinose as carbon source. *Environmental Microbiology* 13, 518-528. doi:10.1111/j.1462-2920.2010.02355.x.
- Zhong, Z., Kwok, L.-Y., Hou, Q., Sun, Y., Li, W., Zhang, H., et al. (2019). Comparative genomic analysis revealed great plasticity and environmental adaptation of the genomes of *Enterococcus faecium*. *BMC Genomics* 20, 602. doi:10.1186/s12864-019-5975-8.
- Zhou, X., Willems, R. J. L., Friedrich, A. W., Rossen, J. W. A., and Bathoorn, E. (2020). *Enterococcus faecium*: from microbiological insights to practical recommendations for infection control and diagnostics. *Antimicrob Resist Infect Control* 9. doi:10.1186/s13756-020-00770-1.

516 **Conflict of Interest:** The authors declare that the research was conducted in the absence  
517 of any commercial or financial relationships that could be constructed as a potential  
518 conflict of interest.

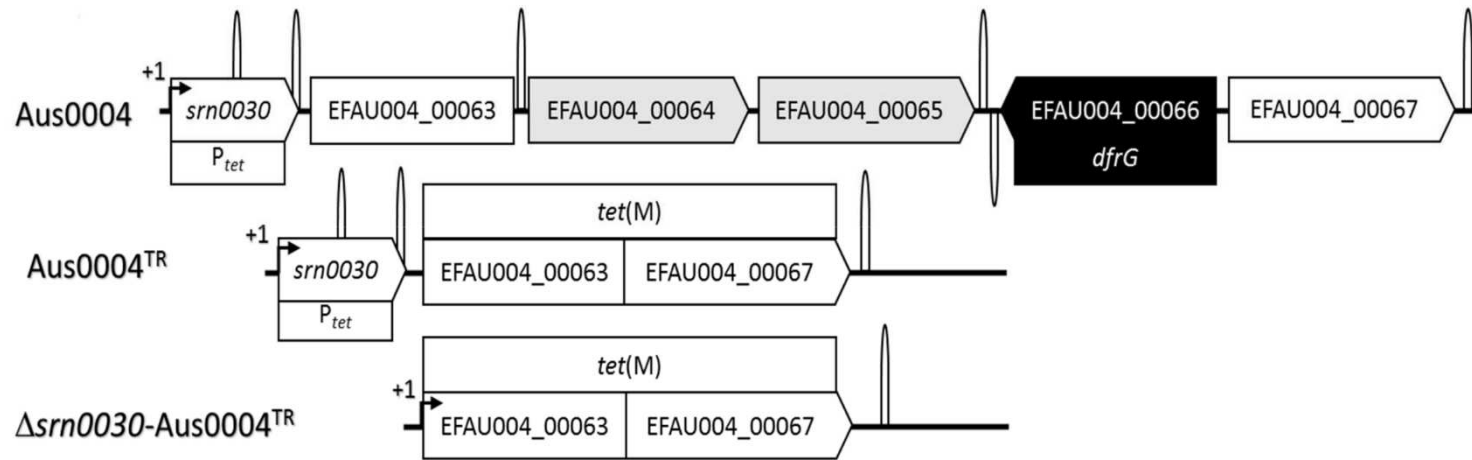
519 **Table 1.** MIC values (mg/L) of tetracycline and tigecycline against the different *E. faecium*  
520 strains.

<i>E. faecium</i>	Tetracycline	Tigecycline
Aus0004	0.12	0.01
Aus0004 <sup>TR</sup>	32	0.03
$\Delta ern0030$ -Aus0004 <sup>TR</sup>	4	0.03
Aus0004 <sup>TR</sup> _pAT29	32	0.03
Aus0004 <sup>TR</sup> _pAT29 $\Omega ern0030$	32	0.03
$\Delta ern0030$ -Aus0004 <sup>TR</sup> _pAT29	4	0.03
$\Delta ern0030$ -Aus0004 <sup>TR</sup> _pAT29 $\Omega ern0030$	4	0.03
Aus0004_pAT29	0.12	0.01
Aus0004_pAT29 $\Omega tetM$	64	0.25
Aus0004_pAT29 $\Omega tetM$ - $\Delta_{204-238} ern0030$	32	0.12
Aus0004_pAT29 $\Omega tetM$ -TGA <sub>172-174</sub>	32	0.03

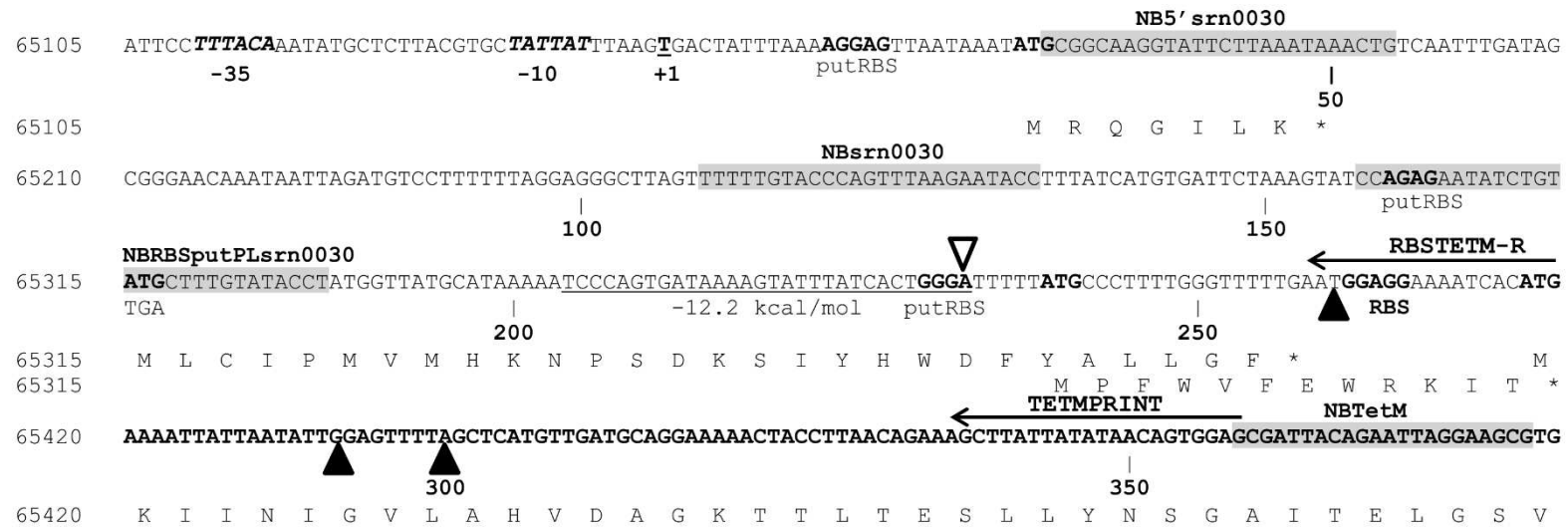
521

522

A



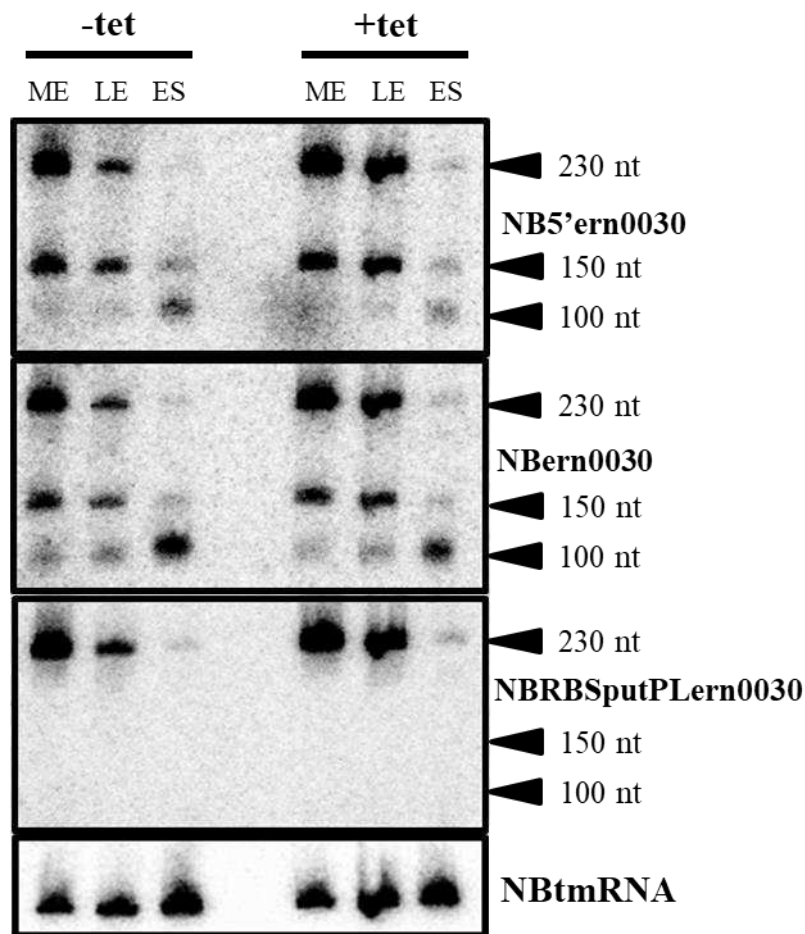
B



524 **Figure 1** | Genetic environment of chromosomal *ern0030* and *tet(M)* genes in *E. faecium* strains.

525 (A) Genetic organization of *ern0030* and *tet(M)* genes in *E. faecium* Aus0004 and derivative strains (Supplementary Table S1). Open reading  
526 frames (ORFs) are indicated by horizontal arrows. Putative transcription terminators, predicted with ARNold finding terminator software  
527 (Naville et al., 2011), were indicated by loops. The *tet(M)* gene actually corresponds to both EFAU004\_00063 and EFAU0004\_00067  
528 annotated genes (Genbank accession no. CP003351.1).

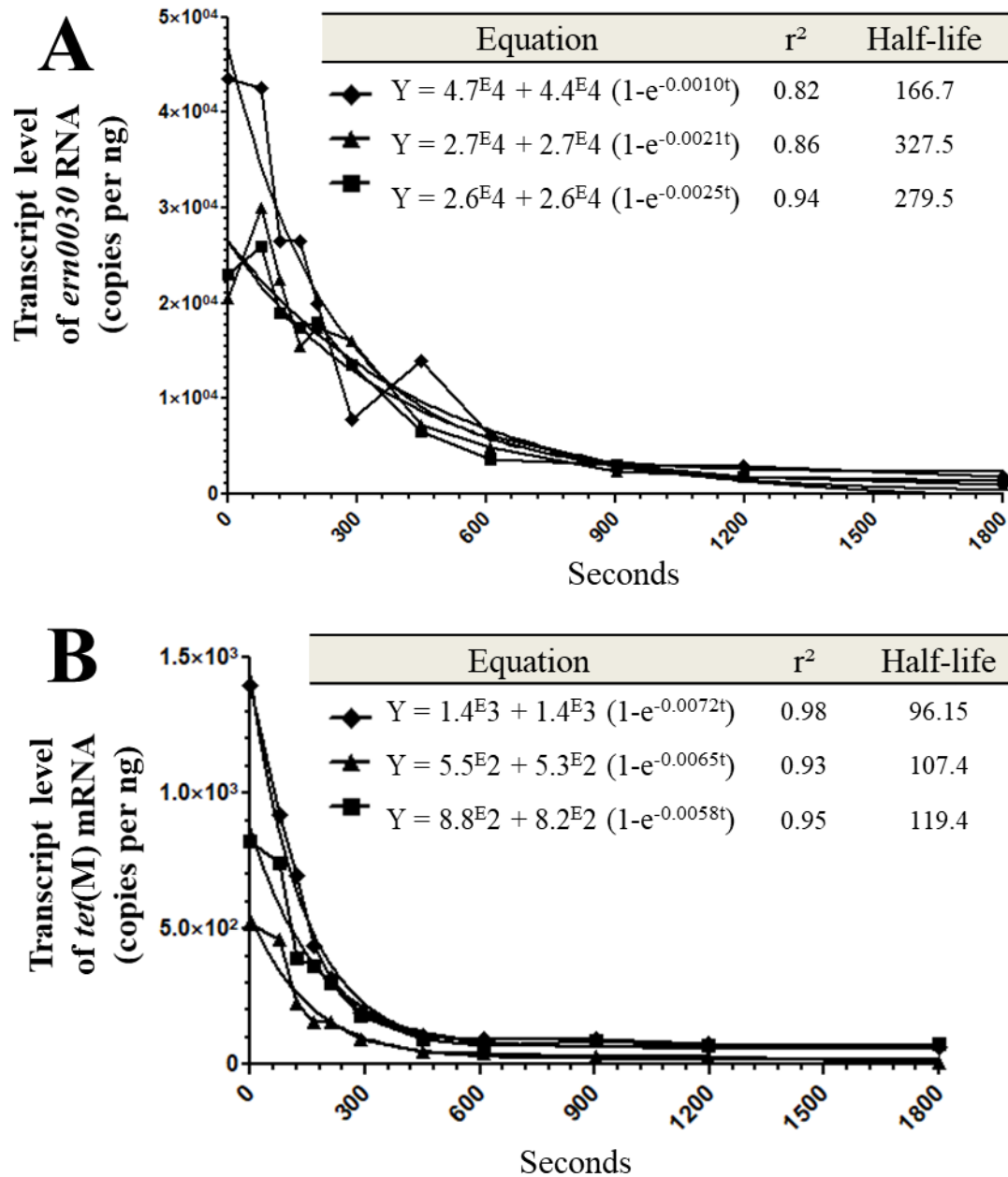
529 (B) Annotated nucleotide sequence of the 5' untranslated region of *tet(M)*. The -35 and -10 promoter boxes are in bold as well as the  
530 transcription start site (+1) from which the numbering starts. The white triangle corresponds to the 3' end of *ern0030*. Putative ribosome  
531 binding sites (RBS) of *ern0030*, the putative leader peptide and *tet(M)* are indicated as well as translation initiation codons of the putative  
532 leader peptide and *tet(M)*. The rho-independent transcriptional terminator (comprising the stem-loop between positions 204 and 233) and  
533 the putative leader peptide sequence are represented. Positions of Northern blot probes and toeprint primers (Supplementary Table S3) are  
534 indicated. Toeprint positions (Figure 6A) are symbolized by black triangles.



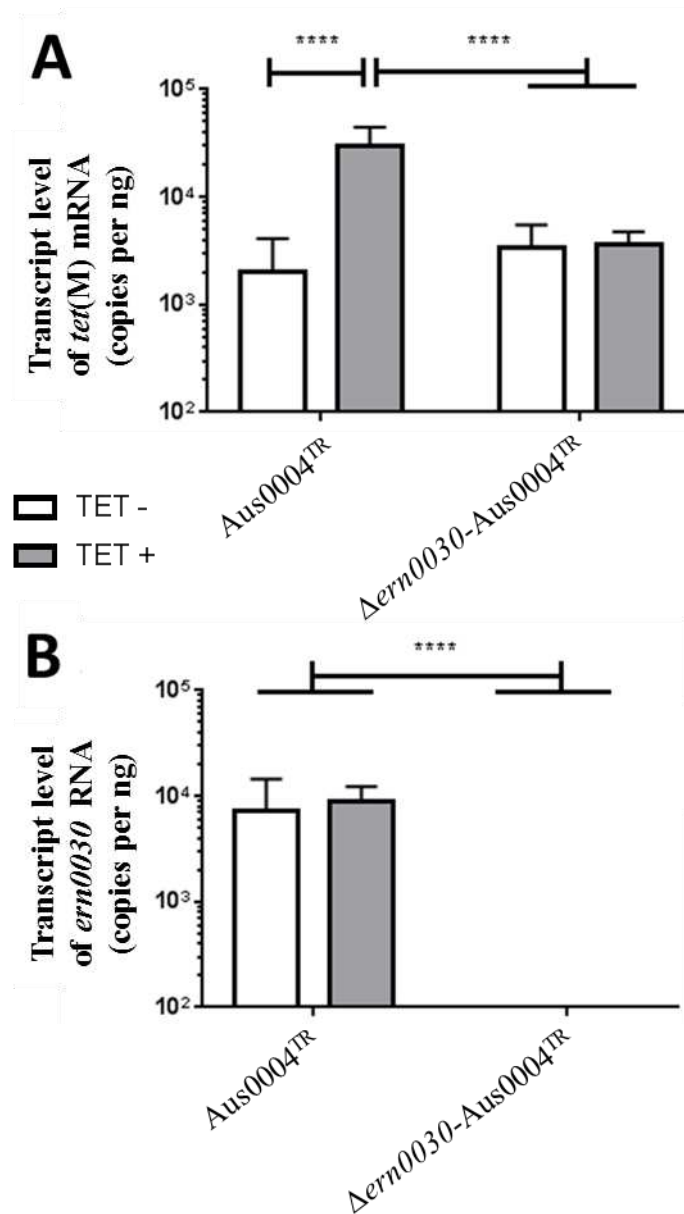
**Figure 2 | Expression of *ern0030* transcripts with or without induction by tetracycline SIC in *E. faecium* Aus0004.**

Northern Blots were performed on total RNAs extracted from Aus0004 cells collected at an ODs of 1, 2.2 and 3 corresponding to mid (ME), late (LE) exponential and early stationary (ES) phases of growth. The RNAs were extracted on cells grown in the absence (-tet) or presence (+tet) of tetracycline SIC (equivalent 1/32 of MIC). The detection of transcripts was performed by using three different probes (Figure 1B, Supplementary Table S3). tmRNA levels were used as internal loading controls. The data show a representative experiment among three independent biological replicates.

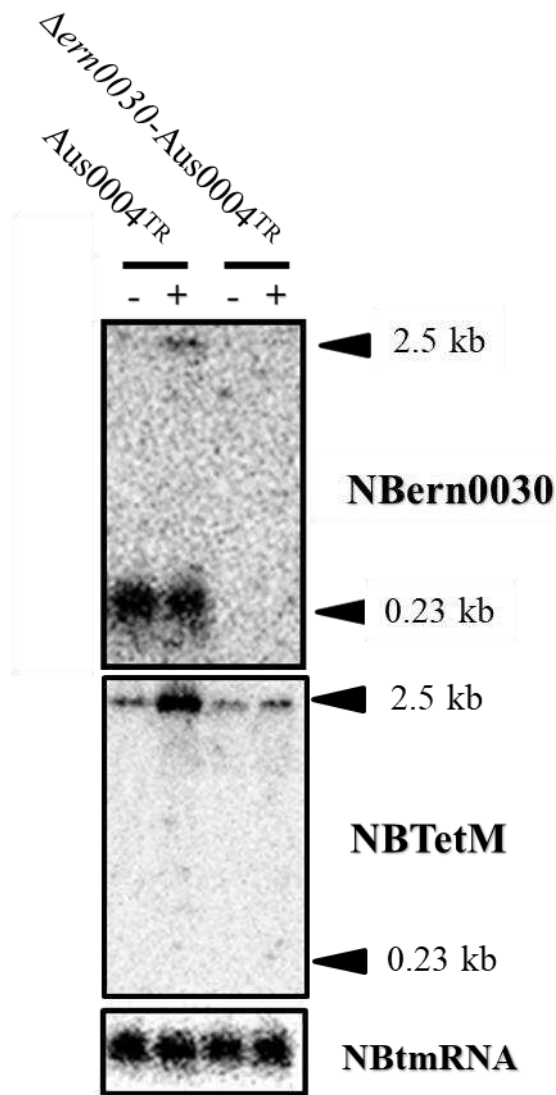




**Figure 3 |** Half-life measurement of *ern0030* (A) and *tet(M)* (B) transcripts in *E. faecium* Aus0004<sup>TR</sup> in the absence of tetracycline. After reaching the ME phase, rifampicin was added at  $t = 0$ . RNA abundance was measured by absolute RT-qPCR (number of copies by ng of total RNAs) and half-lives (in seconds) were determined by nonlinear one phase decay.

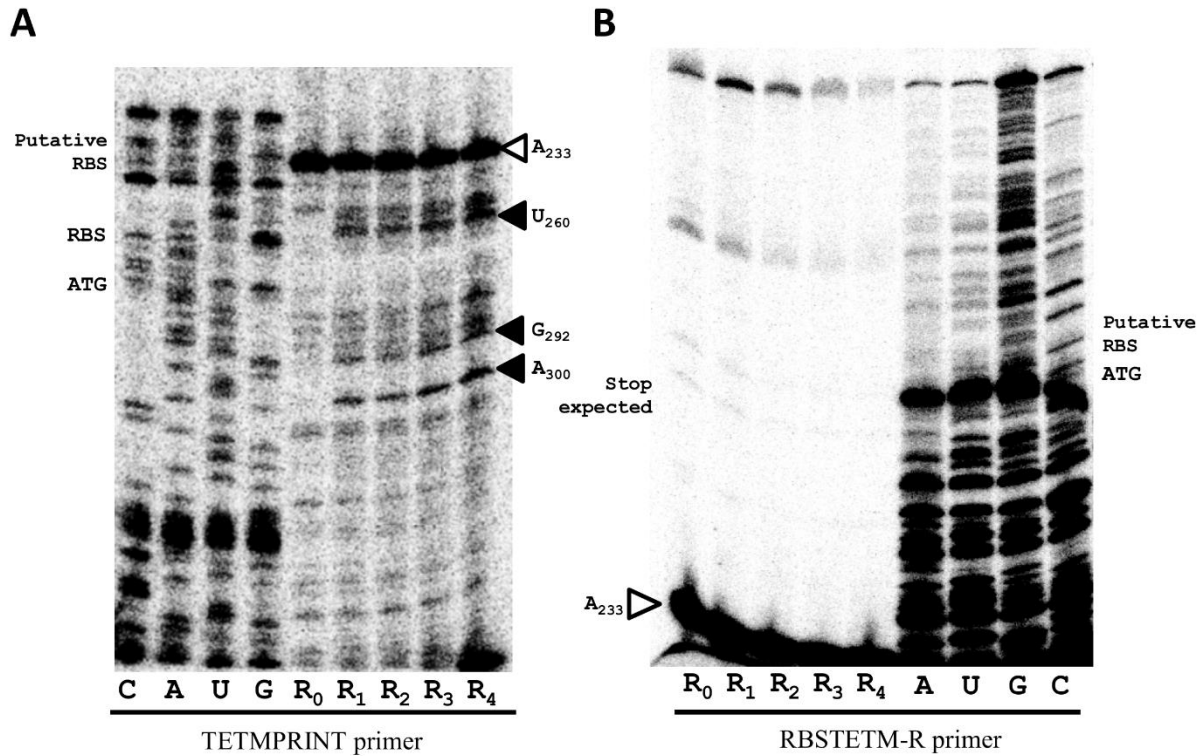


**Figure 4** | Expression levels of *tet(M)* (panels A) and *ern0030* (panel B) in *E. faecium* strains in the presence or absence of tetracycline SIC. Total RNA samples were collected from bacterial cells grown until mid-exponential (ME) phase in BHI broth supplemented (grey bar) or not (white bar) with tetracycline SIC (equivalent to 1/32 of MIC). Transcript levels (expressed as copy numbers per ng of RNA) in each sample were determined by absolute quantification. The multiple comparison between each sample was done by the Tukey test (p-value adjusted: \* < 0.05, \*\*\* < 0.001, \*\*\*\* < 0.0001). Error bar shows 95%CI between three independent biological replicates.



**Figure 5** | Expression of *ern0030* and *tet(M)* transcripts with or without induction by tetracycline SIC in Aus0004<sup>TR</sup> and  $\Delta ern0030$ -Aus0004<sup>TR</sup> strains.

Northern Blots were performed on total RNAs extracted from Aus0004<sup>TR</sup> and  $\Delta ern0030$ -Aus0004<sup>TR</sup> cells collected at an OD of 1 corresponding to mid exponential (ME) phase of growth. The RNAs were extracted on cells grown in the absence (-) or presence (+) of tetracycline SIC (equivalent 1/32 of MIC). The detection of transcripts was performed on two different probes, one with a *ern0030*-specific probe and another with a *tet(M)*-specific probe (Supplementary Table S3). tmRNA levels were used as internal loading controls. The data show a representative experiment among three independent biological replicates.



**Figure 6 | Translational toeprint assays on 5' untranslated region (5' UTR) of *tet(M)*.**

The cDNAs were synthesized using primers TETMPRINT (A) and RBSTETM-R (B) that target the RBS of *tet(M)* and that of the putative leader peptide, respectively. Concentrations of 70S ribosomes were indicated as follows:  $R_0$ : 0 nM,  $R_1$ : 0.25 nM,  $R_2$ : 0.5 nM,  $R_3$ : 1 nM,  $R_4$ : 1.9 nM. Positions of toeprints are represented by black triangles according to *ern0030* numbering (Figure 1B). The position  $A_{233}$  corresponds to the end of the stem-loop of the *ern0030* rho-independent transcription terminator.

Supplementary Table S1 | Bacterial strains used in the study.

Strains	Relevant characteristics	Reference
<b><i>E. faecium</i></b>		
Aus0004	<i>vanB</i> -positive reference strain, isolated in Australia, 1998	(Lam et al., 2012)
Aus0004 <sup>TR</sup>	Mutant derived from Aus0004 with functional <i>tet(M)</i> (see Figure 1A)	This study
$\Delta ern0030$ -Aus0004 <sup>TR</sup>	Aus0004 <sup>TR</sup> deleted for <i>ern0030</i> (see Figure 1A)	This study
Aus0004 <sup>TR</sup> _pAT29	Aus0004 <sup>TR</sup> carrying empty pAT29 vector	This study
Aus0004 <sup>TR</sup> _pAT29 $\Omega ern0030$	Aus0004 <sup>TR</sup> carrying recombinant plasmid pAT29 $\Omega ern0030$	This study
$\Delta ern0030$ -Aus0004 <sup>TR</sup> _pAT29	$\Delta ern0030$ -Aus0004 <sup>TR</sup> carrying empty pAT29 vector	This study
$\Delta ern0030$ -Aus0004 <sup>TR</sup> _pAT29 $\Omega ern0030$	$\Delta ern0030$ -Aus0004 <sup>TR</sup> carrying recombinant plasmid pAT29 $\Omega ern0030$	This study
Aus0004_pAT29	Aus0004 carrying empty pAT29 vector	This study
Aus0004_pAT29 $\Omega tet(M)$	Aus0004 carrying recombinant plasmid pAT29 $\Omega tet(M)$	This study
Aus0004_pAT29 $\Omega tet(M)$ - $\Delta_{204-238} ern0030$	Aus0004 carrying pAT29 $\Omega tet(M)$ - $\Delta_{204-238} ern0030$	This study
Aus0004_pAT29 $\Omega tet(M)$ -TGA <sub>172-174</sub>	Aus0004 carrying pAT29 $\Omega tet(M)$ -TGA <sub>172-174</sub>	This study
<b><i>E. coli</i></b>		
XL1blue	Strain used for 5'3'RACE mapping and cloning	Stratagene®
EC1000	Strain used for cloning	(Leenhouts et al., 1996)

**Supplementary Table S2 | Plasmids used in the study.**

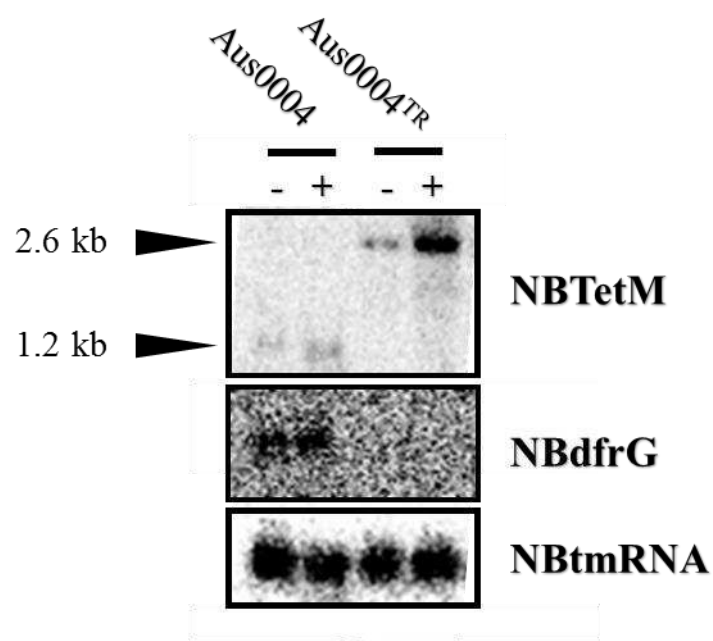
Plasmids	Relevant characteristics <sup>a</sup>	Reference
pGEM-T	High-efficiency TA cloning vector used for 5'3'RACE mapping (Amp <sup>r</sup> )	Promega <sup>®</sup>
pWS3	Temperature-sensitive pG(+)host9-derived shuttle vector used for gene disruption (Spc <sup>r</sup> )	(Zhang et al., 2011)
pAT29	High-copy-number promoterless shuttle vector used for cloning (Spc <sup>r</sup> )	(Trieu-Cuot et al., 1990)
pAT29Ω <i>ern0030</i>	pAT29 containing <i>ern0030</i> with its native promoter and rho-independent terminator (Spc <sup>r</sup> )	This study
pAT29Ω <i>tet</i> (M)	pAT29 containing both <i>ern0030</i> and <i>tet</i> (M) genes (see Figure 1A) (Spc <sup>r</sup> , Tet <sup>r</sup> )	This study
pAT29Ω <i>tet</i> (M)-Δ <sub>204-238</sub> <i>ern0030</i>	pAT29Ω <i>tet</i> (M) with a 35-nt deletion in <i>ern0030</i> comprising its rho-independent terminator (see Figure 1B) (Spc <sup>r</sup> , Tet <sup>r</sup> )	This study
pAT29Ω <i>tet</i> (M)-TGA <sub>172-174</sub>	pAT29Ω <i>tet</i> (M) plasmid with an unique mutation (ATG → TGA) in the translation initiation codon of the putative peptide leader (see Figure 1B) (Spc <sup>r</sup> , Tet <sup>r</sup> )	This study

<sup>a</sup>Amp<sup>r</sup>, ampicillin resistance; Spc<sup>r</sup>, Spectinomycin resistance; Tet<sup>r</sup>, tetracycline resistance.

**Supplementary Table S3 | Oligonucleotides used in the study.**

Name	Nucleotide sequence (5'-3')	Purpose
sRNA0030F1	CATGTGATTCTAAAGTATCCAG	5'3'RACE mapping of <i>ern0030</i>
sRNA0030F2	CTGTATGCTTTGTATACCTATGG	
sRNA0030R1	AAAGGTATTCTTAAACTGGG	
sRNA0030R2	CTAAGCCCTCCTAAAAAAGGACATC	
sRNA0030R3	CCCCTATCAAATTGACAGTT	
KOern0030F1	CTATGAATTCAGCTTTTCATGTGGTTGTCCATATAG	Knockout deletion of <i>ern0030</i>
KOern0030F2	TTTTTATGCCCTTTTGGGTTTTTGAATGG	
KOern0030R1	CCAAAAGGGCATAAAAACTTAAATAATAGCACGTAAGAGC	
KOern0030R2	GTTCTGAATTCCTGTTTGATTACAATTTCCGGCAG	
KOern0030vF	TTGGAGATTCTTTACAAATATGC	
KOern0030vR	TTTCCTGCATCAACATGAGC	
KOdfRGF1	AGAAGGAATTCCAAAGGTACAACGAGGACGG	Restoration of <i>tet(M)</i> function
KOdfRGF2	ATACGCCCTTTATAGTGGAGTACTACATTTACGAGATTTCG	
KOdfRGR1	CTCCACTATAAAGGCGTATATATGCAAGACGTTGTCTTT	
KOdfRGR2	GTTCCAGAATTCCTAAAGGCTTGCTCCAGTAC	
vF-KOdfR	CATCAACACATCGAGGTCCG	
vR-KOdfR	CTCCAGAATAAGCTCTATCAATCTTAC	
pCern0030F	GGTGGTGGATCCCAAATATTGGTACATGATTACAG	Construction of recombinant plasmids
pCern0030R	GGTGGTGAATTCGTGATTTTCTCCATTCAAAAAC	
F-Dloopern0030	ATGCCCTTTTGGGTTTTTGA	
R-Dloopern0030	AACCCAAAAGGGCATTTTTTATGCATAACCATAGGTATAC	
R-StopPLtetM	TACAAAGTCAACAGATATTCTCTGGATACTT	
F-StopPLtetM	TACTGTTGACTTTGTATACCTATGGTTATGC	
R-pCtetM	GGTGGTGAATTCAAAAAGAGCCGATAAGATGAG	
M13F	GTAAAACGACGGCCAGTG	Verification of recombinant plasmid constructions
M13 Reverse short	CAGGAAACAGCTATGAC	
pWS3F1	TTTCCAGTCACGACGTTGT	
pWS3R1	ACTGACAGCTTCCAAGGAGC	
pAT29vF	GTAAAACGACGGCCAG	
pAT29vR	CAGGAAACAGCTATGAC	
qPCRsRNA0030F	CATGTGATTCTAAAGTATCCAGAGA	RT-qPCR
qPCRsRNA0030R	TGCATAACCATAGGTATACAAAGCA	
qPCRTETMF	CGATTACAGAATTAGGAAGCG	
qPCRTETMR	GATGTTACCTTCGTATTTTCC	
adk-F	CCGCACATCTCAACAGGAGA	
adk-R	AGGCCAAGAGCTGTTTCGTT	
T7ern0030F	TAATACGACTCACTATAGGGTGACTATTTAAAGGAGTTAATAAATATGCG	<i>In vitro</i> transcription
R-PCRTETM	GTCCTCGTTGTACCTTTGTC	
F-PCRTETM	GGAGTTTTAGCTCATGTTGATGC	RT-PCR
F-ORF	CCCTTTTGGGTTTTTGAATGGAG	
F-PCRern0030	GGAGTTAATAAATATGCGGCAAGG	
NBtmRNA	TCGCTAGTCTGATGTTCTCTTTTTC	Probes for Northern blot
NBTetM	CGCTTCCTAATTCTGTAATCGC	
NBern0030*	GGTATTCTTAACTGGGTACAAAAA	
NBRBSputPLern0030	AGGTATACAAAGCATACAGATATTCTCTGG	
NB5'ern0030	CAGTTTATTTAAGAATACCTTGCCGC	
NBdfR	GGTTCTTCCTACCTAATATTATCGG	
TETMPRINT	CTCCACTGTTATATAATAAGC	Toeprint assays
RBSTETM-R	CATGTGATTTTCTCCAT	

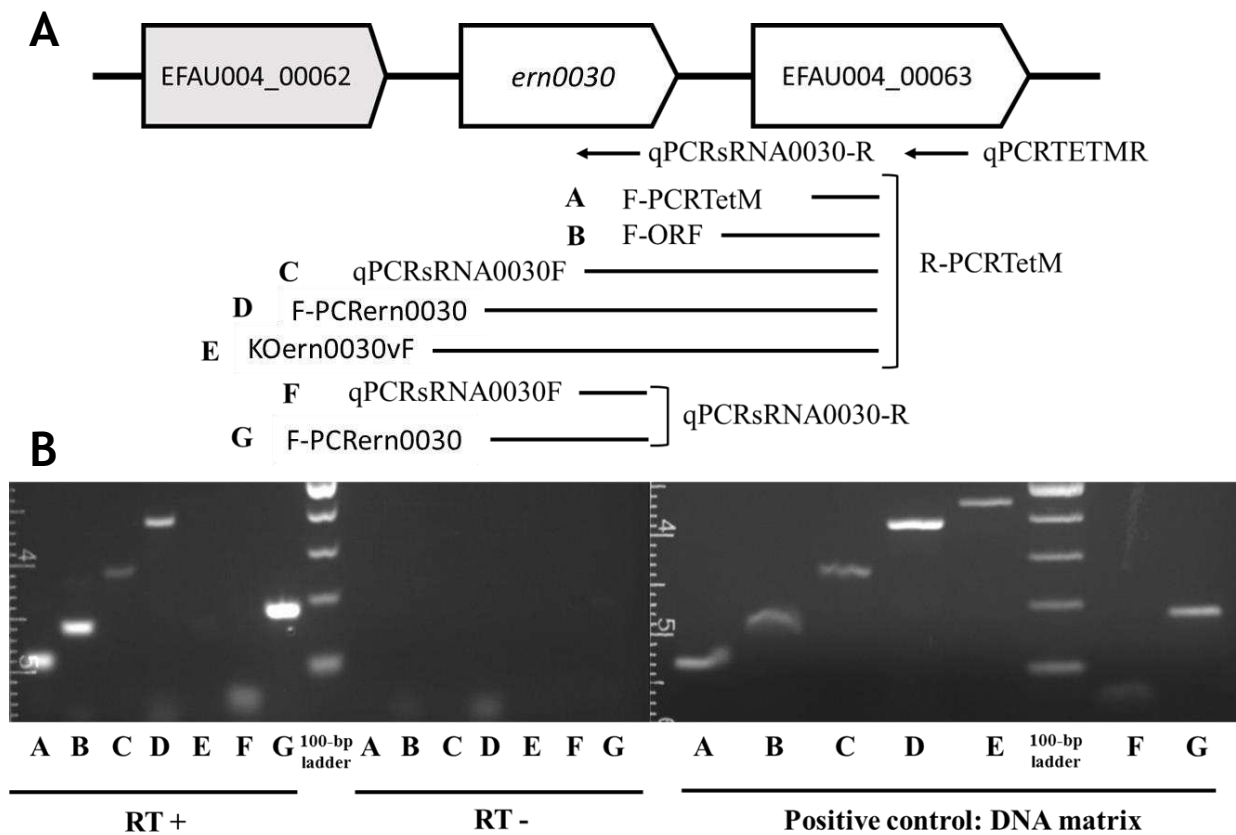
\*Probe for *ern0030* used previously (Sinel et al., 2017).



**Supplementary Figure S1 | Validation of the Aus0004<sup>TR</sup> strain by Northern Blot.**

Northern Blots were performed on total RNAs extracted from Aus0004<sup>TR</sup> cells collected at an OD of 2.1 corresponding to the late exponential phase of growth. The RNAs were extracted on cells grown in the absence (-) or presence (+) of tetracycline SIC (equivalent 1/32 of MIC). The detection of transcripts was performed by using *tet(M)*-specific and *dfrG*-specific probes (Supplementary Table S3). tmRNA levels were used as internal loading controls. The data show a representative experiment among three independent biological replicates.

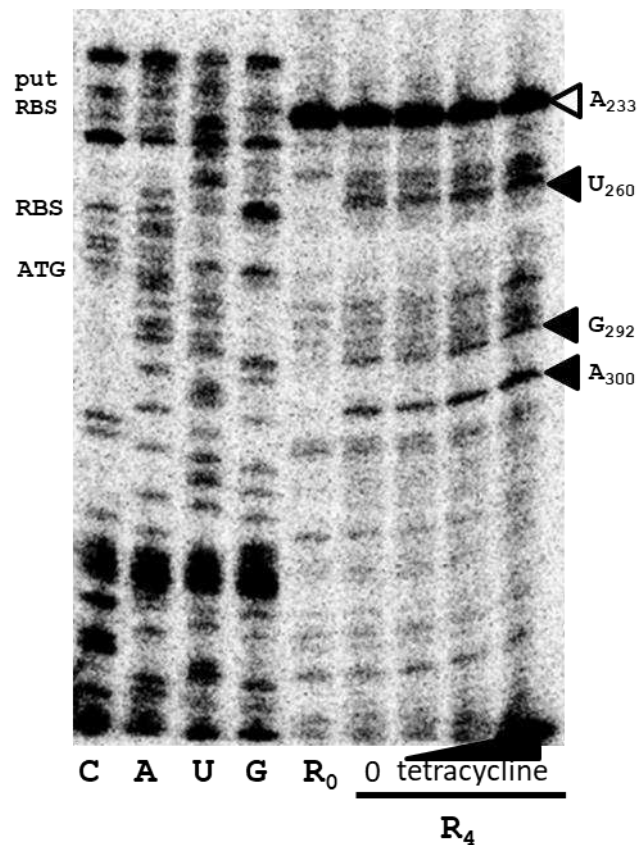




**Supplementary Figure S2 | Mapping of the 5' untranslated region (5' UTR) of *tet(M)* by RT-PCR.**

(A) Schematic representation of the genetic environment of *ern0030* in *E. faecium* Aus0004 (Genbank accession no. CP003351.1). Black labeled lines (A-G) indicate the regions of interest for amplification. The two primers used for reverse transcription are represented by arrows.

(B) Visualization of PCR-amplified products using cDNAs as matrix (RT +). Genomic DNA was used as positive control and DNase-treated RNAs as negative control (RT -).



**Supplementary Figure S3** | Translational toeprint assays on 5' untranslated region (5' UTR) of *tet(M)*.

The cDNAs were synthesized using the TETMPRINT primer that target the RBS of *tet(M)*. Concentrations of 70S ribosomes were indicated as follows: R0: 0 nM, R4: 1.9 nM. Concentrations of tetracycline were from 0 to 288  $\mu$ M (0 M; 281 nM; 9  $\mu$ M; 288  $\mu$ M). Positions of toeprints are represented by black triangles according to *ern0030* numbering (Figure 1B). The position A<sub>233</sub> corresponds to the end of the stem-loop of the *ern0030* rho-independent transcription terminator.

## REFERENCES

- Lam, M. M. C., Seemann, T., Bulach, D. M., Gladman, S. L., Chen, H., Haring, V., et al. (2012). Comparative analysis of the first complete *Enterococcus faecium* genome. *J. Bacteriol.* 194, 2334-2341. doi:10.1128/JB.00259-12.
- Leenhouts, K., Buist, G., Bolhuis, A., ten Berge, A., Kiel, J., Mierau, I., et al. (1996). A general system for generating unlabelled gene replacements in bacterial chromosomes. *Mol. Gen. Genet.* 253, 217-224.
- Sinel, C., Augagneur, Y., Sassi, M., Bronsard, J., Cacaci, M., Guérin, F., et al. (2017). Small RNAs in vancomycin-resistant *Enterococcus faecium* involved in daptomycin response and resistance. *Scientific Reports* 7. doi:10.1038/s41598-017-11265-2.
- Trieu-Cuot, P., Carlier, C., Poyart-Salmeron, C., and Courvalin, P. (1990). A pair of mobilizable shuttle vectors conferring resistance to spectinomycin for molecular cloning in *Escherichia coli* and in gram-positive bacteria. *Nucleic Acids Research* 18, 4296.
- Zhang, X., Vrijenhoek, J. E. P., Bonten, M. J. M., Willems, R. J. L., and van Schaik, W. (2011). A genetic element present on megaplasms allows *Enterococcus faecium* to use raffinose as carbon source. *Environmental Microbiology* 13, 518-528. doi:10.1111/j.1462-2920.2010.02355.

### III. Régulation de deux aaRS par le T-box riboswitch *ern2050*

#### A. Résumé de l'article 2

##### 1. Présentation

Ce travail présente l'investigation fonctionnelle d'un autre des premiers petits ARN mis en évidence chez *E. faecium* : *ern2050* (Sinel *et al.*, 2017a). Après les premières analyses de transcriptomique globale par RNAseq, on a pu décrire *ern2050* comme un régulateur post-transcriptionnel des deux gènes *hisS* et *aspS* codant pour des aminoacyl-ARNt synthétase (aaRS). Cet ARN présente les séquences consensus d'un riboswitch de type T-box. Cet article est le fruit d'une collaboration avec les Dr F.DR. Wencker et Pr W. Ziebuhr (Equipe « nosocomial infection » de l'« institue for molecular infection biology », université de Würzburg, Allemagne), ayant caractérisé expérimentalement la structure secondaire d'*ern2050* et approfondi les expérimentations d'interaction entre les ARNt et *ern2050* (expériences toujours en cours).

Les différents objectifs de ce travail ont été :

- Décrire le rôle fonctionnel de l'ARN *ern2050*,
- Mettre au point un milieu de culture chimiquement défini pour *E. faecium*,
- Valider expérimentalement (caractérisation structurale, transcriptomique, phénotypique et mécanistique) l'ARN *ern2050* comme un régulateur de type T-box riboswitch,
- Décrire le réseau de régulation impliquant *ern2050*.

Comme pour l'étude fonctionnelle des autres ARN, des souches d'étude ont été produites (mutant délété et *trans*-complémenté) [Article 2, figure 1A]. Une analyse transcriptomique globale comparative par RNAseq a été effectuée pour déterminer le type activité (*cis* ou

*trans*). Après avoir catégorisé cet ARN comme un T-box riboswitch, l'objectif a été de valider cette hypothèse expérimentalement notamment par :

- In-line probing pour confirmer qu'*ern2050* soit structurellement proche des autres T-box riboswitchs,
- Utilisation de milieu pauvre dépourvu de différents acides aminés pour déterminer la condition d'adaptation d'*ern2050*,
- Retards sur gel pour mettre en évidence l'interaction entre *ern2050* et les ARNt spécifiques de ce système.

## 2. Principaux résultats

Les principaux résultats de ce travail sont les suivants :

- Données de transcriptomique globale montrant une variation d'expression significative (DEseq analyse, p-value <0,01) pour 22 gènes entre les souches Aus0004 et  $\Delta ern2050$  et seulement pour 5 gènes entre la souche *ern2050 trans*-complémentée et celle avec le plasmide contrôle [Article 2, Table1].
- Validation de la régulation par *ern2050* de deux synthétases co-transcrites (*hisS* et *aspS*) : surexpression dans la souche délétée  $\Delta ern2050$  des deux gènes en RNAseq (Fold-change à 5,5 et 6,1 pour *hisS* et *aspS*, respectivement) et RT-qPCR (Fold-change à 7,2 et 8,1 pour *hisS* et *aspS*, respectivement) [Article 2, Table 1 et Figure 1B].
- Validation expérimentale de la structure secondaire de type T-box riboswitch d'*ern2050* : les séquences consensus sont bien situées sur des régions simple brin, la structure est composée de 3 « stem », dont la première contient bien le « Kink-turn » et la « specifier loop ». La tige-boucle terminatrice prédite est également validée expérimentalement [Article 2, Figure 2].

- Mise en place d'un milieu chimiquement défini et détermination des AA essentiels chez *E. faecium* : la souche Aus0004 est auxotrophe pour 10 AA incluant l'histidine, prototrophe pour 7 AA incluant l'aspartate et braditrophe pour 3 AA [Article 2, Figure 4].
- Validation expérimentale d'une double interaction d'*ern2050* avec les 2 ARNt pris en charge par les deux synthétases régulées : un retard sur gel est observable entre *ern2050* et les deux ARNt<sup>His</sup> et ARNt<sup>Asp</sup>, mais pas pour l'ARNt<sup>Tyr</sup> (utilisé comme contrôle). Néanmoins, la fraction d'ARNt libre reste très importante [Article 2, Figure 3 et S2].
- Mise en évidence de l'importance de l'homéostasie de l'aspartate chez *E. faecium* : la souche délétée a une croissance inférieure à la souche sauvage dans un milieu dépourvu d'aspartate [Article 2, Figure 5].
- Mise en évidence du mécanisme de régulation d'*ern1900* étant également de type T-box riboswitch spécifique de l'histidine, avec la régulation de gènes codant des protéines hypothétiques impliqué dans l'import de l'histidine : la délétion d'*ern2050* semble induire indirectement une répression de l'expression de l'opéron *hisXYZ*, probablement induit par l'augmentation d'ARNt<sup>His</sup> amino-acylé avec l'histidine [Article 2, Figure 6].

## B. Article 2

### **Regulation of *hisS-aspS* operon by T-box riboswitch in *Enterococcus faecium***

Killian Le Neindre<sup>1,2,#</sup>, Freya D. R. Wencker<sup>3,#</sup>, Svetlana Chabelskaya<sup>1</sup>, Sophie Reissier<sup>1</sup>, Loren Dejoies<sup>1,2</sup>, Valérie Bordeau<sup>1</sup>, Brice Felden<sup>1,†</sup>, Vincent Cattoir<sup>1,2,\*</sup> and Wilma Ziebuhr<sup>3</sup>

<sup>1</sup> Inserm UMR\_S 1230, Bacterial Regulatory RNAs and Medicine, University of Rennes 1, Rennes, France

<sup>2</sup> Department of Clinical Microbiology, Rennes University Hospital, Rennes 35000, France.

<sup>3</sup> Institute of Molecular Infection Biology, University of Würzburg, 97080, Germany.

**Article en cours de rédaction**

**Journal envisagé : Nucleic Acid Research**

# Regulation of *hisS-aspS* operon by T-box riboswitch in *Enterococcus faecium*

Killian Le Neindre<sup>1,2,#</sup>, Freya D. R. Wencker<sup>3,#</sup>, Svetlana Chabelskaya<sup>1</sup>, Sophie Reissier<sup>1</sup>, Loren Dejoies<sup>1,2</sup>, Valérie Bordeau<sup>1</sup>, Brice Felden<sup>1,†</sup>, Vincent Cattoir<sup>1,2,\*</sup> and Wilma Ziebuhr<sup>3</sup>

<sup>1</sup> Inserm UMR\_S 1230, Bacterial Regulatory RNAs and Medicine, University of Rennes 1, Rennes 35000, France

<sup>2</sup> Department of Clinical Microbiology, Rennes University Hospital, Rennes 35000, France

<sup>3</sup> Institute of Molecular Infection Biology, University of Würzburg, Würzburg 97080, Germany

\* To whom correspondence should be addressed. Tel: +33(0)299289828 Email: vincent.cattoir@chu-rennes.fr

# Joint Authors

† Deceased

Present Address:

[Killian Le Neindre], Department of Environmental Microbiology, Saint Antoine University Hospital, Paris 75012, France; Inserm UMR\_S 1139, Pathophysiology & Pharmacotoxicology of the Human Placenta, Pre & Postnatal Microbiota, University of Paris, Paris 75006, France.

[Sophie Reissier], Department of Bacteriology-Hygiene, Amiens-Picardie University Hospital, Amiens 80000, France; Microbiology Research Unit AGIR, UR UPJV 4294, Jules Verne University of Picardie, Amiens 80000, France.

[Freya D. R. Wencker], Howard Hughes Medical Institute, Yale University, New Haven, CT 06520-8103; Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06520-8103, United States

## ABSTRACT

A set of small RNAs was previously identified in Vancomycin-resistant *Enterococcus faecium*, a leading cause of nosocomial infections. We describe here the function of *ern2050*, which acts as T-box riboswitch to regulate transcription of downstream *hisS* and *aspS*, aminoacyl-tRNA synthetases genes. Comparative RNAseq between *E. faecium* Aus0004 and isogenic *ern2050* mutant identified genes whose expression is influenced by the RNA. Structural probing of *ern2050* deciphered typical T-box riboswitch-specific elements, including a specifier loop, a T-box consensus sequence and a terminator helix. *In vitro* binding assays between the riboswitch and either tRNA<sup>Asp</sup> or tRNA<sup>His</sup> indicate that each deacylated tRNA interacts with the T-box riboswitch. Their anticodons bind to a GACAC sequence motif within the specifier loop (with GAC and CAC representing Asp and His codons, respectively), whereas tRNA<sup>Tyr</sup> (UA/C-U) does not. Evaluation of *E. faecium* amino acid auxotrophy, with emphasis on *E. faecium* strain Aus0004, revealed auxotrophy for histidine but not for aspartate. Comparative growths between Aus0004 and  $\Delta ern2050$ -Aus0004 further demonstrate that the riboswitch supports growth under aspartate starvation. This is the first example of a functional



riboswitch in *E. faecium* with two overlapping codons allowing a dual tRNA-dependent regulation at transcriptional level.

## INTRODUCTION

Enterococci are commensal Gram-positive bacteria from human and animals guts (1). However, they became major opportunistic pathogens, and are involved in many serious healthcare-associated infections with the emergence of multidrug-resistant (MDR) clinical isolates, especially vancomycin-resistant *Enterococcus faecium* (VREF) (2–4). The epidemiological success of VREF is due to a worldwide dissemination of a specific MDR clonal complex, CC17 (4, 5). This widespread dissemination is because *E. faecium* can cope with numerous environmental stresses thanks to its genomic plasticity, adaptive metabolism, and efficient regulations of gene expression by transcriptional factors, riboswitches and small regulatory RNAs (sRNAs).

sRNAs regulate target gene expression at post-transcriptional levels for fitness, metabolic adaptation or antibiotic resistance (6–8). sRNAs are ~50 to 600 nucleotides-long, usually non coding, and modulate gene expression by *cis*- or *trans*-acting mechanisms (9). The first set of sRNAs expressed in *E. faecium* was recently identified and contains 61 sRNAs candidates, including 10 that were validated by Northern blots and qPCR (10). In the present study, the function of one of them, sRNA\_2050 (*ern2050*) was investigated by a comparative transcriptomic analysis between isogenic strains. We provide structural and biological evidences that *ern2050* is a functional T-box riboswitch (11–13). This RNA was located within the 5' leader of a polycistronic mRNA encoding two aminoacyl-tRNA synthetases (aaRS), i.e. Histidyl-tRNA synthetase HisS and Aspartyl-tRNA synthetase AspS.

T-box riboregulators are *cis*-regulatory RNAs that control the bacterial response to aminoacyl-tRNA availability by binding and assessing the aminoacylation status of specific tRNAs (12, 14, 15). They selectively bind a specific cognate tRNA and monitor its aminoacylation status to either repress or induce expression of downstream gene(s) by controlling transcription termination/antitermination; or translation initiation (16, 17). T-box riboswitches are RNA-based regulatory systems that fine-tune expression of amino acid metabolism-related genes. They are widespread in numerous Gram-positive bacteria, including many human pathogens, but despite being predicted among the *Enterococcaceae* family, their experimental validation is missing (11, 18, 19). In this report, we describe a functional T-box riboswitch in the species *E. faecium*. By experimentally addressing its structure and functions, we show that the *E. faecium* T-box riboswitch, preceding the *hisS-aspS* operon, acts at transcriptional level. Most interestingly, the system accepts two different tRNAs as specific ligands through the presence of two overlapping codons within its specifier loop. Thus, availability of either deacylated tRNA<sup>His</sup> or tRNA<sup>Asp</sup> during T-box riboswitch transcription allows complex formation, antiterminator formation and transcriptional read-through. Moreover, we found this novel riboswitch to be essential for growth under aspartate starvation. Together, *E. faecium* utilizes at least one functional riboswitch able to interact with two cognate tRNAs, allowing dual regulation of two essential genes at transcriptional level.

## MATERIAL AND METHODS

### Strains, growth conditions and amino acid auxotrophy.

Aus0004 *E. faecium* reference strain and isogenic mutants (Table S1) were grown under ambient air at 37°C in brain-heart infusion (BHI, Oxoid) broth, on BHI agar (Oxoid) or in chemically defined medium for *E. faecium* (CDM-EFM). CDM-EFM contains, per liter: 6.7 g YNB with ammonium sulfate (MP biomedical), 10 g glucose (Sigma), 1.72 g KH<sub>2</sub>PO<sub>4</sub> (Merck), 10.4 g of K<sub>2</sub>HPO<sub>4</sub> (Prolabo), 1 g sodium acetate tri-hydrate (Sigma-Aldrich), 2 mg Fe(SO<sub>4</sub>)<sub>2</sub>(NH<sub>4</sub>)<sub>2</sub> (Sigma-Aldrich), 2 mg FeCl<sub>3</sub> (Sigma-Aldrich), 4.6 mg ZnCl<sub>2</sub> (Sigma), 2.5 mg CaCl<sub>2</sub>·2H<sub>2</sub>O (Sigma-Aldrich), 2 mg pyridoxal-HCl 5'P (Sigma-Aldrich), 5 mg pyridoxamine dihydrochloride (Sigma-Aldrich), 1 mg nicotinic acid (Sigma), 600 µg thiamine hydrochloride (Sigma-Aldrich), 800 µg riboflavin (Sigma), 600 µg calcium pantothenate (Sigma-Aldrich), 1 mg biotin (Sigma), 15 mg MnSO<sub>4</sub> (Sigma), 10 mg Xanthine (Sigma), 10 mg Uracil (Sigma), 10 mg Guanine Hydrochloride (Aldrich), 10 mg adenine (Sigma), 2.5 mg Lipoic acid (Sigma), 300 mg L-Arginine (Arg, Sigma), 300 mg L-Histidine (His, Sigma-Aldrich), 300 mg L-Isoleucine (Ile, Sigma-Aldrich), 300 mg L-Leucine (Leu, Sigma), 300 mg L-Methionine (Met, Sigma), 300 mg L-Valine (Val, Sigma-Aldrich), 300 mg L-Glutamine (Glu, Sigma), 300 mg L-Glutamic acid (Glu, Sigma), 300 mg L-Phenylalanine (Phe, Sigma-Aldrich), 300 mg L-Tryptophan (Trp, Sigma-Aldrich), 300 mg L-Aspartic acid (Asp, Sigma), 100 mg L-Tyrosine (Tyr, Sigma), 300 mg L-Cysteine (Cys, Sigma), 300 mg L-Alanine (Ala, Sigma), 300 mg L-Asparagine (Asn, Sigma), 300 mg L-Glycine (Gly, Eurobio), 300 mg L-Lysine (Lys, Sigma-Aldrich), 300 mg L-Proline (Pro, Sigma-Aldrich), 300 mg L-Serine (Ser, Sigma) and 300 mg L-Threonine (Thr, Sigma-Aldrich). *Escherichia coli* strains (Table S1) were grown under ambient air and agitation (160 rpm) at 37°C in Luria-Bertani (LB, ThermoFischer Scientific) or on LB agar (ThermoFischer Scientific). The pAT29 and pWS3 plasmids were used for cloning and knock-out deletion, respectively (Table S1).

*E. faecium* amino acid auxotrophy was done by growth on CDM-EFM, omitting a single amino-acid (-AA) (20). *E. faecium* overnight cultures in BHI were washed by sterilized water and diluted in fresh CDM-EFM-AA to a 0.1 OD<sub>600nm</sub>, then cultured at 37°C.

### RNA extraction, RACE mapping and Northern Blots.

Total RNA extraction was performed as previously described (21). Cells harvesting was done at three growth phases: mid-exponential (ME), late-exponential (LE) and early-stationary (ES) phases, corresponding to OD<sub>600nm</sub> of 1, 2.2, and 3, respectively. Cells were broken using acid-washed glass beads (Sigma-Aldrich) in presence of phenol pH 4 (Sigma-Aldrich), with a FastPrep cell disruptor (FP120, MP biomedical). Total RNAs were extracted by phenol/chloroform and ethanol precipitated overnight. RACE mapping was performed as previously reported (22) on cells collected at ME growth phase in BHI broth. Total RNAs (5 µg) from *E. faecium* Aus0004 were circularized using T4RNA ligase (Promega), reverse transcribed by M-MLV RT (Promega) using primer sRNA2050R1 (Table S2). The cDNAs were used for two consecutive PCR reactions using the sRNA2050R2-F1 and sRNA2050R2-F2 primers (Table S2). The PCR products were cloned into pGEMT vector (Promega), transformed

into *E. coli* XL1blue (Table S1), and sequenced using a BigDye™ Terminator v3.1 Cycle Sequencing Kit (ThermoFischer scientific) with a M13 primer (Table S2). Northern blots were performed with RNA extracts from the three growth phases (ME, LE and ES). The RNA samples (5 to 20 µg) were loaded onto denaturing 8% PAGE and transferred on Hybond N+ membranes (GE Healthcare). The RNAs were probed with <sup>32</sup>P 5'-labeled oligonucleotides (Table S2) in ExpressHyb solution (Clontech) and scanned after an exposure with a Typhoon FLA 9500 scanner (GE Healthcare).

### RNA deletion and overexpression strains construction

An *ern2050*-deleted *E. faecium* Aus0004 mutant was constructed by allelic exchange with a truncated copy of the gene corresponding to 5'-3' positions of *ern2050* using pWS3 and specific primers (Tables S1-S2), as described (23). The *ern2050* gene with its native promoter and rho-independent terminator was cloned into pAT29 shuttle vector (24) using specific primers (Tables S1-S2). The recombinant plasmid was introduced into *E. coli* EC1000, and then into *E. faecium*  $\Delta$ *ern2050*-Aus0004. Empty pAT29 plasmid was used as control. The transformants were selected on media containing 100 mg/L (*E. coli*) or 300 mg/L (*E. faecium*) spectinomycin.

### RNAseq experiment

*E. faecium*  $\Delta$ *ern2050*-Aus0004 and Aus0004,  $\Delta$ *ern2050*-Aus0004\_pAT29 and  $\Delta$ *ern2050*-Aus0004\_pAT29 $\Omega$ *ern2050* strains (Table S1) were cultured at 37°C until LE growth phase (OD<sub>600nm</sub> 2.2) in BHI broth supplemented by 300 mg/L spectinomycin if necessary and total RNAs were extracted by phenol/chloroform and ethanol precipitation. Residual chromosomal DNAs were removed using the Turbo DNA-free kit (Invitrogen). RNA integrity, concentrations and purity were verified by Bioanalyzer (Agilent). The RNA samples were sent to ViroScan3D® (Lyon, France) for ribosomal RNA depletion and sequencing. The strand-specific library preparation was performed using the Ovation Universal Prokaryotic RNA-Seq kit (Nugen) with custom depletion (Anydeplete rRNA design *Enterococcus faecium*) and the samples were sequenced using the Illumina NextSeq500 instrument (multiplex protocol, single-end 75-bp reads). For each strain, independent biological duplicates were performed. For bioinformatic analysis, reads were mapped against the chromosome sequence of *E. faecium* Aus0004 (GenBank accession numbers CP003351.1) using the CLC Genomics Workbench software v10.0.1 (CLCbio, Qiagen, San Diego, CA, USA). Fold-change (FC) values, reads per kilobase per million mapped (RPKM) determination, and statistical analysis were performed using the CLC Genomics Workbench and DESeq2 R package (25). Gene expressions were identified using a log<sub>2</sub> absolute fold change and values greater or lower than 2 were considered as induced or repressed, respectively, while statistical significance was defined when adjusted p-value <0.05. Mean expression and log<sub>2</sub> FC values for each gene were plotted and visualized as a MA plot figure.

### Quantitative RT-PCR

The RNAs were treated by the Turbo DNA-free kit (Invitrogen). The amount of treated RNAs was adjusted to 40 ng per sample, and the reaction used a Power SYBR® Green RNA-to-CT™ 1-Step Kit

(ThermoFischer Scientific). The quantification of gene expression was done with a standard calibration of each gene (Validated if correlation  $r^2 > 0.99$ ). The *adk* gene was used as housekeeping gene to estimate the homogeneity between the samples. The details about RT-qPCR processing are provided in Supplementary Method.

### In-line probing

The DNA templates that contain a T7 promoter sequence for *in vitro* transcription using the MEGAscript® T7 Kit (Ambion) were generated by PCR. The Oligonucleotides and DNA templates used to generate the individual T7 templates were listed in Supplementary Table S2. The details about *in vitro* T7 transcription and in-line probing experiments are provided in supplementary Method and supplementary Table S3.

### tRNA/T-box interaction assays

Transcription was performed using MEGAscript® T7 Kit from the PCR-amplified templates from *E. faecium* Aus0004 (Tables S1-S2). The tRNAs were transcribed and purified by phenol-chloroform and ethanol precipitation. After integrity and length control on denaturing 10% PAGE, these tRNAs were radiolabeled as described (26). The tRNA/T-box binding assay were conducted as previously described (27), and *ern2050 in vitro* transcription was done in the presence of a radiolabeled tRNAs, either tRNA<sup>His</sup>, tRNA<sup>Asp</sup> or tRNA<sup>Tyr</sup>. The transcription products (table S3) were separated onto 6% or 10% native PAGE and scanned after exposure onto a Typhoon FLA 9500 scanner (GE Healthcare).

## RESULTS

### *ern2050* length determination, deletion and *trans*-complementation

By RACE mapping, the size of the *ern2050* RNA was determined at 233 nucleotides (nt), mapping to genomic positions 2519789 (5'-end) and 2519552 (3'-end) on the Aus0004 chromosome (Figure S1). Its 3'-end matches the presence of an upstream stem-loop terminator, predicted by ARNold finding terminator software (28). Using these coordinates, the  $\Delta ern2050$ -Aus0004 deletion mutant was constructed.  $\Delta ern2050$ -Aus0004 was transformed with empty pAT29 ( $\Delta ern2050$ -Aus0004\_pAT29, negative control) and *trans*-complemented with pAT29 $\Omega ern2050$  ( $\Delta ern2050$ -Aus0004\_pAT29 $\Omega ern2050$ , Figure 1A) (Table S1). Northern hybridizations at mid exponential (ME), late-exponential (LE) and early-stationary (ES) growth phases revealed that in strain Aus0004, *ern2050* expression was detected early and decreases during growth (Figure 1A). As expected, no transcript was detected in the deletion mutant (Figure 1) while *ern2050* is overexpressed in the deletion mutant complemented with a plasmid carrying the endogenous promoter of *ern2050*, with an expression profile following that of Aus0004 (Figure 1A).

The growth in either BHI medium or TS medium and BHI supplemented to 0.02% of a deoxycholic-cholic acid mix of strains Aus0004,  $\Delta ern2050$ -Aus0004,  $\Delta ern2050$ -Aus0004\_pAT29 and  $\Delta ern2050$ -Aus0004\_pAT29 $\Omega ern2050$  were comparable.

### ***ern2050* RNA influences expression of two adjacent aaRS genes**

To get an insight into the general function of *ern2050* in *E. faecium*, comparative RNAseq experiments were performed on RNAs extracted at the LE phase of the four *E. faecium* isogenic strains (Table S1). A total number of reads ranging between 16 and 33 million for each strain and conditions were obtained, with a low percentage (from 1.2 to 7.3%) of mapped rRNA genes, implying an efficient rRNA depletion (Table S4). The reproducibility of the experimental duplicates was acceptable ( $r^2 > 0.87$ ), and *ern2050* expression levels of the four strains was in accordance with independent Northern blots performed (Figure 1A). The normalized count mean for *ern2050* expression was 0 in  $\Delta ern2050$ -Aus0004, as expected. Between Aus0004 and  $\Delta ern2050$ -Aus0004, 22 genes displayed altered transcript levels ( $FC < -2$  or  $> 2$ , and adjusted p-value  $< 0.01$ ), with 13 upregulated and 9 downregulated genes (Table 1). 15 of these 22 genes were assembled into five predicted operons. Two were predicted to express permeases, two others operons contained genes involved in pyrimidine and chorismate biosynthesis, and the last operon comprised two genes expressing aminoacyl-tRNA synthetases (aaRS).

Upon over expression of *ern2050* in  $\Delta ern2050$ -Aus0004\_pAT29 $\Omega ern2050$ , only five genes showed significantly altered levels (Table 1). These comparative RNAseq results suggest that *ern2050* RNA is probably not a *trans*-acting sRNA. The RNAseq experiments further revealed that expression of three genes located immediately downstream of *ern2050* was significantly upregulated (4 to 6-fold) when *ern2050* is deleted (Table 1). These genes comprise an operon harbouring the histidyl-tRNA synthetase gene *hisS*, the aspartyl-tRNA synthetase *aspS*, as well as gene EFAU004\_02494, a hypothetical protein with analogy to the YitT family of general stress proteins. We analysed further the expression of these genes by RT-qPCR experiments, confirming their upregulations when *ern2050* was lacking (Figure 1B). These combined data suggested that *ern2050* may act as a *cis*-regulatory element that controls *hisS-aspS-EFAU004\_02494* operon expression.

### **Structure probing in solution indicates that *ern2050* is a T-box riboswitch RNA leader.**

A closer inspection of *ern2050* nucleotide sequence revealed the presence of a 14 nucleotide stretch (5'-UAGGUGGUACCG-3', positions 181-194), with similarity to T-box motifs (14), suggesting that *ern2050* might form a T-box riboswitch upstream the aminoacyl-tRNA synthetase genes (Figure S1). To challenge this hypothesis experimentally, we explored *ern2050* solution structure by in-line probing. *ern2050* secondary structure was determined in its 'off-state' (i.e. without tRNA ligand) by in-line probing experiments (Figure 2A) (29). The approach was augmented by computational folding predictions using mfold (30), with experimental data being entered into the program as constraints to subsequently build an *ern2050* RNA secondary structure model in its 'off-state' that would recapitulate (Figure 2B; the predicted structure of the antiterminator is also shown). In-line probing is a non-enzymatic experimental approach that employs the tendency of RNAs to spontaneously degrade according to their secondary structures. Unpaired nucleotides within single-stranded RNA regions have high flexibility allowing adjacent nucleotides to enter the so-called in-line geometry. This

conformation triggers the slow, non-enzymatic cleavage of the phosphodiester bond in the RNA backbone, by a nucleophilic attack of the 2' oxygen on the adjacent phosphorous center) (31). Double-stranded RNA regions and nucleotides involved in long-range interactions are more rigid which prevents nucleotides to enter the in-line conformation, preventing cleavage. Overall, the approach revealed that *ern2050* harbors structural elements associated with T-box riboswitch function such as stem I, stem II, stem IIA/B pseudoknot and stem III, as well as a terminator helix, as described below (Figure 2) (12).

#### *Stem I structure and motifs.*

Towards *ern2050* 5' end, a kink-turn element was identified at the base of stem I (nt U<sub>9</sub>-A<sub>23</sub>) consisting of an asymmetric loop (G<sub>11</sub>-A<sub>23</sub>/G<sub>109</sub>-A<sub>111</sub>) cut in the in-line experiments flanked by two short helices that are not clived. Despite of an enlarged 5' portion of the loop (13 instead of 6 nucleotides), the overall arrangement of the *ern2050* kink-turn matches the structures conserved in many S- and T-box riboswitches, including a 5'-GA-3' motif flanking the loop on both sides (12, 32). Adjacent to the kink-turn structure a specifier loop is present (G<sub>27</sub>-A<sub>34</sub>/C<sub>96</sub>-C<sub>105</sub>) that harbors two overlapping specifier codons (G<sub>101</sub>AC<sub>103</sub> (Asp), C<sub>103</sub>AC<sub>105</sub> (His) (marked in green in Figure 2B), which are expected to interact with the anticodons of cognate Asp- and His-tRNAs, respectively. Furthermore, the distal region of stem I contains an AG-bulge (nt U<sub>48</sub>-A<sub>56</sub>) with a 5'-AGAGA-3' motif (nt A<sub>52</sub>-A<sub>56</sub>) and a terminal loop (nt C<sub>64</sub>-A<sub>74</sub>) with a 5'-GCUGAAA-3' motif (nt G<sub>68</sub>-A<sub>74</sub>). These motifs (marked in orange in Figure 2B) are highly conserved in T-box-riboswitches and were shown to interact with each other to form a platform for tRNA elbow recognition and binding (33, 34). In agreement with this function, both sequences in *ern2050* are protected from degradation in the in-line experiments (Figure 2A), suggesting their interactions and similar function as in other T-box systems.

#### *Stem II and stem IIA/B pseudoknot.*

Next to stem I, in-line probing is compatible with the presence of stem II and a stem IIA/B pseudoknot, both representing conserved structures in most T-box riboswitches (12). Stem II (nt G<sub>115</sub>-C<sub>134</sub>) of *ern2050* lacks an internal S-turn motif usually present in stem II of other T-box riboswitches (Figure 2B) (12). The nucleotides A<sub>140</sub>-A<sub>143</sub> of the apical loop of stem IIA/B (nt G<sub>138</sub>-A<sub>143</sub>) base pair with nucleotides U<sub>151</sub>-U<sub>154</sub> of a single-stranded region immediately downstream of stem IIA/B, to form a pseudoknot structure that occurs in many T-box riboswitches (Figure 2A, gels #3 & 4, marked in light blue in Figure 2B) (12).

#### *Stem III, T-box and terminator stem.*

Stem III is highly conserved in T-box riboswitches, but of variable size and sequence (12). Stem III of *ern2050* comprises a 6-bp-long helix (G<sub>160</sub>-A<sub>165</sub>/U<sub>170</sub>-C<sub>175</sub>) and a 4-nt loop (G<sub>166</sub>-A<sub>169</sub>) (Figure 2A, gels # 3, 4 & 6). Stem III is followed by a single-stranded region interspaced by a short stem (G<sub>183</sub>-C<sub>191</sub>) involving the core of the T-box sequence (U<sub>181</sub>-G<sub>194</sub>) (marked dark blue in Figure 2B). Further downstream, in-line probing identified a 12-bp-long helix (U<sub>207</sub>-C<sub>218</sub>/G<sub>224</sub>-A<sub>235</sub>) with a 5-nt- loop (A<sub>219</sub>-

U<sub>223</sub>), representing the terminator stem (Figure 2A, gel #5). This structure is essential for gene regulation by transcriptional T-box riboswitches as its immediate formation forces the RNA polymerase to terminate transcription prematurely, which usually occurs in the absence of the uncharged cognate tRNA ligand (i.e. when tRNAs are charged with the amino acid). The *ern2050* terminator stem harbors 6 (conserved) nucleotides (C<sub>208</sub>-U<sub>213</sub>, marked by asterisks in Figure 2B) that can also pair with nucleotides from the T-box sequence as well, leading to the formation of a characteristic T-box loop (nt U<sub>185</sub>-C<sub>191</sub>) as part of an antiterminator (Figure 2B, top). Terminator and antiterminator are mutually exclusive structures. Only an interaction of the free 3'-CCA end of the uncharged tRNA ligands with the T-box bulge (U<sub>185</sub>GG<sub>187</sub>-3') allows a temporal stabilization of the antiterminator, enabling the RNA polymerase to transcribe the downstream genes. Altogether, the structural elements identified in *ern2050* strongly speak in favor of this RNA being a putative T-box riboswitch.

### **The *hisS-aspS* T-box riboswitch binds two tRNAs**

Next, we analyzed whether or not *ern2050* is functionally able to interact with tRNA ligands, which is the indicative hallmark of all T-box riboswitch systems. Most T-box riboswitches specifically recognize and bind one distinct tRNA species, the latter often being associated with the amino acid metabolic pathway controlled by the system. As *ern2050* precedes an operon encoding genes responsible for Histidyl- and Aspartyl-tRNA synthesis, we hypothesized that one of these two tRNAs could interact with the putative T-box riboswitch. While the interaction between tRNA 3'-CCA ends and the T-box bulge is common to all tRNAs, specificity for a distinct tRNA species is usually mediated by a codon present in the specifier loop of the T-box riboswitch that pairs with the anticodon of the uncharged tRNA. Surprisingly, the specifier loop in *ern2050* harbors two overlapping specifier codons, one for aspartate (GAC) and one for histidine (CAC) (Figure 2), promoting the idea of a dual riboswitch capable of accepting both tRNAs as ligands. To challenge this hypothesis, we performed *in vitro* interaction assays between *ern2050* and various tRNA species. Taking into account that tRNA binding is an induced-fit process that takes place while the T-box riboswitch RNA is transcribed (33), *in vitro* transcription of *ern2050* was performed in the presence of either (radioactively labeled) tRNA<sup>Asp</sup>, tRNA<sup>His</sup> or tRNA<sup>Tyr</sup>, with the latter serving as negative control. Putative RNA-RNA complexes were subsequently visualized by autoradiography in native polyacrylamide gels. Figure 3 shows that labeled tRNA<sup>Asp</sup> binds *ern2050*, as expected. Moreover, another shift appears with labeled tRNA<sup>His</sup>. The two tRNA-*ern2050* complexes possess distinct native conformations as they migrate differently on a native gel (Figure 3). However, the complexed tRNA fraction remained low compared to the free tRNAs (Figure S2).

### **Impacts of *ern2050* deletion on amino-acid starvation**

#### *Amino acid auxotrophy determination.*

T-box conformations depend upon the availability of aminoacylated tRNAs (aatRNA) in the bacteria. The level of amino acid (aa) in the growth medium could be influence this availability of aatRNA (35–38). The relationship between *ern2050* and aa starvation requires to determine the aa auxotrophy of *E.*

*faecium* strain Aus0004. The various amino-acid biosynthesis pathways were determined in *E. faecium* Aus0004 using GapMind (39) (except aa not including on GapMind software: Ala, Asp and Glu) and challenged with our RNAseq data collected from Aus0004 (Table S6). Eight pathways were complete and expressed (RPKM > 10): Asn, Cys, Gln, Gly, Lys, Phe, Pro and Tyr (Table S6). Other biosynthetic pathways were incomplete or absent in the Aus0004 genome. All these incomplete or missing biosynthetic pathways corresponded to auxotrophic aa was funded experimentally (Figure 4). The phenotypic assay consisted to growing the Aus0004 and  $\Delta ern2050$ -Aus0004 strains on CDM-EFM medium, where each aa was omitted independently. For both strains, 10 aa were auxotrophs and 7 prototrophs. Despite the expression of *glyA* (EFAU004\_02299), Gly was auxotroph. Lys, Cys and Phe, were braditroph. The lack of Lysine was more detrimental for growth of strain  $\Delta ern2050$ -Aus0004 compared to Aus0004.

#### *Impacts of amino acid starvation on bacterial growth.*

Aspartic acid is prototroph and Histidine is auxotroph in Aus0004 *E. faecium* strain. The minimal concentration of L-Histidine required for a standard growth was 5 mg/L. Tyrosine was considered as negative control since tRNA<sup>Tyr</sup> does not interact with *ern2050* (Figure 3). Growth of strains Aus0004 and  $\Delta ern2050$ -Aus0004 were compared in CDM-EFM medium, with different concentrations of Asp (0 or 5 mg/L), His (5 mg/L) and Tyr (0 mg/L) (Figure 5).

The absence of Aspartate has a negative effect on  $\Delta ern2050$ -Aus0004 growth. When combined with His or Tyr restrictions, growth was more severely impacted, whereas omitting only His and/or Tyr had no effect. Surprisingly, *hisS*, *aspS* and *ern2050* expression levels were not impacted by the different aa starvations (Table S7, Figure S3). In all these media, *hisS* and *aspS* expression increased with  $\Delta ern2050$ -Aus0004 (Table S7), in accordance to RNAseq data.

## DISCUSSION

A T-box mediated transcriptional regulation of *hisS-aspS* operon was reported here to be functional in *E. faecium*. This is the first example of a functional T-box in this specie. A novelty in these findings was that this new T-box RNA leader was dual, and binds each of the two tRNAs, Asp and His, that were aminoacylated by the two downstream aaRSs by the presence of two overlapping codons within the specifier loop. Such a two-codon T-box was reported in *Clostridium acetobutylicum* with tRNA<sup>Asn</sup> and tRNA<sup>Glu</sup> (40).

The *hisS-aspS* operon was present in other firmicutes genomes, and those were predicted to contain an upstream T-box leader with Asp (GAC) as codon specifier (11). Its structure was determined by in-line probing and matched the overall scaffold of *cis*-encoded T-box leaders acting at transcriptional levels. The potential *trans*-activity of *ern2050* was excluded by our comparative RNAseq data, since it influences expression of only a handful of genes (Table 1, Table S5B).



The two aaRS were under the same transcriptional regulation (Figure 1B, table S7). Conserved consensus sequences were present on T-box leader RNAs (14, 41, 42) and they were located in *ern2050* sequence (Figure S1). This RNA contains a typical T-box sequence, AGUA-I box, AG box, GNUG and F-box. This T-box RNA was annotated in TBDB T-box database, (ID F1TD5WKQ) (19). Specifier codon of *hisS-aspS* T-Box RNA leader previously described was GAC (Asp) or AAC (Asn) depending on the strain (11). However, the specifier codon corresponding to His (CAC) overlapped with the first codon. Among other species sequences (Table S9), at least one of the two codons was conserved on all sequences analysed. Aus0004 has two copies of tRNA<sup>His</sup>, both with a GUG anticodon, and one copy of tRNA<sup>Asp</sup> with a GUC anticodon (43, 44).

*In vitro* transcription binding assays in the presence of each of the two tRNAs, showed that each predicted tRNA binds the T-box RNA, whereas an unrelated tRNA<sup>Tyr</sup> does not (Figure 3). Thus, that T-box RNA leader binds the two tRNAs corresponding to the two downstream aaRSs. It suggests functionality of the T-box regulation of expression of the two aminoacyl-tRNA synthetases according to the pool of available charged tRNA<sup>His</sup> and tRNA<sup>Asp</sup>. Genome sequence inspections (Table S9) and tRNA analysis (43, 44), all strains analysed could bind at least one of the two tRNAs, but not always a double binding capacity, that could be considered as a further bacterial adaptation.

The T-box riboswitch deletion influenced expression of other operons (Table 1), likely indirectly. In fact, the operon EFAU004\_00962-00971 was the pyrimidine biosynthetic operon using aspartate as precursor to produce UMP. This operon is regulated by riboswitch mediated by the regulatory protein PyrR (45, 46) (Figure 6A). The complex PyrR-UMP could interact with riboswitch and promotes termination conformation. With  $\Delta ern2050$ -Aus0004 strain, the high *aspS* expression cause decreasing aspartate availability needed for UMP biosynthesis (Figure 6B). In UMP starvation, PyrR has complexed with PRPP or GTP. In absence of the complex PyrR-UMP, the antitermination conformation was promoted (47, 48). The second operon altered by indirect effect of *hisS-aspS* operon high-regulated was EFAU004\_02288-02290 (Table 1). A putative small RNA, sRNA\_1900 (*ern1900*) was located upstream this operon (10), and contain a consensus sequences of typical T-box RNA leader. This predicted T-box RNA leader has two adjacent rho-independent terminators associated with T-box sequence (GGUGGAACCACGU) as described (11). On TBDB T-box database, *ern1900* was annotated with ID LVNDX2SZ (19). The predicted specifier codon was CAC corresponding to tRNA<sup>His</sup> (Figure 6A). Therefore, the level decreases of deacylated tRNA<sup>His</sup> cause a transcription termination upstream the operon 02288-02290. The high expression of *hisS* was the cause of *ern1900* changing conformation (Figure 6B). This operon was annotated as amino-acid transporter and could correspond to *hisXYZ* operon previously described (11). We supposed that *hisXYZ* operon could be implicated in histidine import (Figure 6A).

The T-box mediated regulation of *hisS-aspS* was important for bacterial metabolism, to fuel optimally translation with charged amino acids. Strain  $\Delta ern2050$ -Aus0004 has difficult to growth under aspartate starvation (Figure 5), presumably because the T-box regulation of *aspS* expression is gone. However, *hisS* and *aspS* expression stay unaffected under different aa starvations (Table S7, Figure S3). The

biosynthesis of Aspartate depends to tricarboxylic acid cycle and thus was implicated in many metabolic pathways including Lysine biosynthesis (49) (Figure 6). The high quantity of AspS protein in  $\Delta ern2050$ -Aus0004 causes an important decreasing of cellular aspartate availability. Under aspartate starvation, the aspartate biosynthesis was insufficient to compensate the bacterial metabolic requirement as observed with lysine starvation in  $\Delta ern2050$ -Aus0004 (Figure 4). This aa, Aspartate, was probably important for global metabolism of bacteria (Figure 6). T-box regulation of *aspS* expression is required to ensure Aspartate availability in the cell. Otherwise, bacterial growing was perturbed. Many T box family genes, such as aaRS genes, are essential and disruption of the regulation of these genes can reduce competitive fitness or be lethal.

## DATA AVAILABILITY

Geo-RNAseq data: GSE 161499

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

## ACKNOWLEDGEMENT

We are thankful to Stéphane Dréano (IGDR, Rennes, France) for the DNA sequencing of all constructs in this study and to François Guerin (Rennes hospital, France) for knock-out deletion protocol on *E. faecium* species.

The work is dedicated to the memory of Prof. Brice Felden.

## FUNDING

Sophie Reissier is the recipient of a fellowship funded by the Fondation pour la Recherche Médicale (FRM). This work was funded by the Agence Nationale pour la Recherche (grant number ANR-15-CE12-0003-01 'sRNA-Fit' to Brice Felden), by the FRM (grant number DBF20160635724 'Bactéries et champignons face aux antibiotiques et antifongiques' to Brice Felden; personal grant as a member of the scientific council to Vincent Cattoir) and by the Institut National de la Santé et de la Recherche Médicale (INSERM).

## CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## REFERENCES

1. Willems, R.J. and van Schaik, W. (2009) Transition of *Enterococcus faecium* from commensal organism to nosocomial pathogen. *Future Microbiology*, **4**, 1125–1135.

2. Arias,C.A. and Murray,B.E. (2012) The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat Rev Microbiol*, **10**, 266–278.
3. Ayobami,O., Willrich,N., Reuss,A., Eckmanns,T. and Markwart,R. (2020) The ongoing challenge of vancomycin-resistant *Enterococcus faecium* and *Enterococcus faecalis* in Europe: an epidemiological analysis of bloodstream infections. *Emerg Microbes Infect*, **9**, 1180–1193.
4. Willems,R.J.L., Top,J., Santen,M. van, Robinson,D.A., Coque,T.M., Baquero,F., Grundmann,H. and Bonten,M.J.M. (2005) Global spread of vancomycin-resistant *Enterococcus faecium* from distinct nosocomial genetic complex. *Emerging Infectious Diseases*, **11**, 821.
5. Lee,T., Pang,S., Abraham,S. and Coombs,G.W. (2019) Antimicrobial-resistant CC17 *Enterococcus faecium*: The past, the present and the future. *J Glob Antimicrob Resist*, **16**.
6. Wagner,E.G.H. and Romby,P. (2015) Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. *Adv. Genet.*, **90**, 133–208.
7. Barquist,L. and Vogel,J. (2015) Accelerating discovery and functional analysis of small RNAs with new technologies. *Annu. Rev. Genet.*, **49**, 367–394.
8. Lalaouna,D., Eyraud,A., Chabelskaya,S., Felden,B. and Massé,E. (2014) Regulatory RNAs involved in bacterial antibiotic resistance. *PLoS Pathog.*, **10**, e1004299.
9. Caldelari,I., Chao,Y., Romby,P. and Vogel,J. (2013) RNA-mediated regulation in pathogenic bacteria. *Cold Spring Harb Perspect Med*, **3**.
10. Sinel,C., Augagneur,Y., Sassi,M., Bronsard,J., Cacaci,M., Guérin,F., Sanguinetti,M., Meignen,P., Cattoir,V. and Felden,B. (2017) Small RNAs in vancomycin-resistant *Enterococcus faecium* involved in daptomycin response and resistance. *Scientific Reports*, **7**.
11. Vitreschak,A.G., Mironov,A.A., Lyubetsky,V.A. and Gelfand,M.S. (2008) Comparative genomic analysis of T-box regulatory systems in bacteria. *RNA*, **14**, 717–735.
12. Kreuzer,K.D. and Henkin,T.M. (2018) The T-Box Riboswitch: tRNA as an effector to modulate gene regulation. *Microbiol Spectr*, **6**.
13. Putzer,H. and Laalami,S. (2013) Regulation of the expression of aminoacyl-tRNA synthetases and translation factors. Landes Bioscience.
14. Grundy,F.J., Moir,T.R., Haldeman,M.T. and Henkin,T.M. (2002) Sequence requirements for terminators and antiterminators in the T box transcription antitermination system: disparity between conservation and functional requirements. *Nucleic Acids Res.*, **30**, 1646–1655.
15. Henkin,T.M. (2000) Transcription termination control in bacteria. *Curr. Opin. Microbiol.*, **3**, 149–153.
16. Suddala,K.C. and Zhang,J. (2019) An evolving tale of two interacting RNAs-themes and variations of the T-box riboswitch mechanism. *IUBMB Life*, **71**, 1167–1180.
17. Seliverstov,A.V., Putzer,H., Gelfand,M.S. and Lyubetsky,V.A. (2005) Comparative analysis of RNA regulatory elements of amino acid metabolism genes in Actinobacteria. *BMC Microbiol*, **5**, 54.
18. Novichkov,P.S., Kazakov,A.E., Ravcheev,D.A., Leyn,S.A., Kovaleva,G.Y., Sutormin,R.A., Kazanov,M.D., Riehl,W., Arkin,A.P., Dubchak,I., *et al.* (2013) RegPrecise 3.0—a resource for genome-scale exploration of transcriptional regulation in bacteria. *BMC Genomics*, **14**, 745.

19. Marchand,J.A., Pierson Smela,M.D., Jordan,T.H.H., Narasimhan,K. and Church,G.M. (2020) TBDB: a database of structurally annotated T-box riboswitch:tRNA pairs. *Nucleic Acids Res.*, 10.1093/nar/gkaa721.
20. Veith,N., Solheim,M., van Grinsven,K.W.A., Olivier,B.G., Levering,J., Grosseholz,R., Hugenholtz,J., Holo,H., Nes,I., Teusink,B., *et al.* (2015) Using a genome-scale metabolic model of *Enterococcus faecalis* V583 to assess amino acid uptake and its impact on central metabolism. *Appl. Environ. Microbiol.*, **81**, 1622–1633.
21. Bronsard,J., Pascreau,G., Sassi,M., Mauro,T., Augagneur,Y. and Felden,B. (2017) sRNA and *cis*-antisense sRNA identification in *Staphylococcus aureus* highlights an unusual sRNA gene cluster with one encoding a secreted peptide. *Sci Rep*, **7**, 4565.
22. Pinel-Marie,M.-L., Brielle,R. and Felden,B. (2014) Dual toxic-peptide-coding *Staphylococcus aureus* RNA under antisense regulation targets host cells and bacterial rivals unequally. *Cell Reports*, **7**, 424–435.
23. Zhang,X., Vrijenhoek,J.E.P., Bonten,M.J.M., Willems,R.J.L. and van Schaik,W. (2011) A genetic element present on megaplasms allows *Enterococcus faecium* to use raffinose as carbon source. *Environmental Microbiology*, **13**, 518–528.
24. Trieu-Cuot,P., Carlier,C., Poyart-Salmeron,C. and Courvalin,P. (1990) A pair of mobilizable shuttle vectors conferring resistance to spectinomycin for molecular cloning in *Escherichia coli* and in gram-positive bacteria. *Nucleic Acids Res.*, **18**, 4296.
25. Anders,S. and Huber,W. (2010) Differential expression analysis for sequence count data. *Genome Biology*, **11**, R106.
26. Antal,M., Bordeau,V., Douchin,V. and Felden,B. (2005) A small bacterial RNA regulates a putative ABC transporter. *J. Biol. Chem.*, **280**, 7901–7908.
27. Schoenfelder,S.M. (2014) In vitro Transcription (IVT) and tRNA Binding Assay. *Bio-protocol*, **4**, e1234–e1234.
28. Naville,M., Ghuillot-Gaudeffroy,A., Marchais,A. and Gautheret,D. (2011) ARNold: a web tool for the prediction of Rho-independent transcription terminators. *RNA Biol*, **8**, 11–13.
29. Regulski,E.E. and Breaker,R.R. (2008) In-line probing analysis of riboswitches. *Methods Mol Biol*, **419**, 53–67.
30. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406.
31. Soukup,G.A. and Breaker,R.R. (1999) Relationship between internucleotide linkage geometry and the stability of RNA. *RNA*, **5**, 1308–1325.
32. Winkler,W.C., Grundy,F.J., Murphy,B.A. and Henkin,T.M. (2001) The GA motif: an RNA element common to bacterial antitermination systems, rRNA, and eukaryotic RNAs. *RNA*, **7**, 1165–1172.
33. Zhang,J. and Ferré-D'Amaré,A.R. (2013) Co-crystal structure of a T-box riboswitch stem I domain in complex with its cognate tRNA. *Nature*, **500**, 363–366.
34. Grigg,J.C. and Ke,A. (2013) Sequence, structure, and stacking: specifics of tRNA anchoring to the T box riboswitch. *RNA Biol*, **10**, 1761–1764.

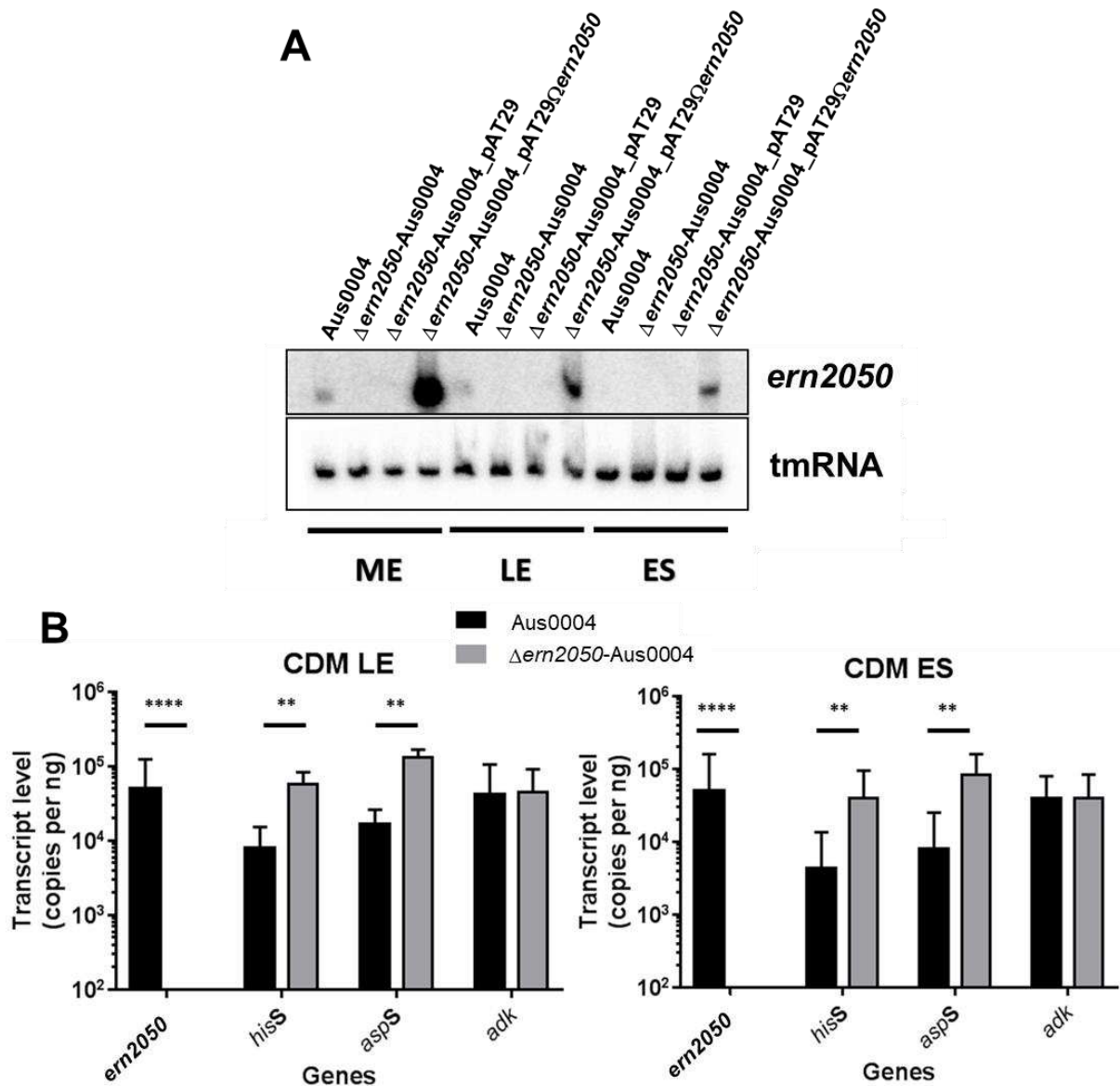
35. Grandoni, J.A., Zahler, S.A. and Calvo, J.M. (1992) Transcriptional regulation of the *ilv-leu* operon of *Bacillus subtilis*. *J. Bacteriol.*, **174**, 3212–3219.
36. Linares, D.M., Fernández, M., Del-Río, B., Ladero, V., Martín, M.C. and Álvarez, M.A. (2012) The tyrosyl-tRNA synthetase like gene located in the tyramine biosynthesis cluster of *Enterococcus durans* is transcriptionally regulated by tyrosine concentration and extracellular pH. *BMC Microbiol.*, **12**, 23.
37. Delorme, C., Ehrlich, S.D. and Renault, P. (1999) Regulation of expression of the *Lactococcus lactis* histidine operon. *J. Bacteriol.*, **181**, 2026–2037.
38. Henkin, T.M., Glass, B.L. and Grundy, F.J. (1992) Analysis of the *Bacillus subtilis* *tyrS* gene: conservation of a regulatory sequence in multiple tRNA synthetase genes. *J. Bacteriol.*, **174**, 1299–1306.
39. Price, M.N., Deutschbauer, A.M. and Arkin, A.P. (2020) GapMind: automated annotation of amino acid biosynthesis. *mSystems*, **5**.
40. Saad, N.Y., Stamatopoulou, V., Brayé, M., Drainas, D., Stathopoulos, C. and Becker, H.D. (2013) Two-codon T-box riboswitch binding two tRNAs. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 12756–12761.
41. Rollins, S.M., Grundy, F.J. and Henkin, T.M. (1997) Analysis of *cis*-acting sequence and structural elements required for antitermination of the *Bacillus subtilis* *tyrS* gene. *Mol. Microbiol.*, **25**, 411–421.
42. Grundy, F.J. and Henkin, T.M. (1994) Conservation of a transcription antitermination mechanism in aminoacyl-tRNA synthetase and amino acid biosynthesis genes in gram-positive bacteria. *J. Mol. Biol.*, **235**, 798–804.
43. Chan, P.P. and Lowe, T.M. (2009) GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res.*, **37**, D93–97.
44. Chan, P.P. and Lowe, T.M. (2016) GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res.*, **44**, D184–D189.
45. Turner, R.J., Lu, Y. and Switzer, R.L. (1994) Regulation of the *Bacillus subtilis* pyrimidine biosynthetic (*pyr*) gene cluster by an autogenous transcriptional attenuation mechanism. *J. Bacteriol.*, **176**, 3708–3722.
46. Turnbough, C.L. and Switzer, R.L. (2008) Regulation of pyrimidine biosynthetic gene expression in bacteria: repression without repressors. *Microbiol. Mol. Biol. Rev.*, **72**, 266–300, table of contents.
47. Zhang, H. and Switzer, R.L. (2003) Transcriptional pausing in the *Bacillus subtilis* *pyr* operon in vitro: a role in transcriptional attenuation? *J. Bacteriol.*, **185**, 4764–4771.
48. Chander, P., Halbig, K.M., Miller, J.K., Fields, C.J., Bonner, H.K.S., Grabner, G.K., Switzer, R.L. and Smith, J.L. (2005) Structure of the nucleotide complex of PyrR, the *pyr* attenuation protein from *Bacillus caldolyticus*, suggests dual regulation by pyrimidine and purine nucleotides. *J. Bacteriol.*, **187**, 1773–1782.
49. Viola, R.E. (2001) The central enzymes of the aspartate family of amino acid biosynthesis. *Acc Chem Res*, **34**, 339–349.

**Table 1. List of transcript genes significantly altered ( $p < 0.01$ ) on RNA-seq experiment.**

Each significant value was bolded. Column “ $\Delta$ /WT” compare  $\Delta ern2050$ -Aus0004 with Aus0004 strains and column “ $\Delta pQ/\Delta p$ ” compare  $\Delta ern2050$ -Aus0004\_pAT29 $\Omega ern2050$  with  $\Delta ern2050$ -Aus0004\_pAT29 strains (Table S1). Fold change and p-value adjusted were performed with DESeq2 R package. Putative operons were framed. Total RNA-seq datas was described in supplementary Table S5.

Gene_name EFAU004_	Product	Strand	Start	End	Fold Change	p-val adj	Fold Change	p-val adj
					$\Delta$ /WT		$\Delta pQ/\Delta p$	
<i>ern0030</i>	5'UTR <i>tet</i> (M)	+	65160	65376	<b>-3.30</b>	<b>1.17<sup>E-03</sup></b>	-1.07	0.97
00239	aldehyde-alcohol dehydrogenase	+	237492	240090	-1.07	1.00	<b>-4.28</b>	<b>4.08<sup>E-05</sup></b>
00338	tyrosine decarboxylase	+	333950	335786	<b>4.70</b>	<b>1.84<sup>E-03</sup></b>	1.48	0.53
00886	alpha amylase	+	897464	899231	<b>-5.62</b>	<b>4.10<sup>E-05</sup></b>	-1.28	0.83
00887	extracellular solute-binding protein	+	899332	900592	<b>-3.60</b>	<b>8.15<sup>E-03</sup></b>	-1.36	0.70
00963	Orotidine 5'-phosphate decarboxylase	-	981442	982159	<b>3.52</b>	<b>4.24<sup>E-04</sup></b>	-1.41	0.77
00964	Dihydroorotate dehydrogenase 1B	-	982155	983088	<b>3.88</b>	<b>4.24<sup>E-04</sup></b>	-1.39	0.79
00965	dihydroorotate dehydrogenase, electron transfer subunit	-	983085	983861	<b>3.32</b>	<b>2.35<sup>E-03</sup></b>	-1.50	0.71
00966	carbamoyl-phosphate synthase large subunit	-	983914	987100	<b>3.70</b>	<b>8.15<sup>E-03</sup></b>	-1.61	0.64
01135	sensory box histidine kinase	+	1164060	1164845	1.26	1.00	<b>-2.85</b>	<b>0.01</b>
01558	rhodanese-like domain-containing protein	-	1575640	1576582	1.12	1.00	<b>3.18</b>	<b>4.08<sup>E-05</sup></b>
01576	prephenate dehydratase	-	1591275	1592112	<b>3.20</b>	<b>1.75<sup>E-03</sup></b>	-1.05	0.97
01577	shikimate kinase	-	1592109	1592618	<b>3.15</b>	<b>4.68<sup>E-03</sup></b>	-1.11	0.95
01582	phospho-2-dehydro-3-deoxyheptonate aldolase	-	1597292	1598314	<b>3.15</b>	<b>1.84<sup>E-03</sup></b>	-1.22	0.85
01663	hypothetical protein	+	1681671	1681826	<b>-3.25</b>	<b>6.52<sup>E-04</sup></b>	1.17	0.91
01742	LacI family transcriptional regulator	-	1769940	1770921	<b>-3.93</b>	<b>4.53<sup>E-03</sup></b>	1.04	0.98
02118	hypothetical protein (LwpA)	-	2146932	2149413	1.93	0.83	<b>2.90</b>	<b>2.79<sup>E-04</sup></b>
02288	family 3 extracellular solute-binding protein	-	2325722	2326538	<b>-6.16</b>	<b>1.99<sup>E-06</sup></b>	1.12	0.95
02289	putative glutamine ABC transporter ATP-binding protein	-	2326555	2327191	<b>-7.57</b>	<b>2.44<sup>E-09</sup></b>	-1.34	NA
02290	amino acid ABC transporter permease	-	2327209	2327851	<b>-5.21</b>	<b>2.67<sup>E-04</sup></b>	-1.52	0.70
02296	hypothetical protein	+	2334103	2334345	<b>-2.91</b>	<b>6.85<sup>E-03</sup></b>	1.22	0.84
02363	hypothetical protein	-	2402295	2402589	<b>3.39</b>	<b>1.81<sup>E-03</sup></b>	<b>4.03</b>	<b>4.08<sup>E-05</sup></b>
02370	phage tail tape measure protein TP901	-	2409209	2411725	<b>2.92</b>	<b>5.83<sup>E-03</sup></b>	2.20	0.16
02494	hypothetical protein	-	2515457	2516333	<b>4.13</b>	<b>7.13<sup>E-07</sup></b>	-1.43	0.58
02495	aspartyl-tRNA synthetase	-	2516441	2518211	<b>6.09</b>	<b>6.78<sup>E-12</sup></b>	-1.24	0.79
02496	histidyl-tRNA synthetase	-	2518225	2519518	<b>5.50</b>	<b>1.05<sup>E-10</sup></b>	-1.25	0.78
<i>ern2050</i>	T box leader RNA	-	2519552	2519789	<b>-15.20</b>	<b>6.78<sup>E-12</sup></b>	<b>294.01</b>	<b>3.13<sup>E-113</sup></b>

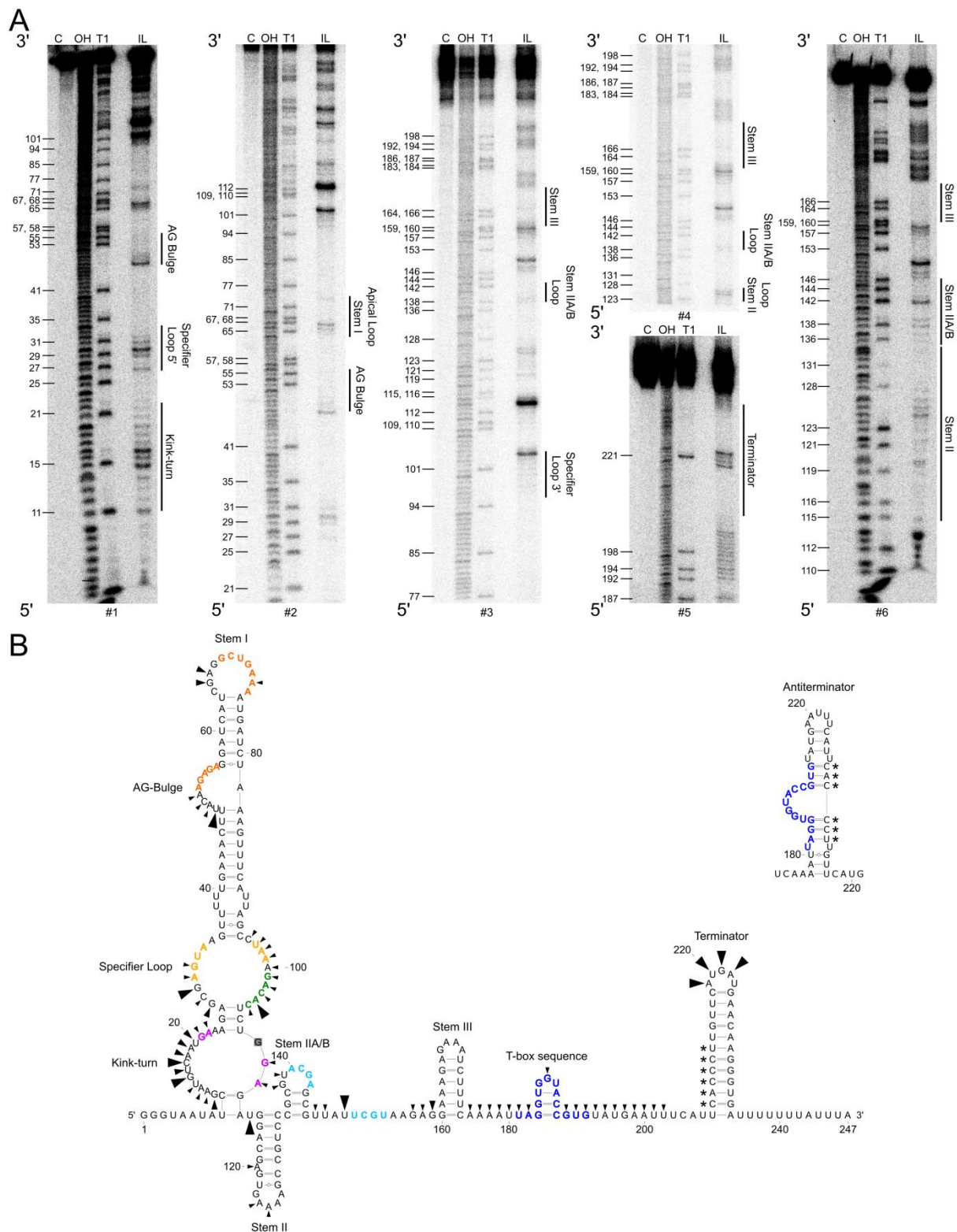
## Figures



**Figure 1. Expression of *ern2050* in *E. faecium* Aus0004.**

(A) Mutant construction validation by Northern blot. Strains used were described in supplementary table S1. Northern blot conditions were described in materials and methods. Using tmRNA as control loading.

(B) Expression analysis of *ern2050*, *hisS* and *aspS* using RT-qPCR. RNA using were extracted from Aus0004 (Black bar) or Δ*ern2050*-Aus0004 (Grey bar) strains on CDM-EFM medium (CDM) at late exponential (LE, left graph) and early stationary (ES, right graph) phase. Adk expression was analysed as control. Error bars correspond to 95% confidence interval between 3 independent biological replicates. Statistical test used was Sidak test (p-value: \* < 0.05, \*\* < 0.01, \*\*\* < 0.001, \*\*\*\* < 0.0001).



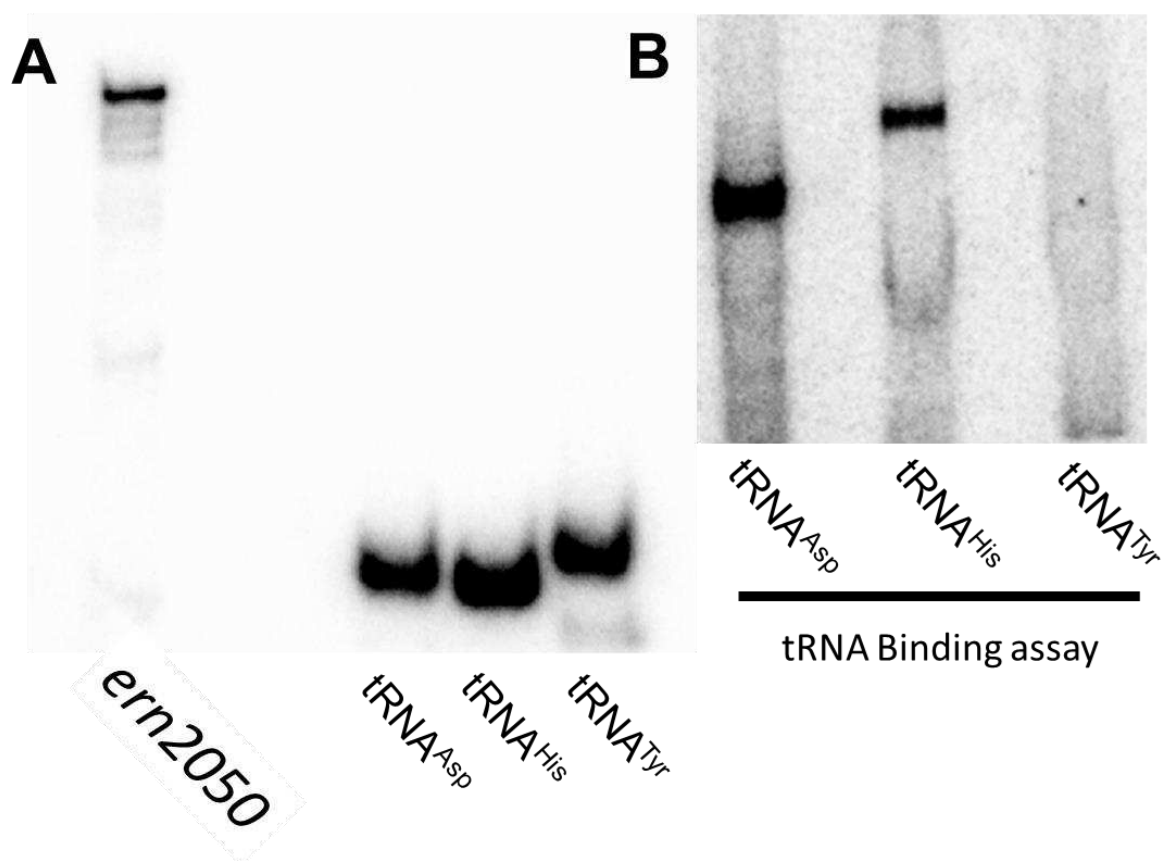
**Figure 2. Secondary structure of *ern2050*.**

(A) In-line probing gels used to build 2D model. Nomenclature of structural motifs according to Kreuzer and Henkin, 2018. Position of guanosines (G) is given on the left of each gel. Structural motifs are specified on right side of each gel. C: control reaction; untreated RNA, OH: alkaline hydrolysis reaction; ladder, T1: RNase T1-treated RNA; G-specific ladder, IL: in-line reaction.

(B) 2D structure model of *ern2050* in its OFF state, predicted antiterminator 2D structure is shown above. Interacting nucleotides of specifier loop are shown in yellow, base pairing nucleotides of AG-bulge and apical loop of stem I in orange, specifier codon in green, GA-motifs of kink-turn in purple,



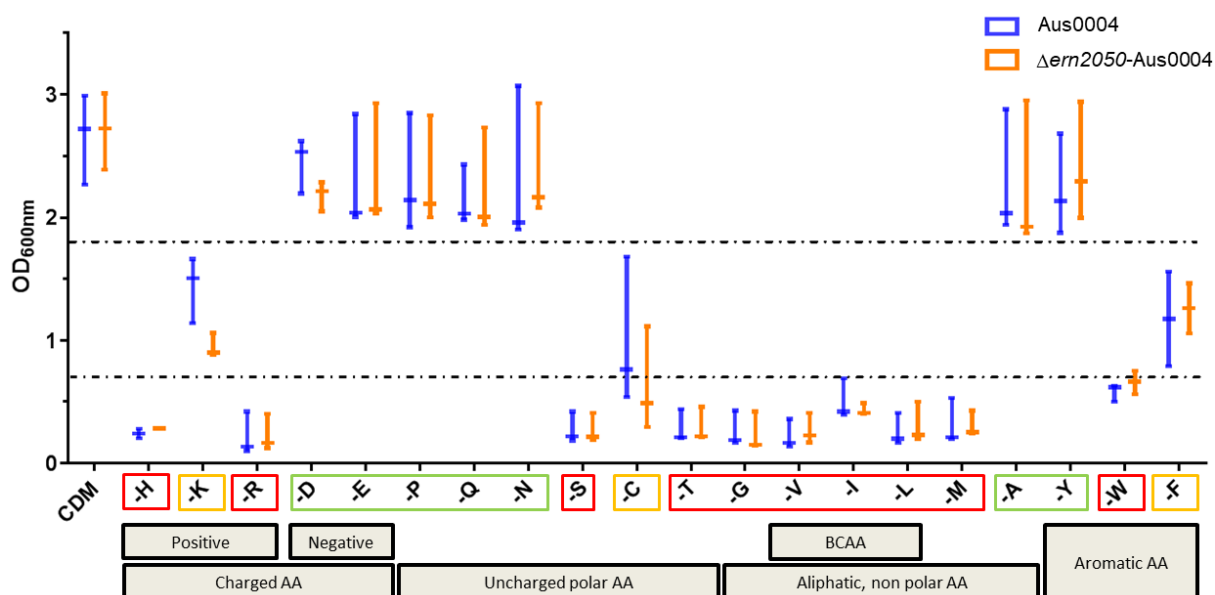
potentially interacting nucleotides of stem IIA/B pseudoknot in light blue, T-box sequence in dark blue. Nucleotides of terminator sequence base pairing with T-box sequence in antiterminator conformation are marked by asterisks (\*). G 109 highlighted in grey indicates 5' end of short *ern2050* RNA used for in-line probing.



**Figure 3. Interaction between tRNA and *ern2050* during *in vitro* transcription.**

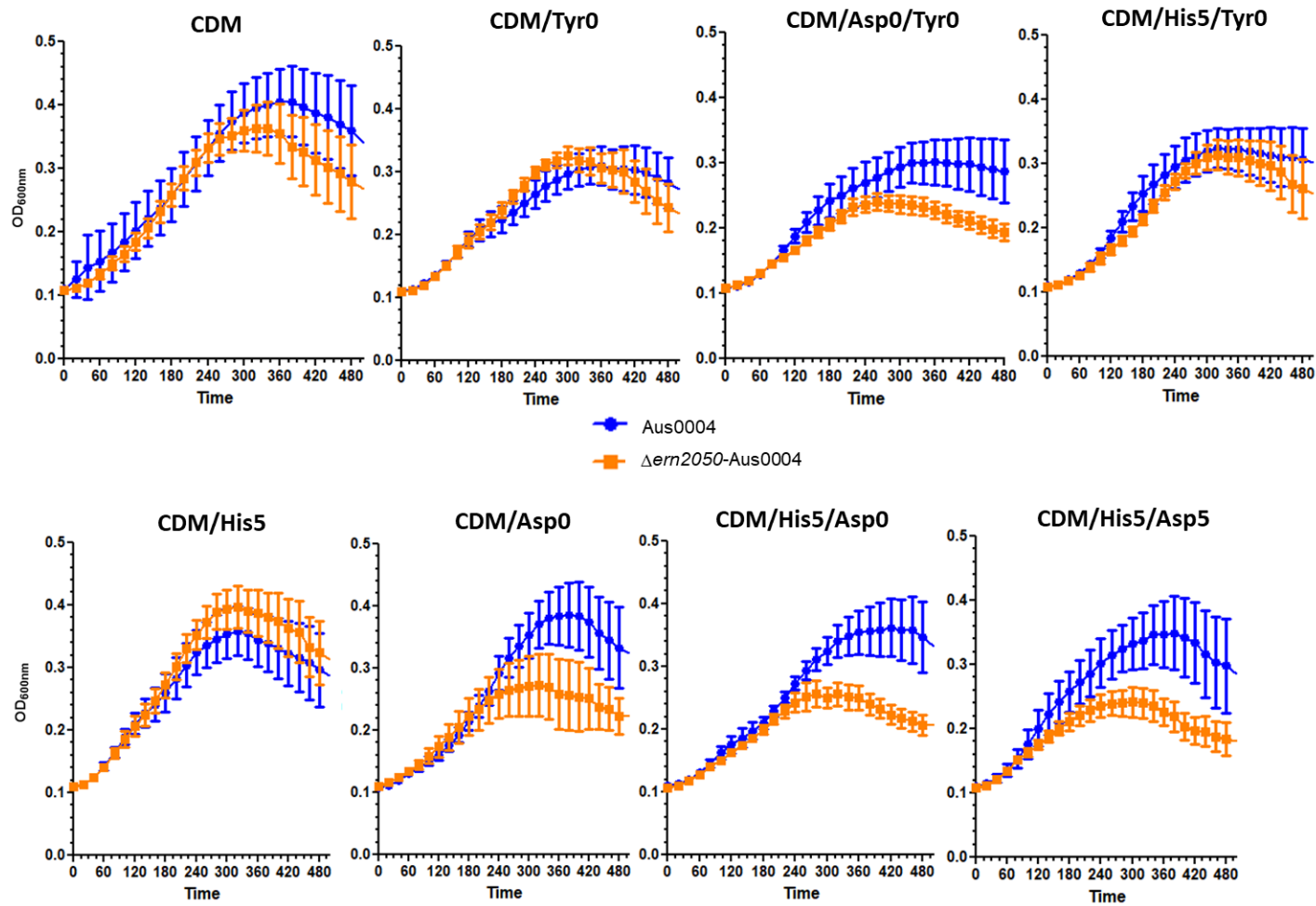
(A) Controlling radiolabeled transcript product *ern2050* and the three tRNA: *tRNA<sup>Asp</sup>*, *tRNA<sup>His</sup>* and *tRNA<sup>Tyr</sup>* on 10 % native gel.

(B) Scanned image of 6 % native gel after *in vitro* transcription binding assay. The *tRNA<sup>Tyr</sup>* was used as negative control. Free tRNA were go out of the gel.



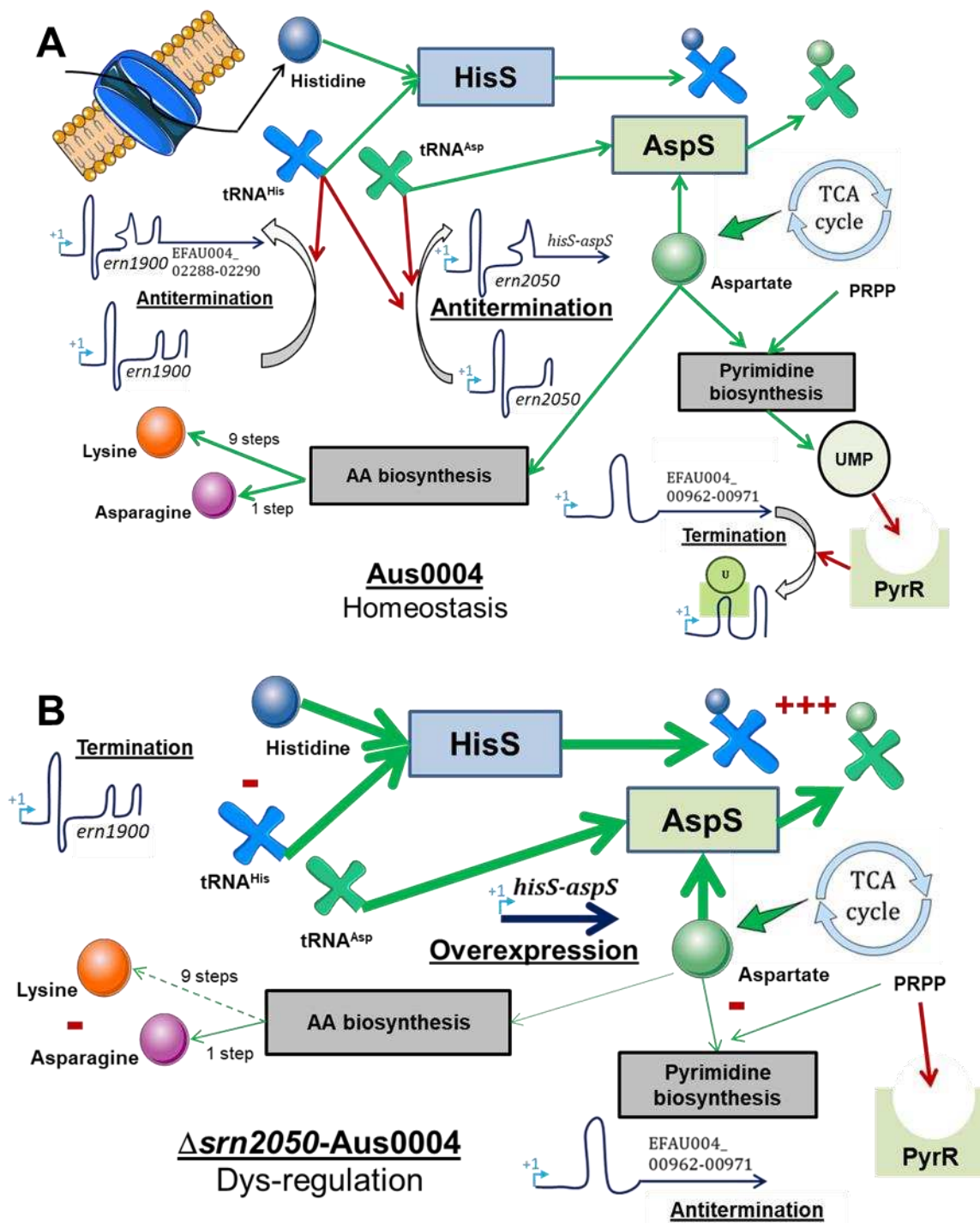
**Figure 4. Amino acid auxotrophic determination in *E. faecium* Aus0004 and  $\Delta$ ern2050-Aus0004 strains.**

Optical density (OD<sub>600nm</sub>) was measured after 8 hours incubation at 37°C. Omitted amino acid was indicated (-H: Histidine, -K: Lysine, -R: Arginine, -D: Aspartate, -E: Glutamate, -P: Proline, -Q: Glutamine, -N: Asparagine, -S: Serine, -C: Cysteine, -T: Threonine, -G: Glycine, -V: Valine, -I: Isoleucine, -L: Leucine, -M: Methionine, -A: Alanine, -Y: Tyrosine, -W: Tryptophan, -F: Phenylalanine)). CDM was the referential growth medium. A threshold (OD<sub>600nm</sub> > 0.8) was defined to distinguish growth from non-growth (red). A second threshold (OD<sub>600nm</sub> > 1.8) was definite to distinguish highly growth (green) from moderate growth (yellow). The error bars representing the range between maximum and minimum of 3 independent biological replicates. AA family was indicated (BCAA: Branched-chained amino-acid).



**Figure 5. Growth curve of Aus0004 and  $\Delta ern2050$ -Aus0004 strains in different medium.**

All growth medium was derived from CDM, with one or two aa with altered concentration: Tyr, His and Asp. The concentration (mg/L) of aa altered was indicated by the number. Blue and orange curves correspond to Aus0004 and  $\Delta ern2050$ -Aus0004 growth respectively. The 95% confidence interval was calculated from OD<sub>600nm</sub> measurements of 21 replicates (GraphPad Prism analysis).




**Figure 6. Regulation and metabolic pathways schematic representation.**

Green arrow corresponds to enzymatic reaction, red arrow corresponds to interaction between two compounds, blue curve represents RNA, large arrow represents major pathway, dotted arrow represents minor pathway.

(A) Regulation and metabolic network predicted in Aus0004 *E. faecium*.

(B) Regulation and metabolic network predicted in Δ*ern2050*-Aus0004 *E. faecium*.


## Supplementary figures

2519832 CAG**ATATCTC**TTGACTTTGAGACTGGTTTT**TAGTA**GGTTATAAGTAATAT  
-35 -10 5' 

2519782 CGAATGTCAATGAAAGAGCG**AGTA**AGTTTTTGAACTTTACA**AGAGAG**GA  
AGUAbox AGbox

2519732 TCATCGAG**GCTG**AAAATGATCTAAAGTTTCATTAGCCTAAAG**GACACT**CTG  
GNUGbox Codon

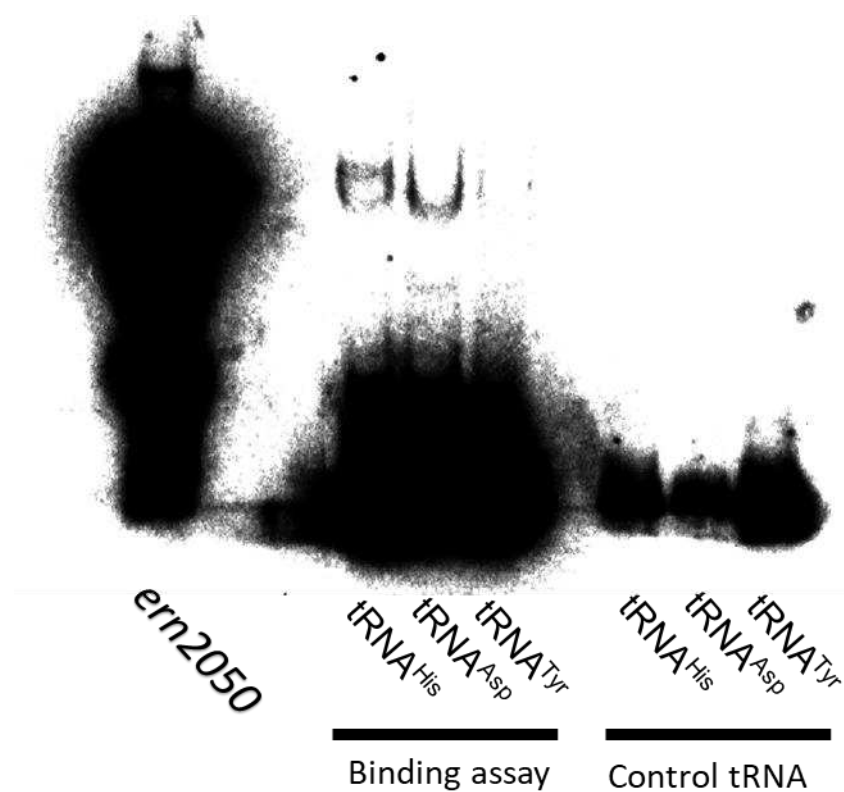
2519682 GAGATGGACGAGTGAAAAGCCGTCCCGCGTACGAG**GCGTTA**TTTCGTAAGAG  
F-box

2519632 GAAAGAGAAATCTTTCAAATTA**GGTGGTACCGTG**TATGAATTTTCATTCA  
3' T-box 

2519582 CCCTTGTTTCATGATGAACAAGGGTGATTTTTTTTATTTACATAGCTAAGAA  
-12 kcal/mol

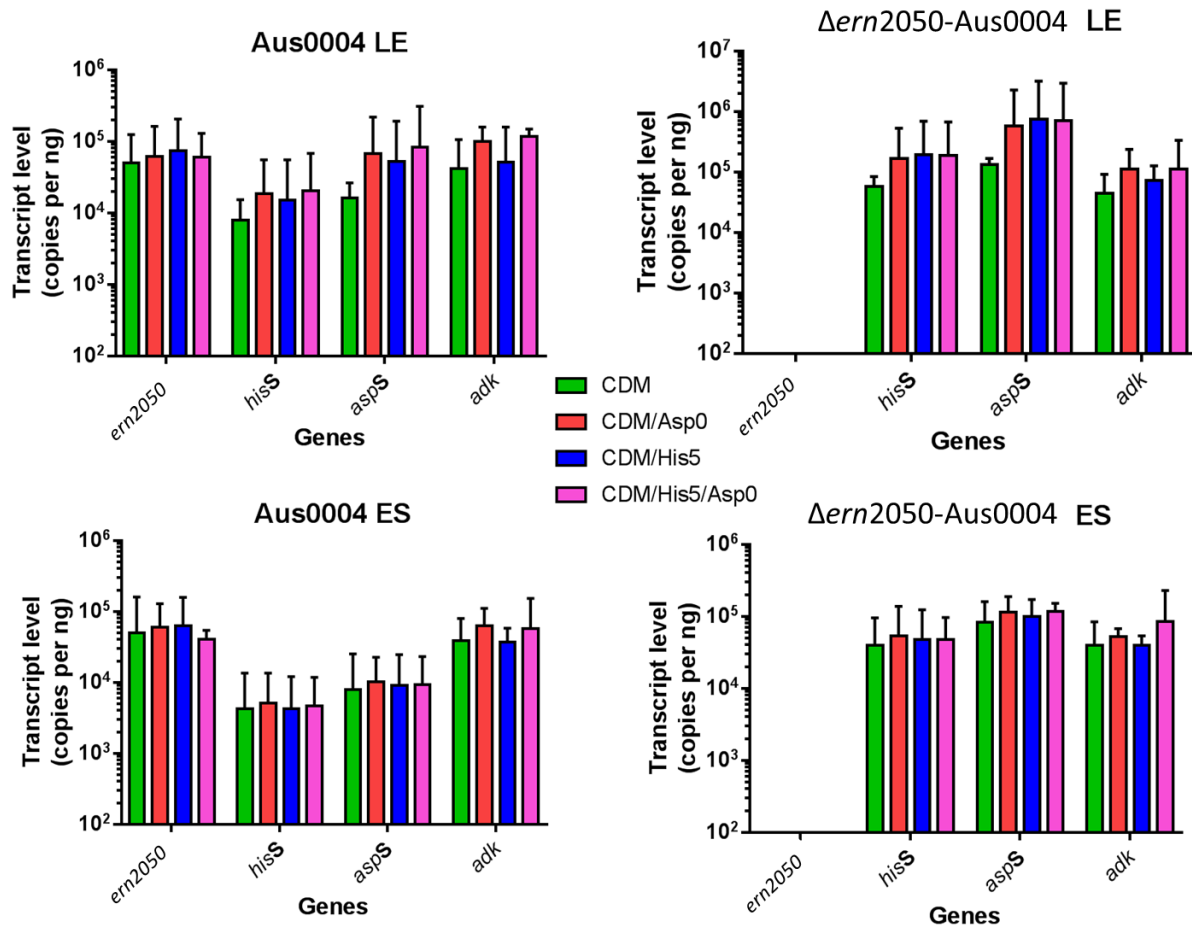
2519532 AGG**GAGAG**AAA**ACTATG**AGTTTCCAAAGACCAAAGGAACAAATGATATT  
RBS M S F Q R P K G T N D I  
HisS protein

**Supplementary Figure S1. Sequence of *ern2050* and flanked region.** 5'3' extremities determined by 5'3' RACE mapping were annotated by blue triangles. T-box consensus sequences were annotated by orange, green (specifier codon) or blue color. Promotor and initiator sequences were indicated in boldface type. Rho-independent terminator sequence predicted (ARNold finding terminator software (1)) was underlined and the free energy of stem-loop region was indicated.



**Supplementary Figure S2. Interaction between tRNA and *ern2050* during *in vitro* transcription.**

*ern2050* loading correspond to controlling radiolabeled transcript *ern2050*. Control tRNA loading corresponds to the three radiolabelled tRNA without *ern2050* product. Binding assay loading correspond to radiolabeled tRNA after binding assay with no-radiolabelled *ern2050* product. These loading were done on 10 % native gel.



**Supplementary Figure S3. Expression analysis of *ern2050*, *hisS* and *aspS* using RT-qPCR according to aa concentration in CDM.**

RNAs using were extracted from Aus0004 or  $\Delta ern2050$ -Aus0004 strains at late exponential (LE) and early stationary (ES) phase. The growth mediums used were the same than Figure 5. Adk expression was analysed as control. Error bars correspond to 95% confidence interval between 3 independent biological replicates. Statistical test used was Sidak test (p-value: \* < 0.05, \*\* < 0.01, \*\*\* < 0.001, \*\*\*\* < 0.0001).

## Supplementary methods

### qRT-PCR supplemental description

#### *Experimental design*

RNA was extracted from Aus0004 strains. Value studied was transcript level of RNA expressed by *E. faecium* at two different growth points (Late exponential and early stationary growth phase). Three genes were investigated: *ern2050* as deletion control, *hisS* and *aspS* as genes of interest. The *adk* gene was used as housekeeping gene to estimate the homogeneity between the samples. Primers used are listed in supplementary Table S2. Two factors were investigated: strain (Aus0004 and  $\Delta ern2050$ -Aus0004) and growth medium (CDM, H5, D0 and H5D0).

#### *RT-qPCR protocol*

Each sample was defined by three parameters: growth point, strain and growth medium. Two technical replicates were done for *ern2050*, *hisS* and *aspS* amplification and one replicate for *adk* amplification. For each gene, an estimation of limit of detection was estimated with sample without RNA. DNase treatment efficiency was estimated by sample without retro-transcriptase. *adk* primers were used for negative control (NTC).

Reverse transcription and amplification were done on one-step. Reaction volume was 20  $\mu$ L with 120 nM of forward and reverse primer (Table S2). The size of qPCR amplicons amounts to 70 nt (*ern2050*), 131 nt (*hisS*), 112 nt (*aspS*) and 62 nt (*adk*). Other composition description was available on Thermofischer scientific site. The qPCR instrument was StepOnePlus™ Real-Time PCR system (Thermofischer scientific) and qPCR data were extracted from StepOne Software v2.3. The reverse transcription was done by 30 minutes at 48°C and 10 minutes at 95°C follow up to amplification with 40 cycles (15 seconds at 95°C and 1 minute at 60°C). This program was concluding by a melt curve with a progressive increasing of 60°C to 95°C.

#### *RT-qPCR validation*

The standard conditions of standard curve to reach were  $r^2 > 0.99$  (acceptable to 0.96) and primer efficiency comprise between 80% to 110% (acceptable to 120%). No specific amplification was estimated with “no RNA” control. DNase treatment was estimated with NTC Ct values, was validated if Ct “NTC” > Ct “no RNA” and was accepted if Ct NTC >30. Deletion of *ern2050* was estimated with Ct values of  $\Delta ern2050$  strain *ern2050* amplification, was validated if Ct *ern2050* > Ct “NTC” and was accepted if Ct *ern2050* > 30. Repeatability was estimated for Ct value < 30 for *ern2050*, *aspS* and *hisS* genes and was validated with SD < 0.1 (acceptable to 0.5).

#### *Data analysis*

Ct values extracted were used for analysis. For quantitative expression, standard curve used DNA amplicons of gene analyzed and contain  $10^2$  to  $10^7$  copies ( $y = \log_{10}(\text{number of copies})$ ;  $x = \text{Ct values}$ ). Extremes values of curve could be excluded. By correlation, each value of technical replicates sample



were expressed by number of copies. These values were adjusted with RNA quantity measured by NanoDrop instrument (ThermoFischer scientific) and were expressed by number of copies per ng of RNA used in the sample.

For statistical test, values compared were the  $\log_{10}$  of copies number per RNA ng. Three biological replicates were used. For global analysis of the two factors investigated (Strain and growth medium), a two-way ANOVA test was used. The Sidak test was used for comparison between significant factors for each gene at growth point medium defined. Software used for statistical test was GraphPad Prism version 7. All raw data concerning RT-qPCR experiment was show in supplementary Table S8.

### **In-line probing assays.**

In-line probing was used to determine the secondary structure of the *ern2050* RNA. The experimental set up was adapted from Regulski,E.E. and Breaker,R.R. (2). *In vitro* transcribed RNA of either the full-length (247 nt) or a shortened (140 nt) *ern2050* version were used for in-line probing. Column-purified PCR products containing a T7 promoter sequence immediately upstream of the *ern2050* sequence were used as template for *in vitro* transcription. To increase transcription efficacy, the full-length *ern2050* template sequence contains two additional 5'-guanosines and the shortened *ern2050* one additional 5'-guanosine. Sequences of the resulting T7 transcripts are listed in Table S3.

*In vitro* transcription was carried out using the MEGAscript T7 Kit (Ambion, #AM1333) according to the manufacturer's instructions for short transcripts. Following, template DNA was digested by incubation with 1  $\mu$ l Turbo DNase for 15 min at 37 °C, reaction volume was increased to 100  $\mu$ l with RNase-free ddH<sub>2</sub>O and RNA was purified with 100  $\mu$ l phenol/chloroform/isoamylalcohol solution (P/C/I, 25:24:1) in a Phase Lock Gel (PLG) 'heavy' tube (5 Prime, #2302830). After vigorous shaking reaction was centrifuged for 12 min at 13,000 rpm and 15 °C and upper aqueous phase cleaned up via a G-25 column (illustra MicroSpin G-25 column, GE Healthcare, #27-5325-01) according to manufacturer's instructions. RNA was precipitated with 300  $\mu$ l 30:1 ethanol/sodium acetate mix for 2 to 3 h at -80 °C or overnight at -20 °C. When precipitated at -80 °C solution was thawed on ice for 20 to 30 min prior to centrifugation for 30 min at 14,000 rpm and 4 °C. RNA pellet was washed with 100  $\mu$ l 70-75 % ethanol, centrifuged for 10 min at 14,000 rpm and 4 °C and then dried at room temperature. To redissolve the RNA 33  $\mu$ l RNase-free ddH<sub>2</sub>O were added and solution was incubated for 5 min at 65 °C and 1,000 rpm in a heating block with shaking function and vortexed one to two times in between. RNA concentration and quality was determined by measuring a 1:10 diluted and an undiluted aliquot of the sample with a spectrophotometer (NanoDrop). 10  $\mu$ M stock solutions of the *in vitro* transcribed RNA were set up with RNase-free ddH<sub>2</sub>O for downstream application and stored at -80 °C until use.

End labelling of the *in vitro* transcribed RNA with radioactive [ $\gamma$ -<sup>32</sup>P]-ATP required preceding dephosphorylation to obtain a 5'-monophosphate. 20 pmol *in vitro* transcribed RNA were mixed with 15  $\mu$ l RNase-free ddH<sub>2</sub>O, incubated at 95 °C for 1 min and chilled on ice. Then 2  $\mu$ l 10x CIP (alkaline phosphatase calf intestinal, NEB) reaction buffer and 2  $\mu$ l CIP (NEB, #M0290) were added and

reaction was incubated for 1 h at 37 °C. Ensuing reaction volume was increased to 110 µl with RNase-free ddH<sub>2</sub>O and RNA purified with 110 µl phenol/chloroform/isoamylalcohol solution (25:24:1) in a PLG heavy tube. After vigorous shaking reaction was centrifuged for 12 min at 13,000 rpm and 15 °C and upper aqueous phase was transferred into a fresh tube. 1.1 µl GlycoBlue™ coprecipitant (15 mg/ml, Ambion) were added and RNA was precipitated with 300 µl 30:1 ethanol/sodium acetate mix for 3 h at -80 °C or overnight at -20 °C. RNA was cleaned up and redissolved in 16 µl RNase-free ddH<sub>2</sub>O as described above. RNA concentration was determined by measuring an undiluted aliquot of the sample with a spectrophotometer (NanoDrop).

20 pmol of dephosphorylated *in vitro* transcribed RNA mixed with RNase-free ddH<sub>2</sub>O to a final volume of 15 µl were incubated at 95 °C for 1 min and chilled on ice. Then 2 µl 10x PNK (T4 polynucleotide kinase, Thermo Scientific) buffer A, 1 µl PNK (Thermo Scientific, #EK0031) and 2 µl [ $\gamma$ -<sup>32</sup>P]-ATP (6000 Ci/mmol, 10 µCi/µl, Hartmann Analytic GmbH, Germany, #SRP-501) were added and incubated for 1 h at 37 °C. Labelling reaction was cleaned up via a G-25 column (illustra MicroSpin G-25 column, GE Healthcare) according to manufacturer's instructions and mixed with 20 µl 2x RNA gel loading dye. Sample was incubated at 95 °C for 1 min, chilled on ice and loaded in two lanes of a denaturing 6 % PAA gel with 7 M urea and deep wells. Electrophoresis proceeded at 300 V for 2 to 2.5 h in 1x TBE buffer at room temperature. Subsequently, bands of wanted RNA (entire transcript) were excised from gel, transferred into a reaction tube and 750 µl RNA elution buffer (0.1 M sodium acetate, 0.1 % (w/v) SDS, 10 mM EDTA) were added. RNA was eluted from gel by incubation at 8 °C and 850 rpm overnight in a cooling block with shaking function. Then reaction was briefly centrifuged, about 700 µl supernatant was transferred into a PLG heavy tube prefilled with 750 µl phenol/chloroform/isoamylalcohol solution (25:24:1) and vigorously shaken. After centrifugation for 12 min at 13,000 rpm and 15 °C upper aqueous phase was transferred equally into 2 fresh tubes and RNA was precipitated with 1 ml absolute ethanol for at least 3 h or overnight at -20 °C. RNA was spun down for 30 min at 13,000 rpm and 4 °C, pellet was air-dried and then redissolved in 25-50 µl RNase-free ddH<sub>2</sub>O. When the signal detected with the Geiger counter was less than 1,000 counts per second, the RNA was regarded as weakly labelled and double the quantity was used for the in-line probing reaction (*i.e.* 0.4 pmol). RNA concentration was determined by measuring an undiluted aliquot of the sample with a spectrophotometer (NanoDrop) and a 0.2 pmol/µl stock was set up with RNase-free ddH<sub>2</sub>O.

For the in-line probing reaction 0.2 pmol [ $\gamma$ -<sup>32</sup>P]-ATP labelled RNA was mixed with RNase-free ddH<sub>2</sub>O to a final volume of 5 µl and incubated at 95 °C for 1 min, then chilled on ice for 5 min and subsequently incubated at 37 °C for 15 min. After addition of 5 µl 2x in-line reaction buffer (100 mM Tris-HCl (pH 8.3), 40 mM MgCl<sub>2</sub>, 200 mM KCl) reaction was incubated for 40 h at room temperature. Then 10 µl 2x colourless gel loading solution (10 M urea, 1.5 mM EDTA (pH 8.0)) were added to in-line probing reaction and kept on ice until loading. For the ladder and control reactions 0.2 pmol [ $\gamma$ -<sup>32</sup>P]-ATP labelled RNA was used. To prepare the T1 ladder 0.2 pmol [ $\gamma$ -<sup>32</sup>P]-ATP labelled RNA was mixed with 8 µl RNA sequencing buffer (Ambion), incubated for 1 min at 95 °C and immediately chilled on ice for 2 min. Then 1 µl RNase T1 (0.1 U/µl, 1:10 dilution of stock) was set up with RNase-free

ddH<sub>2</sub>O, Ambion, #AM2283) was added, reaction incubated for 5 min at 37 °C and stopped by adding 12 µl 2x RNA gel loading dye. Ladder was kept on ice until loading.

To prepare the OH ladder 0.2 pmol [ $\gamma$ -<sup>32</sup>P]-ATP labelled RNA was mixed with 9 µl alkaline hydrolysis buffer (Ambion), incubated for 5 min at 95 °C and reaction stopped by adding 12 µl 2x RNA gel loading dye. Ladder was kept on ice until loading. To prepare the control (RNA integrity) reaction 0.2 pmol [ $\gamma$ -<sup>32</sup>P]-ATP labelled RNA was mixed with 9 µl RNase-free ddH<sub>2</sub>O, 12 µl 2x RNA gel loading dye was added and control reaction was kept on ice until loading. 10 µl of the in-line, ladder and control reactions were loaded per lane of a denaturing PAA sequencing gel with 7 M urea, respectively. 5, 6 and 10 % PAA gels were used to obtain satisfying size separation of the complete RNA molecule. One gel consisted of 50 ml PAA premix (5, 6 or 10 % of a 40 % PAA solution (19:1 acrylamide/bis-acrylamide), 7 M urea, 10 % 10x TBE buffer and purified ('VE') water), 500 µl 10 % APS (ammonium persulfate) and 50 µl TEMED (tetramethylethyldiamin). Gel was pre-warmed to 40-45 °C by running in 1x TBE buffer for 1 h at 40 W. Then samples were loaded and electrophoresis proceeded at 40 W for 1.5 to 3.5 h (depending on PAA percentage and region of the RNA molecule supposed to be well separated) in 1x TBE buffer at room temperature. Ensuing gel was transferred onto blotting paper (Whatman), dried for 45 min at 80 °C with a vacuum applied and exposed to a storage phosphor screen for 1 to 2 days. Screen was read out using the Typhoon™ FLA 7000 laser scanner (GE Healthcare).

## References

1. Naville,M., Ghuillot-Gaudeffroy,A., Marchais,A. and Gautheret,D. (2011) ARNold: a web tool for the prediction of Rho-independent transcription terminators. *RNA Biol*, **8**, 11–13.
2. Regulski,E.E. and Breaker,R.R. (2008) In-line probing analysis of riboswitches. *Methods Mol Biol*, **419**, 53–67.

## Supplementary Tables

**Supplementary Table S1. Bacterial strains and plasmids used in this study. (1–4)**

Strains and plasmid	Characteristics <sup>a</sup>	Reference
<b><i>E. faecium</i> strains</b>		
Aus0004	<i>E. faecium</i> Aus0004 reference strain <i>vanB</i> positive	(1)
$\Delta em2050$ -Aus0004	Aus0004 strain deleted of <i>em2050</i>	This study
$\Delta em2050$ -Aus0004_pAT29	Aus0004 strain deleted of <i>em2050</i> cloning with pAT29	This study
$\Delta sm2050$ -Aus0004_pAT29 $\Omega sm2050$	Aus0004 strain deleted of <i>em2050</i> trans-complemented with pAT29 $\Omega em2050$ plasmid	This study
<b><i>E. coli</i> strains</b>		
XL1blue	Using for 5'3'RACE mapping cloning	Stratagene
EC1000	Using for cloning	(2)
<b>Plasmids</b>		
pWS3	Suicid vector using for Knock-out deletion. <i>Spc</i> <sup>r</sup>	(3)
pAT29	High-copy-number shuttle vector using for cloning. <i>Spc</i> <sup>r</sup>	(4)
pAT29 $\Omega em2050$	pAT29 plasmid complemented by <i>em2050</i> with its native promotor and rho-independent terminator	This study
pGEMT	Using for 5'3'RACE mapping. <i>Amp</i> <sup>r</sup>	Promega
a. <i>Amp</i> <sup>r</sup> Ampicilline resistance; <i>Spc</i> <sup>r</sup> Spectinomycin resistance		

**Supplementary Table S2. Oligonucleotide primers and probes used in this study**

Experiment	Target	Name	Sequence (5'–3')	Position Aus0004	Strand
RACE mapping	<i>ern2050</i>	sRNA2050F1	GGAAAGAGAAATCTTTCAAATTAGG	2519608-2519633	–
		sRNA2050F2	ATTAGGTGGTACCGTGTATG	251594-2519613	–
		sRNA2050R1	CGAATAACGCTCGTACGCGG	2519639-2519658	+
		sRNA2050R2	GCGGGACGGCTTTTCACTCG	2519655-2519674	+
Knock-out		KOsARN2050F1	GATCCGAATTCGTTTGAAGAAATCAAGTGCC	2519056-2519075	+
		KOsARN2050F2	CTTAGCTATGTAAATAATTATAACCTACTAAAAACAGTCTC	2519790-2519814	+
		KOsARN2050R1	TTATTTACATAGCTAAGAAAGGGAGAGAAAACTATGAGTTTCC	2519509-2519551	–
		KOsARN2050R2	GCAAGAATTCATTTTTTGGTGTGCTTTCATC	2520296-2520317	–
		KOsARN2050vF	GGGGAAATAGTCAGATATCTCTTG	2519820-2519843	–
		KOsARN2050vR	CCTTTTGGTCTTTGGAAACTC	2519496-2519516	+
		Cloning	pCsARN2050F	GGTGGTGGATCCAGTTTTTCTCTCCCTTTCTTAG	2519519-2519539
pCsARN2050R			GGTGGTGAATTCAGCTGATTCCTCCTTTTGTCTAG	2519888-2519911	–
Northern Blot (Probes)	<i>ern2050</i>	NBern2050	TGAAAGATTTCCTCTTCCCTCTTACG	2519616-2519640	+
	tmRNA	NBtmRNA	TCGCTAGTCTGATGTTCTCTTTTTTC	1820406-1820430	–
Plasmid	pGEMT	M13F	GTAAACGACGGCCAGTG	///	///
		M13 Reverse short	CAGGAAACAGCTATGAC	///	///
	pWS3	pWS3F1	TTTCCCAGTCACGACGTGTG	///	///
		pWS3R1	ACTGACAGCTTCCAAGGAGC	///	///
	pAT29	pAT29vF	GTAAACGACGGCCAG	///	///
		pAT29vR	CAGGAAACAGCTATGAC	///	///
qPCR	<i>ern2050</i>	qPCRsRNA2050F	GAAATTCATACACGGTACCACCT	2519588-2519610	+
		qPCRsRNA2050R	CGCGTACGAGCGTTATTCG	2519639-2519657	–
	<i>aspS</i>	F-qPCRaspS	CGAATTCACCACTGGCCATT	2517926-2517945	+
		R-qPCRaspS	GCATGGGAAATCGCGGATAA	2518018-2518037	–
	<i>hisS</i>	F-qPCRHisS	TTCATGAAGCGCGCTTTAG	2518913-2518932	+
		R-qPCRHisS	CGCAAATCCGGCTAGTCATC	2519024-2519043	–
	<i>adk</i>	adk-F	CCGCACATCTCAACAGGAGA	108089-108108	+
		adk-R	AGGCCAAGAGCTGTTTCGTT	108131-108150	–
T7 Transcription	tRNA <sup>Asp</sup>	F-T7-tRNA-Asp	TAATACGACTCACTATAGGTTCCGTGGTGTAGGGGTT	1319837-1319856	–
		R-tRNA-Asp	TGGCGGTTCCGACGGGAATCGAAC	1319784-1319804	+
	tRNA <sup>His</sup>	F-T7-tRNA-His	TAATACGACTCACTATAGCGAATGTGGCGAAGTGGTT	1319051-1319070	–
		R-tRNA-His	TGGGCGGAATGATGGGAATCGAACC	1318997-1319019	+
	tRNA <sup>Tyr</sup>	F-T7-tRNA-Tyr	TAATACGACTCACTATAGGAGGAGTAGCGAAGTGGCT	1319652-1319671	–
		R-tRNA-Tyr	TGGTGGAGGAGAGTGGATTGGAAC	1319591-1319611	+
	<i>ern2050</i>	F-T7-ern2050	TAATACGACTCACTATAGGTAATATCGAATGTCAATGAAAG	2519767-2519789	–
		R-3'ern2050	TCACCCTTGTTTCATCATGAACAAGGGTGAATGAAATTC	2519557-2519594	+
Probing	T7 em_2050	FW218	TTAACTAATACGACTCACTATAGGGTAATATCGAATGTCAATGAAAG A	2519766-2519789	–
		FW219	TAAATAAAAAATCACCCTTGTTTCATCATGAACAA	2519545-2519579	+
	T7 short ern_2050	FW238	TTAACTAATACGACTCACTATAGGAGATGGACGAGTGAAAA	2519665-2519682	–
		FW219	TAAATAAAAAATCACCCTTGTTTCATCATGAACAA	2519545-2519579	+

**Supplementary Table S3. Sequences of T7 transcripts used for experiments**

Experiment	Name	Sequence (5'→3')
Interaction assay	<i>em_2050</i>	ggGUAAUAUCGAAUGUCAAGAAAGAGCGAGUAAGUUUUUGAAACUUUACAAGAGAGGAUCAUCGAGGCUGAAAAUGAUCUAAA GUUUCAUUAGCCUAAAAGACACUCUGGAGAUUGGACGAGUGAAAAGCCGUCGCCGCUACGAGCGUUUAUUCGUAAGAGGAAAGAGAA AUCUUUCAAAAUUAGGUGGUACCGUGUAUGAAUUUCAUUCACCCUUGUUCAUGAUGAACAAAGGGUGA
	tRNA <sup>Asp</sup>	GGUUCGUGGUGUAGGGGUUUAACGACGCGGUGUACGACGAGGAUCGCGGGUUCGAUUCGCGGAAACCGCCA
	tRNA <sup>His</sup>	GCGAAUGUGGCGAAGUGGUUAACGCACCGGAUUGUGGCUCCGGCAUUCGUGGGUUCGAUUCGCAUUCGCCCA
	tRNA <sup>Tyr</sup>	GGAGGAGUAGCGAAGUGGCUAAACGCGACGGACUGUAAAUCCGUUCCUAGGGUUCAGUGGUUCGAAUCCACUCUCCUCCACCA
In-line probing	<i>em_2050</i>	ggGUAAUAUCGAAUGUCAAGAAAGAGCGAGUAAGUUUUUGAAACUUUACAAGAGAGGAUCAUCGAGGCUGAAAAUGAUCUAAA GUUUCAUUAGCCUAAAAGACACUCUGGAGAUUGGACGAGUGAAAAGCCGUCGCCGCUACGAGCGUUUAUUCGUAAGAGGAAAGAGAA AUCUUUCAAAAUUAGGUGGUACCGUGUAUGAAUUUCAUUCACCCUUGUUCAUGAUGAACAAAGGGUGAUUUUUUUUAUUUA
	short <i>em_2050</i>	gGGAGAUGGACGAGUGAAAAGCCGUCGCCGCUACGAGCGUUUAUUCGUAAGAGGAAAGAGAAUUCUUUAAAAUUAGGUGGUACC GUGUAUGAAUUUCAUUCACCCUUGUUCAUGAUGAACAAAGGGUGAUUUUUUUUAUUUA

**Supplementary Table S4. Summary of RNA-seq coverage data**

Parameter	Aus0004 WT		$\Delta ern2050$ -Aus0004		$\Delta ern2050$ -Aus0004_pAT29		$\Delta ern2050$ -Aus0004_pAT29 $\Omega ern2050$	
	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2
Total number of reads	25451230	30224263	33999391	31134687	23034399	54046439	22293515	16595607
Number of mapped reads to the chromosome	22535523	27599878	24453037	27141453	14576612	24104821	8061749	9813742
Percentage of mapped reads to the chromosome (%)	88,5	91,3	71,9	87,2	63,3	44,6	36,2	59,1
Average size of reads (nt)	75	75	75	75	75	75	75	75
Average coverage of the chromosome <sup>a</sup> (X)	572	700	621	689	370	612	205	249
Number of reads mapped to rRNA genes	348352	446711	430390	324747	466036	1529737	590140	518312
Percentage of mapped to rRNA genes (%)	1,5	1,6	1,8	1,2	3,2	6,3	7,3	5,3
Number of reads mapped to CDSs	16752682	21034588	19362872	20342794	11513319	18730345	6159211	7597886
Percentage of reads mapped to CDSs (%)	74,3	76,2	79,2	75,0	79,0	77,7	76,4	77,4
Average number of reads per CDS <sup>b</sup>	5858	7355	6770	7113	4026	6549	2154	2657
Number of reads mapped to sRNA candidates	115133	128466	98148	112823	129385	277004	127329	114850
Percentage of reads mapped to sRNA candidates (%)	0,5	0,5	0,4	0,4	0,9	1,1	1,6	1,2
<sup>a</sup> The chromosome size of Aus0004 is 2,955,294 bp.								
<sup>b</sup> The number of predicted CDSs is 2,860.								

**Supplementary Table S5. Total RNA-seq data.** (A) Comparison  $\Delta ern2050$ -Aus0004 with Aus0004 strains. (B) Comparison  $\Delta ern2050$ -Aus0004\_pAT29 with  $\Delta ern2050$ -Aus0004\_pAT29 $\Omega ern2050$  strains.

(Me contacter à l'adresse suivante [Kline1@yahoo.fr](mailto:Kline1@yahoo.fr) si besoin d'obtenir le tableur excel contenant toutes les informations des « supplementary tables ».)

**Supplementary table S6. AA biosynthesis pathway.** List of gene predicted (High and medium confidence level) implicated in aa biosynthesis pathway by GapMind software (5). Essentials aa were indicated in red, semi-essential in orange and non-essential aa in green (experimental data).

Pathway	Gene AUS	Gene RS	5'	Protein	Enzyme	Name annotation	Mean RPKM WT	Confidence level	Substrate	Product	Step	EC
<b>Arginine</b>	EFAU004_00967	RS04780	988181	carA	carbamoyl phosphate synthase subunit alpha	carbamoyl-phosphate synthase small subunit	11,412	H	Glutamine	Carbamoyl-phosphate	5/8	6.3.5.5
<b>Arginine</b>	EFAU004_00966	RS04775	987100	carB	carbamoyl phosphate synthase subunit beta	carbamoyl-phosphate synthase large subunit	6,681	H	Glutamine	Carbamoyl-phosphate	5/8	6.3.5.5
<b>Arginine</b>	EFAU004_01735	RS08715	1763927	argI	ornithine carbamoyltransferase	ornithine carbamoyltransferase	18084,033	H	Ornithine	Citrulline	6/8	2.1.3.3
<b>Asparagine</b>	EFAU004_01761	RS08845	1788265	AsnB	asparagine synthase (glutamine-hydrolysing)	asparagine synthetase B	555,751	M	Aspartate	Asparagine	1/1	6.3.5.4
<b>BCCA</b>	EFAU004_01802	RS09045	1828221	ilvH	2-hydroxybutanoate synthase catalytic subunit	acetolactate synthase	357,525	M	S-adenosyl-L-methionine	S-adenosyl-L-homocysteine	1/9	2.1.1.6
<b>BCCA</b>	EFAU004_01027	RS05075	1049523	ilvE	isoleucine transaminase	branched-chain amino acid aminotransferase	30,836	H	4-methyl-2-oxopentanoate	Leucine	9/9	2.6.1.42
Chorismate	EFAU004_01582	RS07930	1598314	aroG	3-deoxy-7-phosphoheptulonate synthase	phospho-2-dehydro-3-deoxyheptonate aldolase	236,488	H	PEP	DAHP	1/7	2.5.1.54
Chorismate	EFAU004_01581	RS07925	1597259	aroB	3-dehydroquinate synthase	3-dehydroquinate synthase	76,967	H	DAHP	3-Dehydroquinate	2/7	4.2.3.4
Chorismate	EFAU004_01787	RS08965	1811539	aroD	3-dehydroquinate dehydratase	3-dehydroquinate dehydratase	52,082	H	3-Dehydroquinate	3-dehydroshikimate	3/7	4.2.1.10
Chorismate	EFAU004_00952	RS04700	971032	aroE	shikimate dehydrogenase	shikimate 5-dehydrogenase	62,761	H	3-dehydroshikimate	Shikimate	4/7	1.1.1.25
Chorismate	EFAU004_01577	RS07905	15922618	aroL	shikimate kinase	shikimate kinase	133,762	M	Shikimate	3-phosphoshikimate	5/7	2.7.1.71
Chorismate	EFAU004_01578	RS07910	1593906	aroA	3-phosphoshikimate 1-carboxyvinyltransferase	3-phosphoshikimate 1-carboxyvinyltransferase	115,681	H	3-phosphoshikimate	ESP	6/7	2.5.1.19
Chorismate	EFAU004_01580	RS07920	1596193	aroC	chorismate synthase	chorismate synthase	59,751	H	ESP	Chorismate	7/7	4.2.3.5
Chorismate	EFAU004_00695	RS03440	728417	fbp	6-deoxy-5-ketofructose 1-phosphate synthase	fructose-bisphosphate aldolase	761,566	H	glycerone phosphate	D-fructose 1,6-bisphosphate		4.1.2.13
Chorismate	EFAU004_02074	RS10390	2100417	tpiA	D-glyceraldehyde-3-phosphate phospholase	triosephosphate isomerase	3271,541	H	D-glyceraldehyde 3P	glycerone phosphate		5.3.1.1
<b>Cystéine</b>	EFAU004_02656	RS13340	2720386	cysE	serine acetyltransferase	Serine O-acetyltransferase	223,816	H	L-serine	O-acetyl-L-serine	1/2	2.3.1.30
<b>Cystéine</b>	EFAU004_01006	RS04970	1027689	cysK	O-acetylserine sulphydrylase	cysteine synthase A	282,531	H	O-acetyl-L-serine	L-cysteine	2/2	2.5.1.47
<b>Cystéine</b>	EFAU004_00464	RS02300	480152	CBS	cystathionine beta-synthase	cysteine synthase A	1423,166	H	Homocysteine	Cystathionine		4.2.1.22
<b>Glutamine</b>	EFAU004_00692	RS03425	723982	glnA	glutamine synthetase	glutamine synthetase, type I	9,009	H	Glutamate	Glutamine	1/1	6.3.1.2
<b>Glycine</b>	EFAU004_02299	RS11555	2337591	glyA	serine hydroxymethyltransferase	serine hydroxymethyltransferase	249,778	H	Serine	Glycine	1/1	2.1.2.1
<b>Histidine</b>	EFAU004_00537	RS02665	555940	prs	ribose-phosphate diphosphokinase	ribose-phosphate pyrophosphokinase	212,324	H	D-Ribose 5P	PRPP	1/11	2.7.6.1
<b>Histidine</b>	EFAU004_02632	RS13220	2694493	prs	ribose-phosphate diphosphokinase	Ribose-phosphate pyrophosphokinase 1	902,197	H	D-Ribose 5P	PRPP	1/11	2.7.6.1
<b>Histidine</b>	EFAU004_00302	RS01490	298783	HisN	histidinol-phosphate phosphatase	histidinol phosphate phosphatase HisJ family protein	0,000	H	Histidinol-P	Histidinol	8/11	3.1.3.15
<b>Histidine</b>	EFAU004_00869	RS04300	879713	HisN	histidinol-phosphate phosphatase	HisJ family histidinol phosphate phosphatase	0,000	H	Histidinol-P	Histidinol	8/11	3.1.3.15
<b>Lysine</b>	EFAU004_00472	RS02340	486986	asp kinase	aspartate kinase	aspartate kinase	28,678	H	Aspartate	4-P-L-Aspartate	1/9	2.7.2.4
<b>Lysine</b>	EFAU004_00471	RS02335	485988	asd	aspartate semi-aldehyde dehydrogenase	aspartate-semialdehyde dehydrogenase	41,830	H	4-P-L-Aspartate	Aspartate semi-aldehyde	2/9	1.2.1.11
<b>Lysine</b>	EFAU004_00468	RS02320	483677	dapA	4-hydroxy-tetrahydrodipicolinate synthase	dihydrodipicolinate synthase	59,609	H	Aspartate semi-aldehyde	L-2,3-dihydrodipicolinate	3/9	4.3.3.7
<b>Lysine</b>	EFAU004_00469	RS02325	484555	dapB	4-hydroxy-tetrahydrodipicolinate reductase	dihydrodipicolinate reductase	87,679	M	L-2,3-dihydrodipicolinate	L-piperidine-2,6-decarboxylate	4/9	1.17.1.8
<b>Lysine</b>	EFAU004_00470	RS02330	485277	dapH	tetrahydrodipicolinate acetyltransferase	4,4,5-tetrahydrodipyrroline-2,6-carboxylate N-succinyltransferase	57,434	H	L-piperidine-2,6-decarboxylate	L-2-acetamido-6-oxoheptanedioate	5/9	2.3.1.89
<b>Lysine</b>	EFAU004_00474	RS02350	489433	dapX	acetyl-diaminopimelate aminotransferase	aromatic amino acid aminotransferase	75,162	M	L-2-acetamido-6-oxoheptanedioate	N-acetyl-L-LL-2,6-diaminoheptanedioate	6/9	2.6.1.-
<b>Lysine</b>	EFAU004_00473	RS02345	488307	dapL	N-acetyl-diaminopimelate deacetylase	peptidase M20D	30,110	M	N-acetyl-LL-2,6-diaminoheptanedioate	LL-2,6-diaminoheptanedioate	7/9	3.5.1.47
<b>Lysine</b>	EFAU004_02072	RS10380	2097848	dapE	succinyl-diaminopimelate desuccinylase	succinyl-diaminopimelate desuccinylase	57,667	M	N-succinyl-LL-2,6-diaminoheptanedioate	LL-2,6-diaminoheptanedioate	7/9	3.5.1.18
<b>Lysine</b>	EFAU004_00467	RS02315	482874	dapF	diaminopimelate epimerase	diaminopimelate epimerase	45,355	H	LL-2,6-diaminoheptanedioate	meso-diaminoheptanedioate	8/9	5.1.1.7
<b>Lysine</b>	EFAU004_00466	RS02310	481628	lysA	diaminopimelate decarboxylase	diaminopimelate decarboxylase	39,197	H	meso-diaminoheptanedioate	L-Lysine	9/9	4.1.1.20
<b>Methionine</b>	EFAU004_00472	RS02340	486986	asp kinase	aspartate kinase	aspartate kinase	28,678	H	Aspartate	4-P-L-Aspartate	1/7	2.7.2.4
<b>Methionine</b>	EFAU004_00471	RS02335	485988	asd	aspartate semi-aldehyde dehydrogenase	aspartate-semialdehyde dehydrogenase	41,830	H	4-P-L-Aspartate	Aspartate semi-aldehyde	2/7	1.2.1.11
<b>Methionine</b>	EFAU004_00463	RS02295	479222	metB	cystathionine gamma-synthase	cystathionine gamma-synthase	599,599	H	O4-succinyl-L-homoserine	L-cystathionine	5/7	2.5.1.48
<b>Methionine</b>	EFAU004_02223	RS11170	2254852	metC	cystathionine beta-lyase	class II aminotransferase	73,491	H	L-cystathionine	L-Homocysteine	6/7	4.4.1.13
<b>Phenylalanine</b>	EFAU004_00196	RS00960	188384	cmuA	chorismate mutase	maltose O-acetyltransferase	226,272	H	Chorismate	prephenate	1/3	5.4.99.5
<b>Phenylalanine</b>	EFAU004_01576	RS07900	1592112	pdehyd	prephenate dehydratase	prephenate dehydratase	65,336	H	prephenate	phenylpyruvate	2/3	4.2.1.51
<b>Phenylalanine</b>	EFAU004_00474	RS02350	489433	ilvE	phenylalanine transaminase	aromatic amino acid aminotransferase	75,162	H	phenylpyruvate	Phenylalanine	3/3	2.6.1.57
<b>Proline</b>	EFAU004_02662	RS13370	2727500	proB	glutamate 5-kinase	glutamate 5-kinase	407,694	M	Glutamate	L-Glutamate-5P	1/3	2.7.2.11
<b>Proline</b>	EFAU004_02663	RS13375	2728303	proA	gamma-glutamyl phosphate reductase	gamma-glutamyl phosphate reductase	258,190	H	L-Glutamate-5P	L-glutamate 5-semialdehyde	2/3	1.2.1.41
<b>Proline</b>	EFAU004_02640	RS13260	2703849	proC	pyroline-5-carboxylate reductase	delta 1-pyrroline-5-carboxylate reductase	257,251	H	1-pyrroline-5-carboxylate	L-Proline	3/3	1.5.1.2
<b>Serine</b>	EFAU004_01503	RS07545	1520147	serA	3-phosphoglycerate dehydrogenase	4-phosphoglycerate dehydrogenase	54,578	M	3-phospho-D-glycerate	3-phospho-oxypyruvate	1/3	1.1.1.95
<b>Tyrosine</b>	EFAU004_00196	RS00960	188384	cmuA	chorismate mutase	maltose O-acetyltransferase	226,272	H	Chorismate	prephenate	1/3	5.4.99.5
<b>Tyrosine</b>	EFAU004_01579	RS07915	1595010	predehyd	prephenate dehydrogenase	prephenate dehydrogenase	90,021	H	prephenate	4-hydroxyphenylpyruvate	2/3	1.3.1.12
<b>Tyrosine</b>	EFAU004_00474	RS02350	489433	tyrB	tyrosine aminotransferase	aromatic amino acid aminotransferase	75,162	H	4-hydroxyphenylpyruvate	Tyrosine	3/3	2.6.1.57

**Supplementary table S7. Summary of RT-qPCR data.** Test statistics used was two-way ANOVA test with two factors analyzed (Strain and Medium) and log (copy/ng RNA sample) of three independents biologicals replicate as variable (Alpha risk = 0.05; p < \*0.05 \*\*0.01 \*\*\* 0.001 \*\*\*\* 0.0001). Fold change was calculated between  $\Delta em2050$ -Aus0004 and Aus0004 strains.

<b>LE phase</b>		<i>em2050</i>	<i>hisS</i>	<i>aspS</i>	<i>adk</i>
Fold Change $\Delta em2050$ -Aus0004 /Aus0004	CDM	NA	7,22	8,12	1,07
	CDM/His5	NA	12,86	14,35	1,42
	CDM/Asp0	NA	9,01	8,54	1,12
	CDM/His5/Asp0	NA	9,28	8,66	-1,06
2-way ANOVA	Interaction	NS (p 0.53)	NS (p 0.97)	NS (p 0.96)	NS (p 0.60)
	Medium	NS (p 0.80)	NS (p 0.79)	NS (p 0.46)	NS (p 0.14)
	Strain	**** (p < 0.0001)	*** (p 0.0004)	*** (p 0.0006)	NS (p 0.62)
<b>ES phase</b>		<i>em2050</i>	<i>hisS</i>	<i>aspS</i>	<i>adk</i>
Fold Change $\Delta em2050$ -Aus0004 /Aus0004	CDM	NA	9,29	10,45	1,02
	CDM/His5	NA	11,1	11,03	1,08
	CDM/Asp0	NA	10,6	11,34	-1,2
	CDM/His5/Asp0	NA	10,28	12,53	1,48
2-way ANOVA	Interaction	NS (p 0.82)	NS (p 0.99)	NS (p 0.99)	NS (p 0.74)
	Medium	NS (p 0.92)	NS (p 0.95)	NS (p 0.74)	NS (p 0.13)
	Strain	**** (p < 0.0001)	**** (p < 0.0001)	**** (p < 0.0001)	NS (p 0.66)



**Supplementary table S8. Total RT-qPCR data.** RT-qPCR data used for stastical test and transcriptional analysis.

Plate	Strain	Medium	Growth phase	Gene	Standard curve	r <sup>2</sup>	Efficiency	Min curve	Max curve	no RNA control	NTC	Mean Ct	SD Ct	Mean copie per RNA ng	RNA quantity (ng)
1	WT	CDM	ES	Adk	y = -0.28x + 11.10	0,9987	92,33%	NA	14,59	33,58	>40	17,27	NA	4,34E+04	35,6
2	WT	CDM	ES	Adk	y = -0.31x + 10.00	0,9949	102,13%	NA	9,64	>40	32,44	12,17	NA	5,29E+04	36
3	WT	CDM	ES	Adk	y = -0.30x + 10.04	0,9966	101,65%	25,84	9,81	>40	>40	13,53	NA	2,10E+04	39,4
1	Δ	CDM	ES	Adk	y = -0.28x + 11.10	0,9987	92,33%	NA	14,59	33,58	37,09	16,93	NA	5,21E+04	37,2
2	Δ	CDM	ES	Adk	y = -0.31x + 10.00	0,9949	102,13%	NA	9,64	>40	34,98	12,38	NA	4,77E+04	34,6
3	Δ	CDM	ES	Adk	y = -0.30x + 10.04	0,9966	101,65%	25,84	9,81	>40	>40	13,72	NA	1,93E+04	37,6
7	WT	CDM	LE	Adk	y = -0.28x + 9.97	0,9977	92,23%	28,38	10,41	>40	>40	14,95	NA	1,19E+04	45
7	WT	CDM	LE	Adk	y = -0.28x + 9.97	0,9977	92,23%	28,38	10,41	>40	31,93	12,78	NA	5,12E+04	43,4
7	WT	CDM	LE	Adk	y = -0.28x + 9.97	0,9977	92,23%	28,38	10,41	>40	38,55	12,54	NA	6,13E+04	42,4
7	Δ	CDM	LE	Adk	y = -0.28x + 9.97	0,9977	92,23%	28,38	10,41	>40	32,37	14,02	NA	2,35E+04	42
7	Δ	CDM	LE	Adk	y = -0.28x + 9.97	0,9977	92,23%	28,38	10,41	>40	37,13	12,96	NA	4,94E+04	39,8
7	Δ	CDM	LE	Adk	y = -0.28x + 9.97	0,9977	92,23%	28,38	10,41	>40	>40	12,57	NA	6,04E+04	42
1	WT	CDM	ES	AspS	y = -0.32x + 11.79	0,9969	106,57%	30,96	14,75	38,16	>40	22,13	1,49E-03	1,84E+03	35,6
2	WT	CDM	ES	AspS	y = -0.27x + 11.01	0,9999	84,86%	33,77	14,95	>40	32,44	19,71	5,28E-02	1,55E+04	36
3	WT	CDM	ES	AspS	y = -0.27x + 11.17	0,9998	85,80%	33,95	15,43	>40	>40	21,41	1,83E-02	6,55E+03	39,4
1	Δ	CDM	ES	AspS	y = -0.32x + 11.79	0,9969	106,57%	30,96	14,75	38,16	37,09	17,38	2,44E-01	5,55E+04	37,2
2	Δ	CDM	ES	AspS	y = -0.27x + 11.01	0,9999	84,86%	33,77	14,95	>40	34,98	16,48	4,59E-03	1,17E+05	34,6
3	Δ	CDM	ES	AspS	y = -0.27x + 11.17	0,9998	85,80%	33,95	15,43	>40	>40	17,51	3,07E-02	7,71E+04	37,6
7	WT	CDM	LE	AspS	y = -0.27x + 10.20	0,996	87,37%	29,47	11,5	37,12	>40	16,41	2,84E-02	1,19E+04	45
7	WT	CDM	LE	AspS	y = -0.27x + 10.20	0,996	87,37%	29,47	11,5	37,12	38,55	15,92	1,84E-03	1,72E+04	42,4
7	WT	CDM	LE	AspS	y = -0.27x + 10.20	0,996	87,37%	29,47	11,5	37,12	31,93	15,66	2,80E-02	1,98E+04	43,4
7	Δ	CDM	LE	AspS	y = -0.27x + 10.20	0,996	87,37%	29,47	11,5	37,12	32,37	12,88	5,39E-02	1,17E+05	42
7	Δ	CDM	LE	AspS	y = -0.27x + 10.20	0,996	87,37%	29,47	11,5	37,12	37,13	12,74	4,19E-03	1,35E+05	39,8
7	Δ	CDM	LE	AspS	y = -0.27x + 10.20	0,996	87,37%	29,47	11,5	37,12	>40	12,54	7,72E-03	1,45E+05	42
1	WT	CDM	ES	HisS	y = -0.27x + 10.46	0,9957	85,62%	31,94	13,22	36,85	>40	21,94	7,47E-03	1,03E+03	35,6
2	WT	CDM	ES	HisS	y = -0.26x + 10.72	0,9996	83,17%	33,08	14,2	35,84	32,44	19,96	7,86E-03	8,33E+03	36
3	WT	CDM	ES	HisS	y = -0.26x + 10.82	0,998	83,21%	33,11	14,5	33,22	>40	21,66	1,62E-01	3,38E+03	39,4
1	Δ	CDM	ES	HisS	y = -0.27x + 10.46	0,9957	85,62%	31,94	13,22	36,85	37,09	17,33	5,36E-02	1,70E+04	37,2
2	Δ	CDM	ES	HisS	y = -0.26x + 10.72	0,9996	83,17%	33,08	14,2	35,84	34,98	16,72	3,34E-02	6,17E+04	34,6
3	Δ	CDM	ES	HisS	y = -0.26x + 10.82	0,998	83,21%	33,11	14,5	33,22	>40	17,66	3,75E-02	3,97E+04	37,6
7	WT	CDM	LE	HisS	y = -0.25x + 10.67	0,9986	79,56%	33,83	14,38	30,61	>40	21,06	6,73E-02	4,66E+03	45
7	WT	CDM	LE	HisS	y = -0.25x + 10.67	0,9986	79,56%	33,83	14,38	30,61	38,55	20,15	4,39E-02	8,41E+03	42,4
7	WT	CDM	LE	HisS	y = -0.25x + 10.67	0,9986	79,56%	33,83	14,38	30,61	31,93	19,72	4,55E-02	1,06E+04	43,4
7	Δ	CDM	LE	HisS	y = -0.25x + 10.67	0,9986	79,56%	33,83	14,38	30,61	32,37	17,23	2,09E-02	4,69E+04	42
7	Δ	CDM	LE	HisS	y = -0.25x + 10.67	0,9986	79,56%	33,83	14,38	30,61	37,13	17,02	9,19E-02	5,60E+04	39,8
7	Δ	CDM	LE	HisS	y = -0.25x + 10.67	0,9986	79,56%	33,83	14,38	30,61	>40	16,59	4,23E-03	6,81E+04	42
1	WT	CDM	ES	ern2050	y = -0.31x + 11.84	0,9971	103,17%	31,96	15,59	34,25	>40	18,97	9,60E-03	2,80E+04	35,6
2	WT	CDM	ES	ern2050	y = -0.28x + 10.91	0,9995	88,47%	32,46	14,1	37,02	32,44	15,80	3,84E-02	1,01E+05	36
3	WT	CDM	ES	ern2050	y = -0.27x + 11.16	0,9998	85,73%	34,16	15,46	>40	>40	19,52	6,09E-01	2,15E+04	39,4
1	Δ	CDM	ES	ern2050	y = -0.31x + 11.84	0,9971	103,17%	31,96	15,59	34,25	37,09	34,19	1,61E-01	5,55E-01	37,2
2	Δ	CDM	ES	ern2050	y = -0.28x + 10.91	0,9995	88,47%	32,46	14,1	37,02	34,98	32,56	3,68E-01	2,59E+00	34,6
3	Δ	CDM	ES	ern2050	y = -0.27x + 11.16	0,9998	85,73%	34,16	15,46	>40	>40	33,36	4,71E-01	4,21E+00	37,6
7	WT	CDM	LE	ern2050	y = -0.27x + 11.14	0,996	85,53%	34	15,11	36,96	>40	19,28	8,12E-02	2,06E+04	45
7	WT	CDM	LE	ern2050	y = -0.27x + 11.14	0,996	85,53%	34	15,11	36,96	31,93	17,97	2,13E-02	4,81E+04	43,4
7	WT	CDM	LE	ern2050	y = -0.27x + 11.14	0,996	85,53%	34	15,11	36,96	38,55	17,16	1,24E-02	8,08E+04	42,4
7	Δ	CDM	LE	ern2050	y = -0.27x + 11.14	0,996	85,53%	34	15,11	36,96	>40	36,18	9,10E-01	6,93E-01	42
7	Δ	CDM	LE	ern2050	y = -0.27x + 11.14	0,996	85,53%	34	15,11	36,96	37,13	33,02	4,53E-01	4,86E+00	39,8
7	Δ	CDM	LE	ern2050	y = -0.27x + 11.14	0,996	85,53%	34	15,11	36,96	32,37	32,70	1,28E-01	5,52E+00	42
1	WT	D0	ES	Adk	y = -0.28x + 11.10	0,9987	92,33%	NA	14,59	33,58	32,37	16,41	NA	6,53E+04	41,6
2	WT	D0	ES	Adk	y = -0.31x + 10.00	0,9949	102,13%	NA	9,64	>40	33,12	11,42	NA	8,05E+04	40,2
3	WT	D0	ES	Adk	y = -0.30x + 10.04	0,9966	101,65%	25,84	9,81	>40	36,39	12,52	NA	4,15E+04	40,6
1	Δ	D0	ES	Adk	y = -0.28x + 11.10	0,9987	92,33%	NA	14,59	33,58	31,63	16,52	NA	5,79E+04	43,8
2	Δ	D0	ES	Adk	y = -0.31x + 10.00	0,9949	102,13%	NA	9,64	>40	>40	12,25	NA	5,21E+04	34,6
3	Δ	D0	ES	Adk	y = -0.30x + 10.04	0,9966	101,65%	25,84	9,81	>40	37,11	12,51	NA	4,54E+04	37,4
4	WT	D0	LE	Adk	y = -0.34x + 12.31	0,9785	119,00%	30,96	15,52	34,95	35,51	16,69	NA	1,10E+05	39,6
5	WT	D0	LE	Adk	y = -0.27x + 10.68	0,9977	84,40%	32,54	14,09	35,83	>40	15,15	NA	1,16E+05	39,4
6	WT	D0	LE	Adk	y = -0.27x + 10.88	0,9957	87,35%	32,38	NA	>40	>40	16,18	NA	7,28E+04	40,2
4	Δ	D0	LE	Adk	y = -0.34x + 12.31	0,9785	119,00%	30,96	15,52	34,95	36,59	16,19	NA	1,59E+05	41,2
5	Δ	D0	LE	Adk	y = -0.27x + 10.68	0,9977	84,40%	32,54	14,09	35,83	35,90	15,29	NA	1,15E+05	36,2
6	Δ	D0	LE	Adk	y = -0.27x + 10.88	0,9957	87,35%	32,38	NA	>40	>40	16,50	NA	5,94E+04	40,4
1	WT	D0	ES	AspS	y = -0.32x + 11.79	0,9969	106,57%	30,96	14,75	38,16	32,37	20,72	3,79E-02	4,38E+03	41,6
2	WT	D0	ES	AspS	y = -0.27x + 11.01	0,9999	84,86%	33,77	14,95	>40	33,12	19,80	5,51E-02	1,32E+04	40,2
3	WT	D0	ES	AspS	y = -0.27x + 11.17	0,9998	85,80%	33,95	15,43	>40	36,39	20,26	1,18E-02	1,30E+04	40,6
1	Δ	D0	ES	AspS	y = -0.32x + 11.79	0,9969	106,57%	30,96	14,75	38,16	31,63	16,61	1,63E-02	8,19E+04	43,8

Plate	Strain	Medium	Growth phase	Gene	Standard curve	r <sup>2</sup>	Efficiency	Min curve	Max curve	no RNA control	NTC	Mean Ct	SD Ct	Mean copie per RNA ng	RNA quantity (ng)
2	Δ	D0	ES	AspS	$y = -0.27x + 11.01$	0,9999	84,86%	33,77	14,95	>40	>40	16,27	5,09E-02	1,34E+05	34,6
3	Δ	D0	ES	AspS	$y = -0.27x + 11.17$	0,9998	85,80%	33,95	15,43	>40	37,11	16,66	4,09E-02	1,31E+05	37,4
4	WT	D0	LE	AspS	$y = -0.30x + 10.50$	0,9982	101,44%	27,94	11,19	37,11	35,51	15,13	2,37E-02	2,04E+04	39,6
5	WT	D0	LE	AspS	$y = -0.26x + 11.18$	0,9997	83,53%	34,74	15,88	>40	>40	18,65	5,44E-02	4,62E+04	39,4
6	WT	D0	LE	AspS	$y = -0.38x + 13.68$	0,9625	138,61%	31,34	NA	34,93	>40	18,39	2,87E-02	1,36E+05	40,2
4	Δ	D0	LE	AspS	$y = -0.30x + 10.50$	0,9982	101,44%	27,94	11,19	37,11	36,59	12,68	2,92E-02	1,11E+05	41,2
5	Δ	D0	LE	AspS	$y = -0.26x + 11.18$	0,9997	83,53%	34,74	15,88	>40	35,90	15,94	9,21E-04	2,60E+05	36,2
6	Δ	D0	LE	AspS	$y = -0.38x + 13.68$	0,9625	138,61%	31,34	NA	34,93	>40	15,74	1,87E-02	1,36E+06	40,4
1	WT	D0	ES	HisS	$y = -0.27x + 10.46$	0,9957	85,62%	31,94	13,22	36,85	32,37	21,43	5,38E-02	1,20E+03	41,6
2	WT	D0	ES	HisS	$y = -0.26x + 10.72$	0,9996	83,17%	33,08	14,2	35,84	33,12	19,93	6,62E-02	7,61E+03	40,2
3	WT	D0	ES	HisS	$y = -0.26x + 10.82$	0,998	83,21%	33,11	14,5	33,22	36,39	20,54	2,16E-02	6,44E+03	40,6
1	Δ	D0	ES	HisS	$y = -0.27x + 10.46$	0,9957	85,62%	31,94	13,22	36,85	31,63	17,21	8,11E-03	1,56E+04	43,8
2	Δ	D0	ES	HisS	$y = -0.26x + 10.72$	0,9996	83,17%	33,08	14,2	35,84	>40	16,54	9,48E-04	6,86E+04	34,6
3	Δ	D0	ES	HisS	$y = -0.26x + 10.82$	0,998	83,21%	33,11	14,5	33,22	37,11	16,56	2,47E-02	7,75E+04	37,4
4	WT	D0	LE	HisS	$y = -0.26x + 10.25$	0,9975	80,99%	32,21	12,92	36,38	35,36	19,25	1,10E-03	4,96E+03	39,2
5	WT	D0	LE	HisS	$y = -0.31x + 11.69$	0,9933	102,57%	30,94	15,03	35,90	>40	19,23	1,89E-02	1,57E+04	39,4
6	WT	D0	LE	HisS	$y = -0.35x + 12.86$	0,9872	123,70%	31,32	NA	36,91	>40	19,23	1,01E-02	3,42E+04	40,2
4	Δ	D0	LE	HisS	$y = -0.26x + 10.25$	0,9975	80,99%	32,21	12,92	36,38	>40	16,61	6,52E-02	2,33E+04	40
5	Δ	D0	LE	HisS	$y = -0.31x + 11.69$	0,9933	102,57%	30,94	15,03	35,90	35,90	16,07	8,31E-03	1,59E+05	36,2
6	Δ	D0	LE	HisS	$y = -0.35x + 12.86$	0,9872	123,70%	31,32	NA	36,91	>40	16,48	5,13E-03	3,14E+05	40,4
1	WT	D0	ES	ern2050	$y = -0.31x + 11.84$	0,9971	103,17%	31,96	15,59	34,25	32,37	18,59	2,01E-03	3,13E+04	41,6
2	WT	D0	ES	ern2050	$y = -0.28x + 10.91$	0,9995	88,47%	32,46	14,1	37,02	33,12	15,88	5,34E-02	8,56E+04	40,2
3	WT	D0	ES	ern2050	$y = -0.27x + 11.16$	0,9998	85,73%	34,16	15,46	>40	36,39	17,65	6,29E-02	6,42E+04	40,6
1	Δ	D0	ES	ern2050	$y = -0.31x + 11.84$	0,9971	103,17%	31,96	15,59	34,25	31,63	33,31	4,53E-01	9,01E-01	43,8
2	Δ	D0	ES	ern2050	$y = -0.28x + 10.91$	0,9995	88,47%	32,46	14,1	37,02	>40	>40	NA	2,29E-02	34,6
3	Δ	D0	ES	ern2050	$y = -0.27x + 11.16$	0,9998	85,73%	34,16	15,46	>40	37,11	32,11	1,27E-01	9,02E+00	37,4
4	WT	D0	LE	ern2050	$y = -0.30x + 11.18$	0,9998	97,80%	30,88	14,07	>40	35,36	17,62	3,52E-02	2,32E+04	39,2
5	WT	D0	LE	ern2050	$y = -0.28x + 11.28$	0,9969	91,71%	32,52	14,95	37,00	>40	16,45	2,60E-02	1,03E+05	39,4
6	WT	D0	LE	ern2050	$y = -0.30x + 11.74$	0,9924	99,18%	32,78	NA	36,47	>40	17,93	9,51E-03	5,93E+04	40,2
4	Δ	D0	LE	ern2050	$y = -0.30x + 11.18$	0,9998	97,80%	30,88	14,07	>40	>40	35,13	1,02E+00	1,67E-01	40
5	Δ	D0	LE	ern2050	$y = -0.28x + 11.32$	0,9969	91,71%	32,52	14,95	37,00	35,90	30,15	4,84E-01	1,53E+01	36,2
6	Δ	D0	LE	ern2050	$y = -0.30x + 11.74$	0,9924	99,18%	32,78	NA	36,47	>40	33,79	9,53E-02	1,06E+00	40,4
1	WT	H5	ES	Adk	$y = -0.28x + 11.10$	0,9987	92,33%	NA	14,59	33,58	33,64	17,58	NA	2,78E+04	45,6
2	WT	H5	ES	Adk	$y = -0.31x + 10.00$	0,9949	102,13%	NA	9,64	>40	37,12	12,50	NA	4,43E+04	34,2
3	WT	H5	ES	Adk	$y = -0.30x + 10.04$	0,9966	101,65%	25,84	9,81	>40	32,26	12,69	NA	3,94E+04	38
1	Δ	H5	ES	Adk	$y = -0.28x + 11.10$	0,9987	92,33%	NA	14,59	33,58	32,24	16,96	NA	4,61E+04	41,2
2	Δ	H5	ES	Adk	$y = -0.31x + 10.00$	0,9949	102,13%	NA	9,64	>40	35,79	12,71	NA	3,56E+04	36,6
3	Δ	H5	ES	Adk	$y = -0.30x + 10.04$	0,9966	101,65%	25,84	9,81	>40	>40	12,70	NA	3,84E+04	38,6
4	WT	H5	LE	Adk	$y = -0.34x + 12.31$	0,9785	119,00%	30,96	15,52	34,95	35,36	17,44	NA	6,04E+04	39,2
5	WT	H5	LE	Adk	$y = -0.27x + 10.68$	0,9977	84,40%	32,54	14,09	35,83	36,08	15,68	NA	8,88E+04	37,2
6	WT	H5	LE	Adk	$y = -0.27x + 10.88$	0,9957	87,35%	32,38	NA	>40	36,97	20,50	NA	4,52E+03	43
4	Δ	H5	LE	Adk	$y = -0.34x + 12.31$	0,9785	119,00%	30,96	15,52	34,95	>40	17,23	NA	6,81E+04	40
5	Δ	H5	LE	Adk	$y = -0.27x + 10.68$	0,9977	84,40%	32,54	14,09	35,83	36,09	15,40	NA	9,65E+04	40,6
6	Δ	H5	LE	Adk	$y = -0.27x + 10.88$	0,9957	87,35%	32,38	NA	>40	35,32	16,75	NA	5,40E+04	38
1	WT	H5	ES	AspS	$y = -0.32x + 11.79$	0,9969	106,57%	30,96	14,75	38,16	33,64	21,85	3,21E-03	1,75E+03	45,6
2	WT	H5	ES	AspS	$y = -0.27x + 11.01$	0,9999	84,86%	33,77	14,95	>40	37,12	20,10	1,36E-02	1,29E+04	34,2
3	WT	H5	ES	AspS	$y = -0.27x + 11.17$	0,9998	85,80%	33,95	15,43	>40	32,26	20,45	1,67E-02	1,23E+04	38
1	Δ	H5	ES	AspS	$y = -0.32x + 11.79$	0,9969	106,57%	30,96	14,75	38,16	32,24	16,99	2,71E-02	6,62E+04	41,2
2	Δ	H5	ES	AspS	$y = -0.27x + 11.01$	0,9999	84,86%	33,77	14,95	>40	35,79	16,31	9,09E-03	1,23E+05	36,6
3	Δ	H5	ES	AspS	$y = -0.27x + 11.17$	0,9998	85,80%	33,95	15,43	>40	>40	16,92	4,75E-02	1,08E+05	38,6
4	WT	H5	LE	AspS	$y = -0.30x + 10.50$	0,9982	101,44%	27,94	11,19	37,11	35,36	17,79	1,73E-02	3,13E+03	39,2
5	WT	H5	LE	AspS	$y = -0.26x + 11.18$	0,9997	83,53%	34,74	15,88	>40	36,08	19,02	2,26E-03	3,92E+04	37,2
6	WT	H5	LE	AspS	$y = -0.38x + 13.68$	0,9625	138,61%	31,34	NA	34,93	36,97	18,53	2,88E-02	1,13E+05	43
4	Δ	H5	LE	AspS	$y = -0.30x + 10.50$	0,9982	101,44%	27,94	11,19	37,11	>40	12,97	1,21E-02	8,80E+04	40
5	Δ	H5	LE	AspS	$y = -0.26x + 11.18$	0,9997	83,53%	34,74	15,88	>40	36,09	15,68	2,67E-03	2,71E+05	40,6
6	Δ	H5	LE	AspS	$y = -0.38x + 13.68$	0,9625	138,61%	31,34	NA	34,93	35,32	15,44	3,95E-02	1,87E+06	38
1	WT	H5	ES	HisS	$y = -0.27x + 10.46$	0,9957	85,62%	31,94	13,22	36,85	33,64	22,19	3,45E-02	6,86E+02	45,6
2	WT	H5	ES	HisS	$y = -0.26x + 10.72$	0,9996	83,17%	33,08	14,2	35,84	37,12	20,47	1,16E-01	6,45E+03	34,2
3	WT	H5	ES	HisS	$y = -0.26x + 10.82$	0,998	83,21%	33,11	14,5	33,22	32,26	20,84	1,89E-01	5,74E+03	38
1	Δ	H5	ES	HisS	$y = -0.27x + 10.46$	0,9957	85,62%	31,94	13,22	36,85	32,24	17,68	2,24E-02	1,24E+04	41,2
2	Δ	H5	ES	HisS	$y = -0.26x + 10.72$	0,9996	83,17%	33,08	14,2	35,84	35,79	16,55	4,20E-02	6,46E+04	36,6
3	Δ	H5	ES	HisS	$y = -0.26x + 10.82$	0,998	83,21%	33,11	14,5	33,22	>40	16,77	3,89E-02	6,61E+04	38,6
4	WT	H5	LE	HisS	$y = -0.26x + 10.25$	0,9975	80,99%	32,21	12,92	36,38	35,51	21,91	4,09E-02	1,01E+03	39,6
5	WT	H5	LE	HisS	$y = -0.31x + 11.69$	0,9933	102,57%	30,94	15,03	35,90	36,08	19,73	2,28E-02	1,17E+04	37,2
6	WT	H5	LE	HisS	$y = -0.35x + 12.86$	0,9872	123,70%	31,32	NA	36,91	36,97	19,21	6,00E-02	3,28E+04	43
4	Δ	H5	LE	HisS	$y = -0.26x + 10.25$	0,9975	80,99%	32,21	12,92	36,38	36,59	17,03	1,22E-02	1,76E+04	41,2

Plate	Strain	Medium	Growth phase	Gene	Standard curve	r <sup>2</sup>	Efficiency	Min curve	Max curve	no RNA control	NTC	Mean Ct	SD Ct	Mean copie per RNA ng	RNA quantity (ng)
5	Δ	H5	LE	HisS	$y = -0.31x + 11.69$	0,9933	102,57%	30,94	15,03	35,90	36,09	15,92	4,14E-03	1,57E+05	40,6
6	Δ	H5	LE	HisS	$y = -0.35x + 12.86$	0,9872	123,70%	31,32	NA	36,91	35,32	16,22	1,70E-02	4,11E+05	38
1	WT	H5	ES	ern2050	$y = -0.31x + 11.84$	0,9971	103,17%	31,96	15,59	34,25	33,64	18,96	2,97E-02	2,20E+04	45,6
2	WT	H5	ES	ern2050	$y = -0.28x + 10.91$	0,9995	88,47%	32,46	14,1	37,02	37,12	15,95	1,93E-02	9,62E+04	34,2
3	WT	H5	ES	ern2050	$y = -0.27x + 11.16$	0,9998	85,73%	34,16	15,46	>40	32,26	17,57	7,51E-02	7,20E+04	38
1	Δ	H5	ES	ern2050	$y = -0.31x + 11.84$	0,9971	103,17%	31,96	15,59	34,25	32,24	33,99	7,05E-03	5,73E-01	41,2
2	Δ	H5	ES	ern2050	$y = -0.28x + 10.91$	0,9995	88,47%	32,46	14,1	37,02	35,79	35,18	4,02E-01	4,68E-01	36,6
3	Δ	H5	ES	ern2050	$y = -0.27x + 11.16$	0,9998	85,73%	34,16	15,46	>40	>40	34,08	5,26E-01	2,64E+00	38,6
4	WT	H5	LE	ern2050	$y = -0.30x + 11.18$	0,9998	97,80%	30,88	14,07	>40	35,51	17,39	5,85E-02	2,70E+04	39,6
5	WT	H5	LE	ern2050	$y = -0.28x + 11.27$	0,9969	91,71%	32,52	14,95	37,00	36,08	16,17	7,47E-02	1,31E+05	37,2
6	WT	H5	LE	ern2050	$y = -0.30x + 11.74$	0,9924	99,18%	32,78	NA	36,47	36,97	17,73	9,14E-02	6,35E+04	43
4	Δ	H5	LE	ern2050	$y = -0.30x + 11.18$	0,9998	97,80%	30,88	14,07	>40	36,59	34,30	8,75E-01	2,77E-01	41,2
5	Δ	H5	LE	ern2050	$y = -0.28x + 11.31$	0,9969	91,71%	32,52	14,95	37,00	36,09	33,45	5,85E-02	1,58E+00	40,6
6	Δ	H5	LE	ern2050	$y = -0.30x + 11.74$	0,9924	99,18%	32,78	NA	36,47	35,32	36,01	7,92E-01	2,61E-01	38
1	WT	H5D0	ES	Adk	$y = -0.28x + 11.10$	0,9987	92,33%	NA	14,59	33,58	34,33	15,89	NA	1,02E+05	37,4
2	WT	H5D0	ES	Adk	$y = -0.31x + 10.00$	0,9949	102,13%	NA	9,64	>40	31,46	12,64	NA	3,40E+04	40,2
3	WT	H5D0	ES	Adk	$y = -0.30x + 10.04$	0,9966	101,65%	25,84	9,81	>40	>40	12,62	NA	3,65E+04	43
1	Δ	H5D0	ES	Adk	$y = -0.28x + 11.10$	0,9987	92,33%	NA	14,59	33,58	35,24	15,18	NA	1,52E+05	40
2	Δ	H5D0	ES	Adk	$y = -0.31x + 10.00$	0,9949	102,13%	NA	9,64	>40	>40	12,20	NA	4,91E+04	38
3	Δ	H5D0	ES	Adk	$y = -0.30x + 10.04$	0,9966	101,65%	25,84	9,81	>40	>40	12,22	NA	5,48E+04	38
4	WT	H5D0	LE	Adk	$y = -0.34x + 12.31$	0,9785	119,00%	30,96	15,52	34,95	>40	16,68	NA	1,14E+05	38
5	WT	H5D0	LE	Adk	$y = -0.27x + 10.68$	0,9977	84,40%	32,54	14,09	35,83	36,12	15,36	NA	1,07E+05	37,6
6	WT	H5D0	LE	Adk	$y = -0.27x + 10.88$	0,9957	87,35%	32,38	NA	>40	34,97	15,35	NA	1,31E+05	37,6
4	Δ	H5D0	LE	Adk	$y = -0.34x + 12.31$	0,9785	119,00%	30,96	15,52	34,95	35,84	16,09	NA	1,96E+05	35
5	Δ	H5D0	LE	Adk	$y = -0.27x + 10.68$	0,9977	84,40%	32,54	14,09	35,83	>40	15,17	NA	1,20E+05	37,6
6	Δ	H5D0	LE	Adk	$y = -0.27x + 10.88$	0,9957	87,35%	32,38	NA	>40	36,62	18,36	NA	1,75E+04	42,6
1	WT	H5D0	ES	AspS	$y = -0.32x + 11.79$	0,9969	106,57%	30,96	14,75	38,16	34,33	21,40	8,54E-03	2,98E+03	37,4
2	WT	H5D0	ES	AspS	$y = -0.27x + 11.01$	0,9999	84,86%	33,77	14,95	>40	31,46	19,82	2,93E-02	1,30E+04	40,2
3	WT	H5D0	ES	AspS	$y = -0.27x + 11.17$	0,9998	85,80%	33,95	15,43	>40	>40	20,27	6,45E-03	1,22E+04	43
1	Δ	H5D0	ES	AspS	$y = -0.32x + 11.79$	0,9969	106,57%	30,96	14,75	38,16	35,24	16,44	1,47E-02	1,02E+05	40
2	Δ	H5D0	ES	AspS	$y = -0.27x + 11.01$	0,9999	84,86%	33,77	14,95	>40	>40	16,20	3,42E-02	1,28E+05	38
3	Δ	H5D0	ES	AspS	$y = -0.27x + 11.17$	0,9998	85,80%	33,95	15,43	>40	>40	16,74	3,56E-02	1,23E+05	38
4	WT	H5D0	LE	AspS	$y = -0.30x + 10.50$	0,9982	101,44%	27,94	11,19	37,11	>40	16,48	1,12E-02	8,16E+03	38
5	WT	H5D0	LE	AspS	$y = -0.26x + 11.18$	0,9997	83,53%	34,74	15,88	>40	36,12	18,47	5,37E-02	5,39E+04	37,6
6	WT	H5D0	LE	AspS	$y = -0.38x + 13.68$	0,9625	138,61%	31,34	NA	34,93	34,97	18,12	8,17E-03	1,84E+05	37,6
4	Δ	H5D0	LE	AspS	$y = -0.30x + 10.50$	0,9982	101,44%	27,94	11,19	37,11	35,84	12,81	5,56E-03	1,16E+05	35
5	Δ	H5D0	LE	AspS	$y = -0.26x + 11.18$	0,9997	83,53%	34,74	15,88	>40	>40	15,85	1,32E-02	2,64E+05	37,6
6	Δ	H5D0	LE	AspS	$y = -0.38x + 13.68$	0,9625	138,61%	31,34	NA	34,93	36,62	15,39	2,82E-02	1,75E+06	42,6
1	WT	H5D0	ES	HisS	$y = -0.27x + 10.46$	0,9957	85,62%	31,94	13,22	36,85	34,33	21,36	4,02E-02	1,39E+03	37,4
2	WT	H5D0	ES	HisS	$y = -0.26x + 10.72$	0,9996	83,17%	33,08	14,2	35,84	31,46	20,20	2,05E-02	6,44E+03	40,2
3	WT	H5D0	ES	HisS	$y = -0.26x + 10.82$	0,998	83,21%	33,11	14,5	33,22	>40	20,50	2,51E-02	6,23E+03	43
1	Δ	H5D0	ES	HisS	$y = -0.27x + 10.46$	0,9957	85,62%	31,94	13,22	36,85	35,24	16,55	3,72E-02	2,57E+04	40
2	Δ	H5D0	ES	HisS	$y = -0.26x + 10.72$	0,9996	83,17%	33,08	14,2	35,84	>40	16,66	2,80E-02	5,82E+04	38
3	Δ	H5D0	ES	HisS	$y = -0.26x + 10.82$	0,998	83,21%	33,11	14,5	33,22	>40	16,94	3,96E-02	6,06E+04	38
4	WT	H5D0	LE	HisS	$y = -0.26x + 10.25$	0,9975	80,99%	32,21	12,92	36,38	>40	20,22	5,12E-02	2,88E+03	38
5	WT	H5D0	LE	HisS	$y = -0.31x + 11.69$	0,9933	102,57%	30,94	15,03	35,90	36,12	19,23	3,23E-02	1,65E+04	37,6
6	WT	H5D0	LE	HisS	$y = -0.35x + 12.86$	0,9872	123,70%	31,32	NA	36,91	34,97	19,09	1,84E-03	4,10E+04	37,6
4	Δ	H5D0	LE	HisS	$y = -0.26x + 10.25$	0,9975	80,99%	32,21	12,92	36,38	35,84	16,38	8,62E-02	3,04E+04	35
5	Δ	H5D0	LE	HisS	$y = -0.31x + 11.69$	0,9933	102,57%	30,94	15,03	35,90	>40	16,37	4,05E-02	1,24E+05	37,6
6	Δ	H5D0	LE	HisS	$y = -0.35x + 12.86$	0,9872	123,70%	31,32	NA	36,91	36,62	16,09	3,68E-02	4,06E+05	42,6
1	WT	H5D0	ES	ern2050	$y = -0.31x + 11.84$	0,9971	103,17%	31,96	15,59	34,25	34,33	18,19	4,90E-03	4,64E+04	37,4
2	WT	H5D0	ES	ern2050	$y = -0.28x + 10.91$	0,9995	88,47%	32,46	14,1	37,02	31,46	17,05	4,01E-02	4,09E+04	40,2
3	WT	H5D0	ES	ern2050	$y = -0.27x + 11.16$	0,9998	85,73%	34,16	15,46	>40	>40	18,51	4,90E-02	3,56E+04	43
1	Δ	H5D0	ES	ern2050	$y = -0.31x + 11.84$	0,9971	103,17%	31,96	15,59	34,25	35,24	33,66	2,93E-01	7,58E-01	40
2	Δ	H5D0	ES	ern2050	$y = -0.28x + 10.91$	0,9995	88,47%	32,46	14,1	37,02	>40	38,45	2,19E+00	8,44E-02	38
3	Δ	H5D0	ES	ern2050	$y = -0.27x + 11.16$	0,9998	85,73%	34,16	15,46	>40	>40	33,51	8,95E-02	3,73E+00	38
4	WT	H5D0	LE	ern2050	$y = -0.30x + 11.18$	0,9998	97,80%	30,88	14,07	>40	>40	17,03	5,23E-02	3,60E+04	38
5	WT	H5D0	LE	ern2050	$y = -0.28x + 11.30$	0,9969	91,71%	32,52	14,95	37,00	36,12	16,71	3,53E-02	9,07E+04	37,6
6	WT	H5D0	LE	ern2050	$y = -0.30x + 11.74$	0,9924	99,18%	32,78	NA	36,47	34,97	18,18	1,57E-02	5,34E+04	37,6
4	Δ	H5D0	LE	ern2050	$y = -0.30x + 11.18$	0,9998	97,80%	30,88	14,07	>40	35,84	35,86	1,65E+00	1,37E-01	35
5	Δ	H5D0	LE	ern2050	$y = -0.28x + 11.33$	0,9969	91,71%	32,52	14,95	37,00	>40	33,38	6,87E-01	1,86E+00	37,6
6	Δ	H5D0	LE	ern2050	$y = -0.30x + 11.74$	0,9924	99,18%	32,78	NA	36,47	36,62	32,60	1,11E-01	2,27E+00	42,6

**Supplementary table S9. Specifier codon sequence comparison between different bacterial species.** Specific codon of T-box RNA leader of *hisS-aspS* operon was identified in different bacterial species (6). Alternative codon was annotated by prime symbol.

Specie	Strain	Id	Strand	5'position	Sequence	Codon 1	Codon 2
<i>Enterococcus faecium</i>	Aus0004	CP003351.1	-	2519789	GACAC	Asp	His
<i>Bacillus subtilis</i>	ATCC11774	CP026010.1	-	3760318	AACAC	Asn	His
<i>Bacillus stearothermophilus</i>	10	CP008934.1	-	802269	GACAC	Asp	His
<i>Bacillus cereus</i>	ATCC14579	CP034551.1	-	4340932	GACAC	Asp	His
<i>Bacillus halodurans</i>	C-125	BA000004.3	+	1345310	AACAC	Asn	His
<i>Oceanobacillus iheyensis</i>	HTE831	BA000028.3	-	2047383	GACAT	Asp	His'
<i>Listeria monocytogenes</i>	ATCC51779	CP025567	-	1574194	GACAC	Asp	His
<i>Lactobacillus casei</i>	12A	CP006690.1	-	1494835	GACAC	Asp	His
<i>Lactobacillus brevis</i>	100D8	CP015338.1	-	2102143	GACAC	Asp	His
<i>Lactobacillus johnsonii</i>	G2A	CP040854.1	+	769431	GACAC	Asp	His
<i>Enterococcus faecalis</i>	V583	AE016830.1	-	1910447	GACAT	Asp	His'
<i>Clostridium thermocellum</i>	ATCC27405	CP000568.1	-	1622720	GACAC	Asp	His
<i>Carboxydotherrmus hydrogenoformans</i>	Z-2901	CP000141.1	-	1987895	AACAC	Asn	His
<i>Thermoanaerobacter tengcongensis</i>	MB4	AE008691.1	+	1231030	GACAC	Asp	His
<i>Bacillus clausii</i>	KSM-K16	AP006627.1	+	1698282	AACAC	Asn	His
<i>Bacillus licheniformis</i>	ATCC14580	CP034569.1	-	2789821	AACAC	Asn	His
<i>Desulfotomaculum reducens</i>	MI-1	CP000612.1	+	807921	AACAC	Asn	His
<i>Exiguobacterium sibiricum</i>	255-15	CP001022.1	-	2041444	GACCC	Asp	Pro
<i>Lactobacillus plantarum</i>	ATCC8014	CP024413.1	-	1775842	GACGC	Asp	Arg

#### References (supplementary Tables)

1. Lam,M.M.C., Seemann,T., Bulach,D.M., Gladman,S.L., Chen,H., Haring,V., Moore,R.J., Ballard,S., Grayson,M.L., Johnson,P.D.R., *et al.* (2012) Comparative analysis of the first complete *Enterococcus faecium* genome. *J. Bacteriol.*, **194**, 2334–2341.
2. Leenhouts,K., Buist,G., Bolhuis,A., ten Berge,A., Kiel,J., Mierau,I., Dabrowska,M., Venema,G. and Kok,J. (1996) A general system for generating unlabelled gene replacements in bacterial chromosomes. *Mol. Gen. Genet.*, **253**, 217–224.
3. Zhang,X., Vrijenhoek,J.E.P., Bonten,M.J.M., Willems,R.J.L. and van Schaik,W. (2011) A genetic element present on megaplasms allows *Enterococcus faecium* to use raffinose as carbon source. *Environmental Microbiology*, **13**, 518–528.
4. Trieu-Cuot,P., Carlier,C., Poyart-Salmeron,C. and Courvalin,P. (1990) A pair of mobilizable shuttle vectors conferring resistance to spectinomycin for molecular cloning in *Escherichia coli* and in gram-positive bacteria. *Nucleic Acids Res.*, **18**, 4296.
5. Price,M.N., Deutschbauer,A.M. and Arkin,A.P. (2020) GapMind: Automated Annotation of Amino Acid Biosynthesis. *mSystems*, **5**.
6. Vitreschak,A.G., Mironov,A.A., Lyubetsky,V.A. and Gelfand,M.S. (2008) Comparative genomic analysis of T-box regulatory systems in bacteria. *RNA*, **14**, 717–735.

## IV. Discussion

Cette première investigation fonctionnelle a permis de préparer les premières pistes pour une caractérisation plus approfondie des petits ARN décrits chez *E. faecium*. Nous avons ici au moins 9 candidats avec des hypothèses à explorer, dont 3 ARN avec des données expérimentales sur leurs fonctions [Tableau 10]. Nous allons revenir sur les différentes expérimentations à effectuer pour approfondir l'étude fonctionnelle de ces ARN.

Tout d'abord, l'expression des 6 ARN n'ayant pas encore été validée par Northern blot et RT-qPCR devra être vérifiée. Ensuite, la caractérisation génétique des 16 ARN sélectionnés lors de ce manuscrit devra être complétée par 5'3'RACE mapping. Pour les ARN présentant des incohérences de taille entre les résultats obtenus par 5'3'RACE mapping et par Northern blot, l'utilisation d'un plasmide comme pAT29 pourrait être une solution. Ce plasmide semblant être à haut niveau de copies, les expérimentations de 5'3'RACE mapping sur des souches *trans*-complémentées devraient être plus aisées, en confirmant en parallèle l'expression de l'ARN dans ces souches par Northern-Blot. Des essais de demi-vie sur chacun des 16 ARN semblent indispensables pour finaliser la première caractérisation génétique de ces ARN. En effet, ces données sont nécessaires pour évaluer le type de régulateur et estimer si ces données sont cohérentes avec les hypothèses posées lors de ce manuscrit [Tableau 10]. Des extraits d'ARN totaux ont déjà été obtenus dans la souche Aus0004 avec ou en absence de concentration subinhibitrice d'antiseptiques dans les mêmes conditions que l'essai de demi-vie avec Aus0004<sup>TR</sup> décrit dans l'article 1 (Le Neindre *et al.*, 2021) mais ne sont pas encore exploités. Une investigation en tant qu'ARNreg agissant en *trans* pourra donc être sérieusement envisagée pour les ARN ayant une demi-vie d'au moins dix minutes. Malheureusement, les prédictions de cibles par bio-informatique ne sont pas informatives seules et ne permettent pas de distinguer un régulateur en 5'UTR d'un ARNreg. En effet, le nombre de cibles prédites reste important quelque soit l'ARN étudié. Ces données nécessitent d'être couplées à des données

expérimentales. L'identification globale des cibles peut être effectuée par l'approche MAPS (Lalaouna et Massé, 2015). Cette approche repose sur une chromatographie d'affinité couplée à un séquençage haut débit [Figure 43]. Pour cela, l'ARN étudié est fusionné avec une étiquette MS2 puis cloné dans la souche d'intérêt. Pour augmenter la sensibilité, l'étiquette peut être doublée. Pour améliorer la spécificité, plusieurs ARN fusionnés pourront être produits, avec des fusions en 5' et/ou en 3'. Lors de la croissance bactérienne, l'ARN-MS2 transcrit peut interagir avec diverses cibles. Au temps de croissance souhaité, les bactéries sont lysées, et le lysat est élué dans une colonne avec une phase immobile contenant une résine associée aux protéines MS2-MBP. Ces protéines fixent les étiquettes MS2. Les complexes ARN-MS2 avec leurs cibles sont donc retenus sur la colonne. On peut alors ajouter un éluant pour récupérer les ARN (phénol pH = 4). Les ARN sont ensuite identifiés par séquençage à haut débit permettant d'obtenir une liste de cibles de l'ARN étudié. Cette approche couplée aux prédictions informatiques est un bon moyen d'identifier les cibles d'un ARNreg. Les cibles les plus intéressantes devront être confirmées par des retards sur gel. Le principe d'un retard sur gel est de mettre en évidence l'interaction directe *in vitro* entre l'ARNreg et sa cible en comparant la différence de migration sur un gel natif entre un complexe et l'ARN seul. En s'appuyant sur les prédictions d'interaction, les zones d'interaction peuvent être testées en produisant des transcrits partiellement délétés des séquences prédites. La perte d'interaction suite à cette délétion ciblée confirmera la zone d'interaction.

Au vu des données obtenues, l'approche MAPS avec confirmation des cibles par retard sur gel peut déjà être effectuée avec les ARN *ern0120* et *ern0160*. Selon la fonction des cibles validées, des tests phénotypiques comparant les souches d'étude pourront être réalisés pour estimer l'impact de la régulation au niveau phénotypique. Selon la zone d'interaction, le type de régulation (post-transcriptionnelle ou traductionnelle) sera à confirmer.

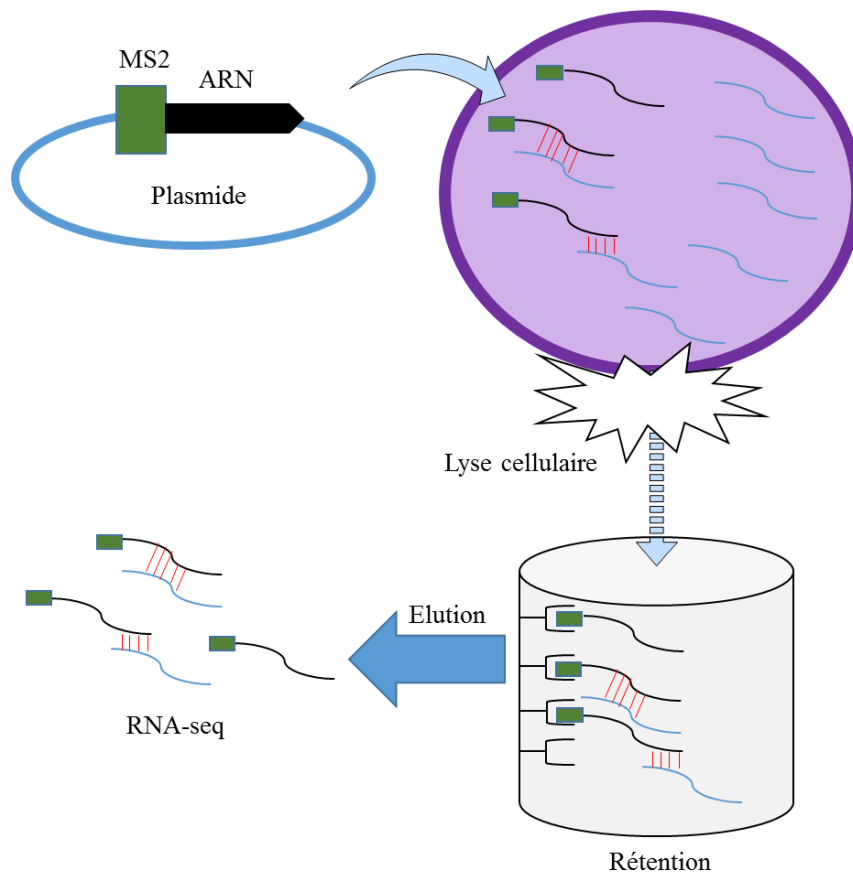


Figure 43. Représentation de l'approche MAPS permettant l'identification des cibles d'un ARNreg.  
L'ARNreg est représenté en noir, tandis que les ARNm bactériens sont représentés en bleu.  
L'étiquette MS2 est représentée par le rectangle vert.

La régulation post-transcriptionnelle peut être évaluée par RT-qPCR et/ou RNAseq en comparant les souches d'étude et par un système double-plasmide [Figure 44]. Une différence d'expression entre la souche sauvage et délétée de l'ARN étudié (ou *trans*-complémentée et plasmide vide) du gène cible permet la mise en évidence d'une régulation post-transcriptionnelle directe ou indirecte. Le système double-plasmide permet de mettre en évidence la régulation directe de l'ARN régulateur étudié sur l'ARN cible [Figure 44]. Le principe repose sur l'utilisation de deux plasmides, l'un inductible contenant l'ARN régulateur, et le second contenant un gène rapporteur de type *gfp* ou *blaZ* soumis à la région promotrice de l'ARNm cible. Si l'induction de l'expression de l'ARN

régulateur induit une modification de l'activité mesurée (enzymatique, fluorescence ou autre), alors l'ARN a bien une activité de régulation fonctionnelle. La régulation traductionnelle peut être évaluée par des essais de TOE print. Un essai de TOE print permet de révéler la fixation d'un ribosome sur l'ARNm [Figure 44]. Une régulation traductionnelle est fortement probable dans le cas où l'ajout de l'ARN régulateur modifie la position du TOE print. Cette méthode présente également l'intérêt de valider la localisation de l'interaction sur l'ARNm.

Dans le cas d'*ern0120*, la co-transcription avec le gène en aval devra être explorée par RT-PCR. Si cette co-transcription est avérée, confirmant donc la présence d'*ern0120* dans la région 3'UTR de EFAU004\_00212, l'acteur provoquant la scission d'*ern0120* devra être recherché. Si les données expérimentales (MAPS et retard sur gel) confirment la ribonucléase HIII comme cible, la fonction de cette enzyme dans la scission d'*ern0120* devra être explorée. Pour cela, la production de la protéine RnhC et de l'ARN comprenant EFAU004\_00212 et *ern0120* est nécessaire. On pourra ensuite, sur un gel dénaturant, observer si l'ajout de la ribonucléase provoque l'apparition d'*ern0120* à partir du transcrit total.



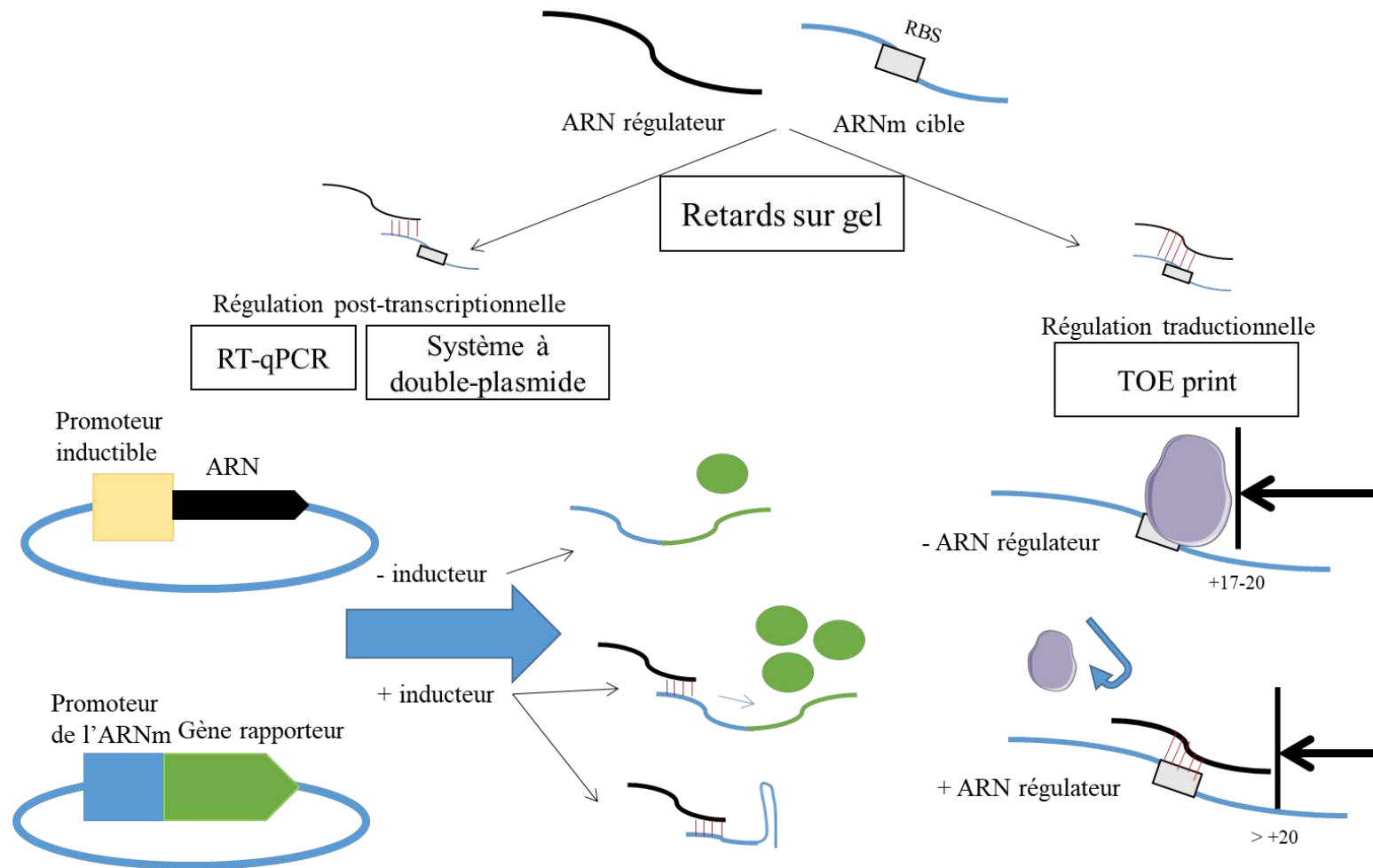


Figure 44. Représentation des méthodes de validation des cibles d'ARNreg.

L'ARNreg est représenté en noir, l'ARNm cible en bleu et l'ARN du gène rapporteur en vert. La protéine du gène rapporteur est représentée par le cercle vert. Le ribosome est représenté par la protéine violette. Le RBS de l'ARNm cible est représenté par le rectangle gris.

Pour l'investigation d'*ern0160*, plusieurs pistes peuvent être envisagées. Tout d'abord, l'interaction avec l'ARN en antisens (*ern0170*) peut être expérimentée par retard sur gel. Comme *ern0170* se situe en aval du gène *mgtA*, leur co-transcription va devoir être évaluée (RT-PCR et/ou Northern blot). Malheureusement, dans les conditions de cultures utilisées, *ern0170* et *mgtA* sont faiblement exprimés, rendant plus difficile l'investigation fonctionnelle. En utilisant un système rapporteur avec un plasmide à haut niveau de copies, il pourrait être possible de forcer la transcription d'*ern0170*. On pourrait alors utiliser un système double-plasmide pour vérifier si *ern0160* influe sur la continuité de transcription d'*ern0170*. Une stabilisation d'*ern0170* par *ern0160* peut également être envisagée. En plus de l'activité en *cis*, une activité en *trans* d'*ern0160* semble plausible dû à sa grande stabilité [Figure 31]. Les données de transcriptomique globale comparant la souche sauvage avec  $\Delta ern0160$  et  $\Delta ern0160\_pAT29$  avec  $\Delta ern0160\_pAT29\Omega ern0160$  semblent cohérentes avec une activité en *trans* (données non publiées). En effet, contrairement à ce qui a été observé avec les données de transcriptomique globale des souches d'étude d'*ern2050*, la souche surexprimant *ern0160* cause une variation d'expression significative d'un grand nombre de gènes (>300). Néanmoins, cette observation peut être causée par un ensemble de régulations directes et indirectes, nécessitant d'effectuer une recherche des cibles directes par l'approche MAPS. Rappelons également que le plasmide  $pAT29\Omega ern0160$  surexprime *ern0170*, devant donc être pris en compte dans les interprétations futures. Le rôle fonctionnel d'*ern0160* a été étudié dans un modèle murin de colonisation intestinale. Les premiers résultats suggèrent une possible implication du *ern0160* dans la colonisation intestinale mais des études complémentaires sont nécessaires afin d'approfondir ces premiers résultats (Reissier *et al.* 2021).

Pour les autres ARN à demi-vie plus courte, l'investigation en tant qu'ARNreg agissant en *trans* n'est pas à favoriser dans un premier temps. Seuls les ARN *ern0030*, *ern1260* et

*ern2050* appartiennent à cette catégorie à ce jour, dû à l'absence de données sur les demi-vies des autres ARN candidats.

Récemment, de nouveaux petits ARN ont été mis en évidence chez *E. faecium* (souche Aus0004), avec des prédictions bio-informatiques supplémentaires sur leurs mécanismes de régulation (Michaux *et al.*, 2020). Cinq de nos 16 ARN sélectionnés n'ont pas été mis en évidence par ces auteurs (*ern0510*, *ern0980*, *ern1060*, *ern2050* et *ern2210*). En rapport avec ce manuscrit, on peut distinguer 6 de leurs ARN prédits comme T-box régulateurs avec sRNA\_016 régulant *tyrS*, sRNA\_095 régulant *valRS*, sRNA\_100 correspondant à l'ARN *ern1900* régulant l'opéron *hisXYZ*, sRNA\_110 régulant *leuS* et deux derniers (sRNA\_022 et sRNA\_026 régulant des transporteurs). Ces auteurs mettent également en évidence en tant que riboswitch, sRNA\_025 correspondant à *ern0690* et prédit comme un riboswitch avec un motif « ykok » (Michaux *et al.*, 2020). Le motif « ykok » est retrouvé dans les riboswitchs ayant comme molécule effectrice  $Mg^{2+}$  (McCown *et al.*, 2017). Ce type de riboswitch est notamment retrouvé dans la régulation des gènes *mgtA* et *mgtE*, et pourrait donc correspondre au mécanisme de régulation médié par *ern0170*. Le riboswitch médié par PyrR a également été mis en évidence en tant que sRNA\_043 (Michaux *et al.*, 2020). On retrouve sRNA\_002 prédit comme 5'UTR correspondant à *ern0030*, soit le régulateur de *tet(M)*. Ils ont considéré sRNA\_069 comme un ARN intergénique alors que cet ARN correspond à *ern1300*, régulateur de *mef(A)*. Ils ont localisé *ern2410* sur deux positions, avec sRNA\_116 en 5'UTR de EFAU004\_02833 codant un transporteur ABC et sRNA\_068 en région intergénique (EFAU004\_01582 codant AroGA et EFAU004\_01583 codant la petite sous-unité d'une  $\beta$ -galactosidase). Enfin, ils ont identifié sRNA\_067, correspondant à *ern1260*, comme un analogue à BsrC, ARN non codant exprimé chez *B. subtilis* (Saito *et al.*, 2009).

L'ARN BsrC est exprimé exclusivement à un temps de croissance de 4h, avec une taille de 280 nucléotides et l'extrémité 5' a été identifiée par primer extension mais aucune

fonction ne lui a encore été attribuée (Saito *et al.*, 2009). Le terminateur prédit par mfold (Zuker, 2003) est également proche du codon stop du gène *ydaG* étant sur le brin opposé. En comparaison avec *E. faecium*, le transcrit *ern1260* est légèrement plus petit, le profil d'expression est semblable mais reste exprimé plus longtemps chez *E. faecium* [Figure 13] et a un arrangement semblable avec celui de BsrC. Avec les données initiales d'essai de demi-vie, on a considéré qu'*ern1260* était un ARNreg ne pouvant agir qu'en *cis*, au niveau du 3'UTR du gène codé sur le brin opposé, gène étant *fsaA* qui code une aldolase putative chez *E. faecium*. L'hypothèse ayant été retenue est une protection de l'ARNm par interaction avec *ern1260* au niveau du 3'UTR ou avec une hypothétique seconde forme d'*ern1260* ayant une interaction plus grande avec l'ARNm de *fsaA*. Mais le gène *fsaA* étant faiblement exprimé chez *E. faecium*, les conditions d'expression de ce gène devront être explorées. De plus, la délétion de BsrC chez *B. subtilis* n'a eu aucun impact sur la capacité de croissance de la bactérie suggérant un rôle mineur de cet ARN (Saito *et al.*, 2009).

L'exploration du système de régulation peut néanmoins être menée en utilisant des plasmides avec promoteur constitutif. La première étape serait de déterminer où s'arrête la 3'UTR de *fsaA*. On pourrait insérer le gène *fsaA* et la région intergénique suivant ce gène dans un plasmide avec un promoteur constitutif pour pouvoir exprimer le gène. Après vérification de l'expression, la taille déterminée par Northern blot comparée à celle de la séquence codante nous donnera l'information de la taille de la région 3'UTR. La longueur de cette 3'UTR pourra être vérifiée par RT-PCR pour s'assurer de la position par rapport à la séquence d'*ern1260*. On pourra alors vérifier si une interaction directe existe bien entre *ern1260* et le 3'UTR de *fsaA* par retard sur gel. En absence de « long » 3'UTR, la seconde forme d'*ern1260* pourra être recherchée, par Northern blot et RT-PCR, en condition subinhibitrice de daptomycine [Figure 41]. Si cette seconde forme est mise en évidence, l'interaction entre cette forme et l'ARNm de *fsaA* pourra également être testée par retard sur gel. Ensuite, la fonction d'*ern1260* dans la stabilité de l'ARNm de *fsaA* pourrait être

évaluée en utilisant un plasmide avec promoteur inductible contenant *ern1260* et un plasmide avec un promoteur constitutionnel contenant *fsaA* avec la séquence en 3'UTR impliquée dans l'interaction. Une variation de stabilité de l'ARNm de *fsaA* en présence de l'inducteur démontrerait la fonction d'*ern1260*.

Un autre ARNreg, *ern0280*, pour lequel nous avons aujourd'hui peu de données expérimentales, doit être exploré. En effet, c'est le 2<sup>ème</sup> ARN le plus exprimé des 16 candidats sélectionnés, et celui-ci code pour une toxine putative *fst*-like. Étonnamment, son antitoxine putative, *ern0290*, est faiblement exprimée. La prédiction de ce système comme T-AT de type I est très probable, d'autant que ce système est largement retrouvé dans différents plasmides d'*E. faecium*. La première étape pour l'investigation serait de déterminer la demi-vie d'*ern0280* et *ern0290*, en s'attendant à une demi-vie plus importante d'*ern0280*. La seconde étape est de vérifier si une interaction est bien existante entre la toxine et l'antitoxine, puis de vérifier par TOE print, que l'ORF de la toxine est bien capable de fixer un ribosome et que la présence de l'antitoxine altère la position du TOE print d'*ern0280*. Enfin, la toxine pourra être produite *in vitro* pour évaluer sa cytotoxicité et son activité bactéricide.

Les ARN *ern2210* et *ern2410* ont des profils d'expression semblables, c'est-à-dire une expression de deux formes distinctes, plutôt lors de la phase exponentielle de croissance. En se basant sur leurs environnements génétiques, on a considéré ces deux ARN comme de potentiels régulateurs en 5'UTR [Figure 29]. La première étape dans leur investigation sera donc de vérifier par RT-PCR et/ou Northern Blot si ces ARN sont bien co-transcrits avec le gène retrouvé en aval. Ensuite, la régulation en 5' de ces 2 ARN pourra être mise en évidence en utilisant des souches délétées  $\Delta ern2210$  et  $\Delta ern2410$ . Dans l'hypothèse que ce soient des riboswitchs, la mise en évidence de la molécule effectrice risque d'être indispensable pour étudier leurs mécanismes. Pour *ern2210*, on peut supposer que c'est un riboswitch régulant le gène *glmU*. Le mécanisme de régulation de *glmU* n'est pas connu à

ce jour, mais il semble exister une autorégulation provoquée par la production de N-acétylglucosamine. En utilisant le milieu chimiquement défini mis en place lors de ce manuscrit, on pourra essayer de provoquer une variation d'expression de *glmU* en ajoutant du GlcNAc-1P ou de l'UDP-GlcNAc dans le milieu, ou au contraire, faire varier la concentration de glucose dans le milieu. Malheureusement, *ern2210* contient des séquences répétées risquant de rendre son investigation plus complexe [Tableau 7]. Sur le même principe, *ern2410* semble être dans le 5'UTR d'un opéron codant 3 transporteurs et une protéine de régulation probablement impliquée dans le métabolisme de la biotine. On pourra donc faire varier la concentration de biotine du milieu chimiquement défini pour étudier une potentielle variation d'expression de l'opéron. Cette variation d'expression peut également être révélée en utilisant un système rapporteur utilisant comme promoteur les riboswitchs. L'avantage de ces systèmes rapporteurs est la possibilité de produire facilement plusieurs construits avec des délétions et des mutations ciblées pour détecter les séquences indispensables au riboswitch. Au niveau du mécanisme, l'induction de l'élongation par les molécules effectrices pourra être mise en évidence par transcription *in vitro*. Dans l'hypothèse que ces riboswitchs soient de type ribozyme, l'observation d'une potentielle catalyse devra se faire en présence d'ions en plus de la molécule effectrice. Cet essai pourra se faire lors ou après la transcription *in vitro*.

Dans ce manuscrit, nous mettons en évidence 2 T-box riboswitchs, *ern1900* et *ern2050*. La régulation en 5' de *ern2050* sur l'opéron *hisS-aspS* a bien été mise en évidence à l'aide des souches d'étude [Article 2, Figure 1B et Table 1]. En plus des séquences consensus, la détermination expérimentale de la structure secondaire d'*ern2050* est typique d'un T-box riboswitch [Article 2, Figure 2]. Néanmoins, nous n'avons pas réussi à induire une variation d'expression de l'opéron avec les restrictions d'acides aminés dans le milieu [Article 2, Figure S3 et Table S7]. Malgré des premiers résultats d'interaction T-box/ARNt encourageant, des expérimentations supplémentaires sont également nécessaires pour

compléter l'investigation mécanistique du T-box riboswitch. Pour cela, nous allons utiliser plusieurs construits d'ARNt et d'*ern2050* pour tester les zones d'interaction. Par mutation ciblée, nous avons produit plusieurs *ern2050* avec diverses mutations au niveau du codon « specifier », permettant de vérifier la spécificité du T-box riboswitch. Pour les ARNt, nous avons produit des ARNt avec des modifications au niveau de l'anticodon, et pour mimer l'amino-acylation, l'ajout d'un C à l'extrémité -CCA 3'. Les expérimentations d'interaction *in vitro* sont actuellement en cours (Dr F.DR. Wencker, Institute for molecular infection biology). Les résultats obtenus à ce jour sont : l'absence d'impact de l'ajout de C à l'extrémité -CCA 3', une confirmation de l'interaction de l'ARNt<sup>Asp</sup>, une absence d'interaction avec l'ARNt<sup>His</sup> mais une interaction avec l'ARNt<sup>Lys</sup>. Effectivement, le codon correspondant à l'histidine (CAC) est positionné juste avant une séquence appariée (U<sub>106</sub>CU<sub>108</sub>) et le 3<sup>ème</sup> C du codon (C<sub>105</sub>) peut potentiellement s'apparier avec un G sur le brin opposé (G<sub>28</sub>) malgré l'observation expérimentale de la structure secondaire [Article 2, Figure 2]. L'interaction avec l'ARNt<sup>Lys</sup> est plus étonnante mais possible, vu la présence d'un codon AAA au niveau de la « specifier loop ». De plus, l'aspartate est précurseur de la biosynthèse de la lysine [Figure 45] et pourrait donc impacter indirectement l'amino-acylation des ARNt<sup>Lys</sup> et donc la régulation d'*ern2050*.

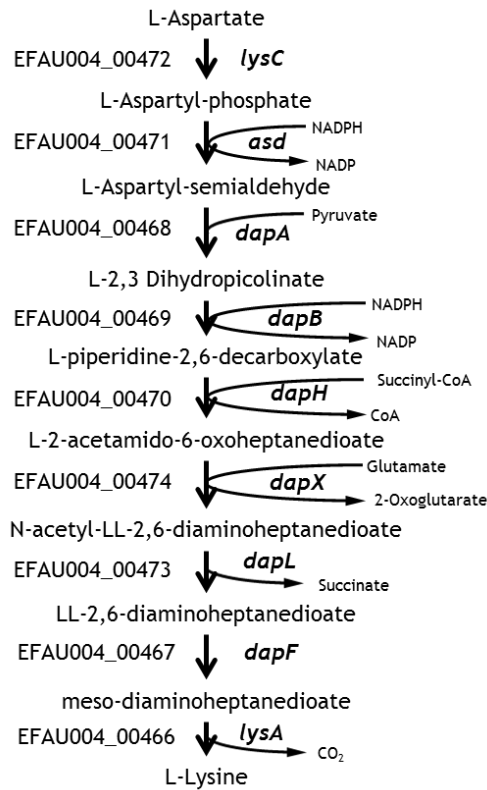


Figure 45. Voie de biosynthèse de la L-Lysine chez *E. faecium* Aus0004.

*lysC* : Aspartate kinase ; *asd* : aspartate semi-aldehyde déhydrogénase ; *dapA* : dihydrodipicolinate synthase ; *dapB* : dihydrodipicolinate reductase ; *dapH* : tétrahydrodipicolinate acétyltransférase ; *dapX* : acétyl-diaminopimelate aminotransférase ; *dapL* : N-acétyl-diaminopimelate déacétylase ; *dapF* : diaminopimelate épimérase ; *lysA* : diaminopimelate décarboxylase.

Dans nos essais de détermination des acides aminés essentiels, nous avons remarqué un impact de la restriction de la lysine sur la souche  $\Delta ern2050$ -Aus0004 [Article 2, Figure 4], qui sera donc à explorer plus précisément, notamment au niveau de la transcription des deux synthétases. Ces résultats sont préliminaires et restent à confirmer. En parallèle, il serait intéressant de valider officiellement *ern1900* comme T-box riboswitch. La première étape serait de regarder la variation d'expression de l'opéron *hisXYZ* selon la concentration d'histidine dans le milieu chimiquement défini. Sur le même modèle qu'*ern2050*, il faudrait déterminer les extrémités 5' et 3' de cet ARN, sachant que cet ARN a la particularité de présenter deux motifs répétés « T-box/terminateur ». A ce jour, ce type de T-box riboswitch dit « en tandem » n'a pas encore été validé expérimentalement,



bien que prédit bio-informatiquement (Vitreschak *et al.*, 2008). Il sera intéressant de déterminer expérimentalement la structure secondaire de ce type d'ARN pour vérifier si la structure est telle qu'attendue (2 tige-boucles terminatrices). Néanmoins, avec un seul codon spécifier mais deux séquences T-box, le mécanisme de régulation par l'ARNt non aminoacylé est plus difficile à appréhender.

Enfin, l'étude des petits ARN a permis de remettre en avant le mécanisme de régulation de *tet(M)*, reposant sur l'hypothèse « originale » d'une accélération de synthèse d'un peptide leader putatif causée par une altération globale de la traduction due à la tétracycline (Su *et al.*, 1992). Jusqu'à ce jour, les expériences sur les mécanismes d'atténuation démontrent plutôt des changements conformationnels dus à une pause ribosomale lors de la synthèse du peptide leader (Yanofsky, 1981; Lovett, 1990; Ramu *et al.*, 2009; Chancey *et al.*, 2015). Dans la région régulatrice *Ptet*, le peptide leader putatif initialement décrit ne semble pas avoir de rôle dans la régulation de *tet(M)*, démontré par nos essais de TOE print [Article 1, Figure 6b]. Si nous réfutons l'hypothèse initialement décrite du mécanisme de régulation par atténuation transcriptionnelle et donc l'absence de recrutement ribosomal par l'*orf*<sub>12</sub>, alors la conformation la plus stable est la conformation antiterminatrice [Figure 18].

Une première hypothèse de mécanisme de régulation serait donc un mécanisme de type riboswitch médié par une protéine régulatrice [Figure 46]. Cette hypothèse est cohérente avec l'induction par la tétracycline de la conformation antiterminatrice dépendante de la présence du transposon Tn916 (Scornec *et al.*, 2017). On peut donc supposer qu'un acteur de ce transposon (protéine ou ARNreg) soit impliquée dans la régulation au niveau de *Ptet*. La fonction de certaines protéines du module de régulation est encore inconnue aujourd'hui (Celli et Trieu-Cuot, 1998). On aurait donc une transcription basale jusqu'au niveau de la séquence codante de ce module de régulation [Figure 10]. La protéine pourra

donc interagir avec la région *P<sub>tet</sub>* pour favoriser la conformation terminatrice, soit *ern0030*.

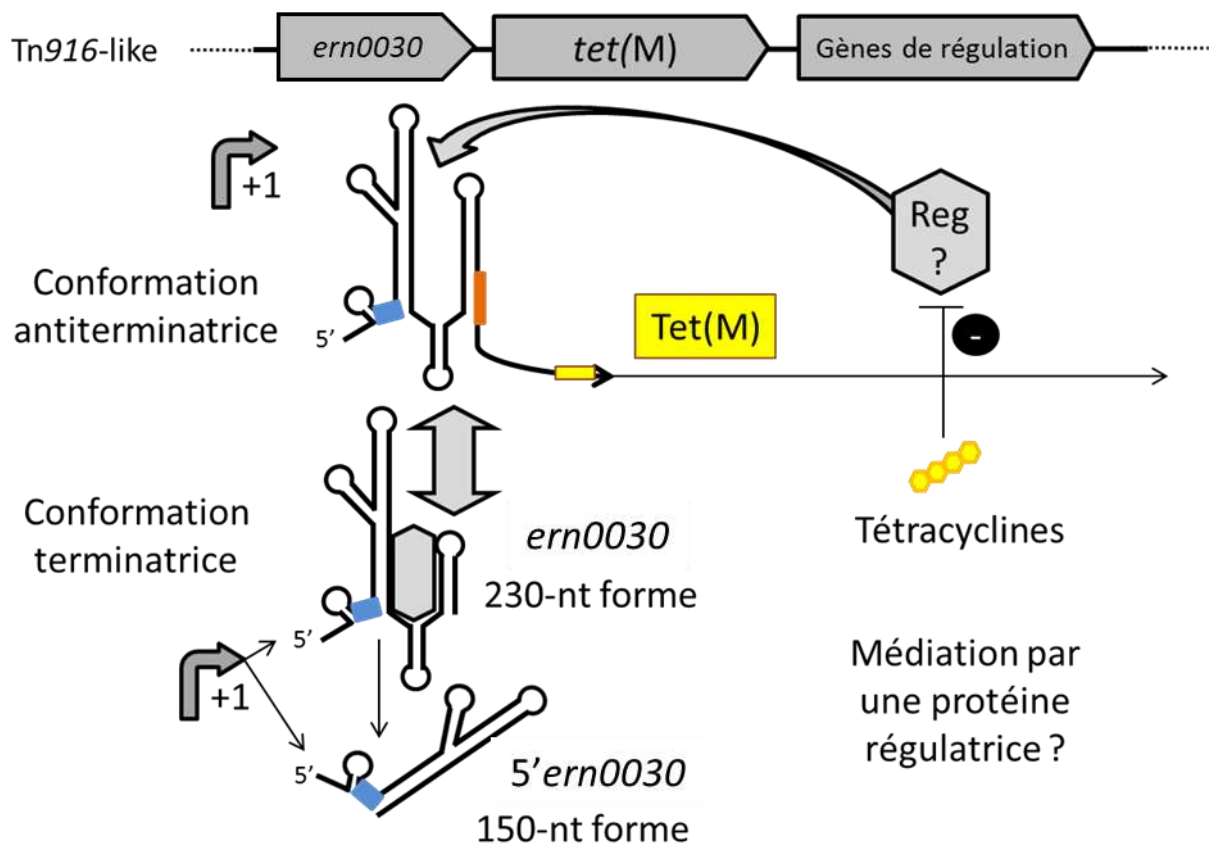


Figure 46. Hypothèse du mécanisme de régulation de *tet(M)* Tn916-dépendant : médiation par une protéine régulatrice.  
La protéine régulatrice hypothétique est représentée par l'héxagone gris. Le RBS correspondant au petit peptide positionné en 5' d'*ern0030* est représenté par le rectangle bleu. Le RBS correspondant au petit peptide révélé par TOE print [Article 1, Figure 6A] est représenté par le rectangle orange.

Cette hypothèse serait donc en accord avec l'expression basale de *tet(M)* en absence de tétracycline [Article 1, Figure 4A et 5]. Egalement, la présence d'une telle protéine expliquerait l'absence de différence majeure de sensibilité aux tétracyclines dans notre système plasmidique [Article 1, Table 1]. En effet, l'expression du système *ern0030-tet(M)* plasmidique est bien plus importante que l'expression chromosomique de la protéine de régulation potentielle. La conformation antiterminatrice est donc favorisée malgré les

mutations ou délétions occasionnées au niveau de la séquence *ern0030*. Le seul impact potentiel serait une fragilisation de la structure, cohérent avec nos légères différences de sensibilité aux tétracyclines observées [Article 1, Table 1]. La mise en évidence de cette protéine régulatrice et de sa fonction nécessiterait donc la synthèse *in vitro* des protéines codées sur le module de régulation. L'interaction entre la protéine régulatrice et la région *Ptet* devra être révélée par retard sur gel. Si une interaction existe, le mécanisme plus précis pourrait être mis en évidence en effectuant une transcription *in vitro* du +1 de transcription d'*ern0030* jusqu'au début de la séquence codante de *tet(M)* (matrice 5'*tetM*) en présence ou non de la protéine régulatrice. Nos données préliminaires de transcription *in vitro* ayant montré la présence d'un seul transcrit (400nt) à partir de la matrice 5'*tetM* [Figure 32], l'ajout de la protéine régulatrice devrait favoriser au moins l'apparition du transcrit *ern0030* à 230nt. Enfin, on pourrait vérifier par un plasmide, si la surexpression de cette protéine au sein de la bactérie diminue l'expression de *tet(M)*. Les séquences critiques dans l'interaction avec cette protéine régulatrice pourront donc être recherchées.

L'observation d'une induction par les macrolides et la spectinomycine de la conformation antiterminatrice indépendamment de la présence du transposon Tn916 suggère l'existence d'un mécanisme de régulation complémentaire (Scornec *et al.*, 2017). Étonnamment, l'effet le plus prononcé est observé avec des antibiotiques pouvant provoquer un arrêt ribosomal lors de l'élongation (lincosamides, streptogramines et spectinomycine) [Figure 7.C et D]. On peut donc supposer un mécanisme de régulation impliquant un ribosome, comme un mécanisme d'atténuation ou un mécanisme de protection ribosomal. La présence de deux autres peptides putatifs que l'*orf*<sub>12</sub> renforce cette hypothèse, d'autant que celui positionné juste après l'extrémité 3' d'*ern0030* est capable de fixer un ribosome [Article 1, Figure 6A]. Au vu de la position, un mécanisme par protection ribosomale serait plus probable [Figure 47]. Cela expliquerait les différences d'expression et de sensibilité

aux tétracyclines entre notre souche  $\Delta ern0030$ -Aus0004<sup>TR</sup> et les souches mutantes BTR (Beabout *et al.*, 2015b). En effet, les deux types de souches abolissent l'induction par la tétracycline car le site de fixation putatif de la protéine régulatrice est supprimé. Mais, la délétion totale ne laisse qu'une petite séquence 5'UTR au gène *tet(M)* ne permettant pas la traduction du petit peptide putatif. On peut donc imaginer que cette petite séquence est plus facilement et rapidement clivable par d'hypothétiques RNases, contrairement aux souches BTR gardant un 5'UTR de *tet(M)* plus structuré et une séquence pouvant recruter un ribosome au niveau de ce petit peptide. Un autre argument en faveur de ce mécanisme est la multiplicité de transcrit d'*ern0030* [Article 1, Figure 2]. Ce mécanisme expliquerait la production du transcrit à 100nt exprimé en phase stationnaire, phase où la traduction est ralentie. En effet, un peptide putatif est localisé en 5' d'*ern0030* et pourrait donc protéger la séquence 5' d'*ern0030* d'une dégradation. L'investigation de ce mécanisme nécessitera la mise en évidence des RNases impliquées dans le clivage potentiel des différentes séquences simple brin d'*ern0030*. La capacité de fixation du peptide situé en 5' devra être évaluée par TOE print. Enfin, dans la compréhension des deux mécanismes de régulation supposés, la détermination de la structure secondaire par in-line probing d'*ern0030* et de la conformation antiterminatrice de *Ptet* semble importante.

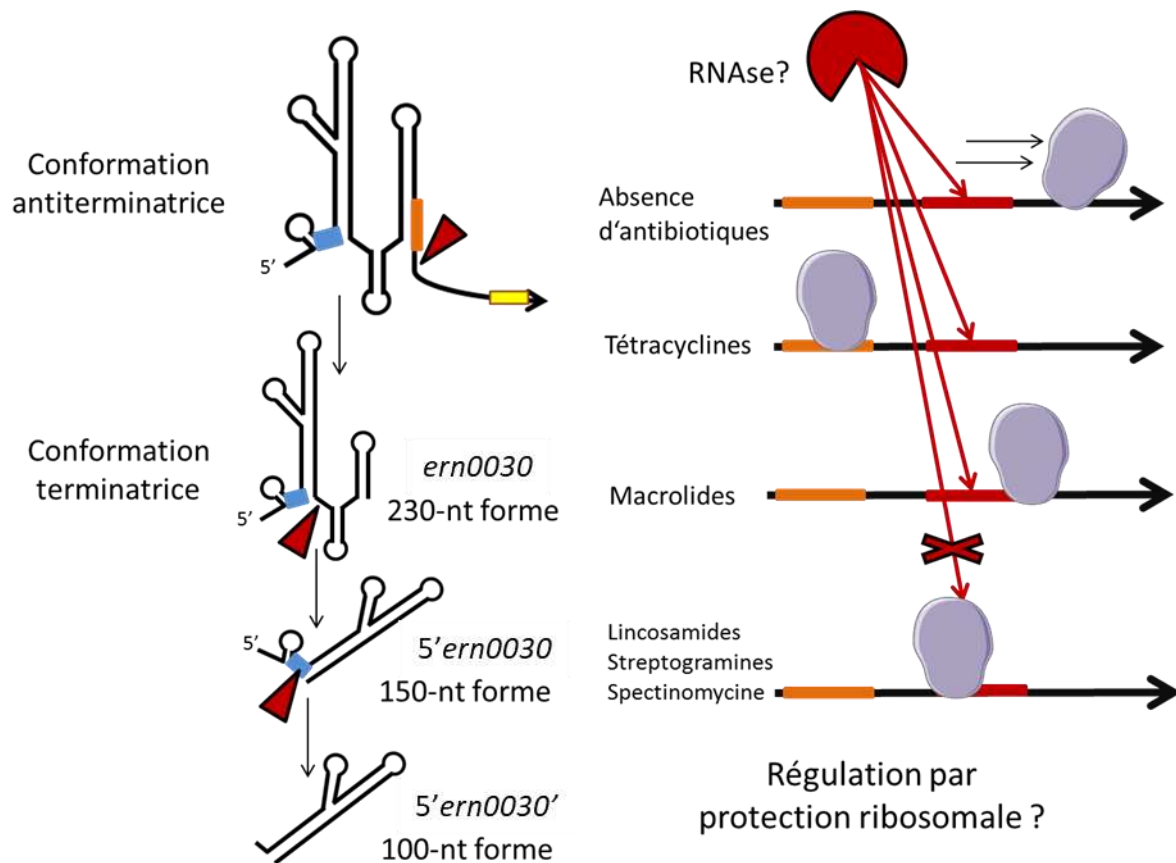


Figure 47. Hypothèse du mécanisme de régulation de *tet(M)* Tn916-indépendant : médiation par protection ribosomale.  
Les sites de clivage putatifs (par RNAse ou non) sont indiqués par les triangles rouges.

**Tableau 10. Liste récapitulative des fonctions des 16 ARN sélectionnés initialement.**

	Type	Cible	Mécanisme	Fonction	Arguments
<i>ern0030</i>	5'UTR***	<i>tet(M)</i> ***	- Atténuation transcriptionnelle* - Riboswitch** - Protection ribosomale*	<b>Antibiorésistance***</b> <b>Mobilité génétique***</b>	- 5'3'RACE, variation expression et sensibilité ATB si délétion - Mécanisme remis en cause (peptide leader non mis en évidence) - Littérature : (Su <i>et al.</i> , 1992; Celli et Trieu-Cuot, 1998; Beabout <i>et al.</i> , 2015b; Scornec <i>et al.</i> , 2017; Michaux <i>et al.</i> , 2020; Le Neindre <i>et al.</i> , 2021)
<i>ern1300</i>	5'UTR***	<i>mef(A)</i> ***	- Atténuation transcriptionnelle*** - Protection ribosomale*	<b>Antibiorésistance***</b>	- 5'3'RACE, homologie avec 5'UTR de <i>mef(E)</i> - Littérature : (Chancey <i>et al.</i> , 2011, 2015; Michaux <i>et al.</i> , 2020)
<i>ern1030</i>	5'UTR*	<i>ugpB</i> *	Riboswitch*	Métabolique* (T)	- RNAseq
<i>ern1900</i>	5'UTR***	<i>hisXYZ</i> ***	T-box riboswitch**	Métabolique** (T)	- Référéncé dans la base de données TBDB ( <a href="#">LVNDX2SZ</a> ) - Variation d'expression de l'opéron <i>hisXYZ</i> chez $\Delta ern2050$ (régulation indirecte) - Littérature : (Vitreschak <i>et al.</i> , 2008; Marchand <i>et al.</i> , 2020)
<i>ern2410</i>	5'UTR*	ABC transporteurs*	Riboswitch*	Métabolique* (T)	- 5'3'RACE - Littérature : (Michaux <i>et al.</i> , 2020)
<i>ern2050</i>	5'UTR***	<i>hisS-aspS</i> ***	T-box riboswitch***	<b>Métabolique*** (Trad)</b>	- Référéncé dans la base de données TBDB ( <a href="#">F1TD5WKQ</a> ) - Structure secondaire et interaction validées expérimentalement - Variation d'expression de l'opéron <i>hisS-aspS</i> chez $\Delta ern2050$ (Régulation directe) - Littérature : (Vitreschak <i>et al.</i> , 2008; Marchand <i>et al.</i> , 2020)
<i>ern2210</i>	5'UTR*	<i>glmU</i> *	Riboswitch*	Métabolique* (BioΣ)	- 5'3'RACE
<i>ern0280</i>	ARNreg**	<i>ern0290</i> **	<i>cis</i> ** (T-AT type 1)	Maintien d'EGM**	- 5'3'RACE, prédiction toxine Fst-like - Conservation au sein de plasmide
<i>ern0510</i>	ARNreg*	Transposase*	<i>cis</i> *	Mobilité génétique*	- RNAseq
<i>ern0980</i>	ARNreg*	<i>xkdP</i> *	<i>cis</i> *	?	- RNAseq, Northern Blot
<i>ern1060</i>	ARNreg*	<i>gloA</i> *	<i>cis</i> *	Métabolique* (Dtx)	- RNAseq
<i>ern1260</i>	ARNreg*	<i>fsaA</i> *	<i>cis</i> *	Métabolique* (BioΣ)	- 5'3'RACE, stabilité faible - Littérature : (Saito <i>et al.</i> , 2009)
<i>ern0160</i>	ARNreg**	<i>ern0170</i> * Transporteurs*	<i>cis-trans</i> *	Métabolique* (T)	- 5'3'RACE, stabilité élevée
<i>ern0120</i>	ARNreg**	<i>rnhC</i> *	<i>trans</i> *	Métabolique*	- 5'3'RACE, stabilité élevée
<i>ern1940</i>	ARNreg*	?	?	?	- RNAseq
<i>ern2020</i>	ARNreg*	?	?	?	- RNAseq, Northern Blot

\* : Faible confiance, Aucun argument convaincant

\*\* : Confiance modérée, Localisation génétique et/ou prédiction bio-informatique et/ou données transcriptomiques

\*\*\* : Forte confiance, Arguments expérimentaux phénotypiques et/ou transcriptomiques et/ou mécanistiques et/ou déjà décrit dans la littérature

Fonction métabolique : T (transport), Dtx (Détoxification), BioΣ (Biosynthèse), Trad (Traduction)

## Conclusion

Ce travail de thèse est la première investigation fonctionnelle des petits ARN décrit chez *E. faecium*. A l'aide des premières caractérisations génétiques, les premières hypothèses concernant les ARN les plus intéressants ont pu être posées, nécessitant des validations expérimentales. Un des ARN ayant eu une investigation fonctionnelle approfondie est un T-box riboswitch régulant un opéron de deux aminoacyl-ARNt synthétases. La validation expérimentale de ce nouveau T-box riboswitch a permis la mise en évidence d'une autre T-box riboswitch et la mise en place d'un milieu de culture chimiquement défini pour *E. faecium*. De plus, ce système de régulation s'avère important pour l'adaptation lors de la croissance bactérienne, notamment en milieu pauvre en aspartate. Nos travaux ont également permis de revisiter un mécanisme de régulation du gène de résistance aux tétracyclines *tet(M)*. Bien que le mécanisme de régulation soit post-transcriptionnel, l'hypothèse initiale d'atténuation transcriptionnelle n'est pas supportée par nos résultats, suggérant un mécanisme de régulation alternatif et nécessitant de plus amples investigations.



## Bibliographie

- Aarestrup, F. M., Agerso, Y., Gerner-Smidt, P., Madsen, M., et Jensen, L. B. (2000). Comparison of antimicrobial resistance phenotypes and resistance genes in *Enterococcus faecalis* and *Enterococcus faecium* from humans in the community, broilers, and pigs in Denmark. *Diagn Microbiol Infect Dis* 37, 127-137. doi:10.1016/s0732-8893(00)00130-9.
- Agwuh, K. N., et MacGowan, A. (2006). Pharmacokinetics and pharmacodynamics of the tetracyclines including glycylcyclines. *J Antimicrob Chemother* 58, 256-265. doi:10.1093/jac/dkl224.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., et Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* 215, 403-410. doi:10.1016/S0022-2836(05)80360-2.
- Anderson, M. S., et Raetz, C. R. (1987). Biosynthesis of lipid A precursors in *Escherichia coli*. A cytoplasmic acyltransferase that converts UDP-N-acetylglucosamine to UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine. *J Biol Chem* 262, 5159-5169.
- Anokhina, M. M., Barta, A., Nierhaus, K. H., Spiridonova, V. A., et Kopylov, A. M. (2004). Mapping of the second tetracycline binding site on the ribosomal small subunit of *E. coli*. *Nucleic Acids Res* 32, 2594-2597. doi:10.1093/nar/gkh583.
- Antonelli, A., D'Andrea, M. M., Brenciani, A., Galeotti, C. L., Morroni, G., Pollini, S., et al. (2018). Characterization of *poxTA*, a novel phenicol-oxazolidinone-tetracycline resistance gene from an MRSA of clinical origin. *J Antimicrob Chemother* 73, 1763-1769. doi:10.1093/jac/dky088.
- Arias, C. A., et Murray, B. E. (2012). The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat Rev Microbiol* 10, 266-278. doi:10.1038/nrmicro2761.
- Arsène, S., et Leclercq, R. (2007). Role of a *qnr*-Like gene in the intrinsic resistance of *Enterococcus faecalis* to fluoroquinolones. *Antimicrob Agents Chemother* 51, 3254-3258. doi:10.1128/AAC.00274-07.
- Arthur, M., Depardieu, F., Molinas, C., Reynolds, P., et Courvalin, P. (1995). The *vanZ* gene of Tn1546 from *Enterococcus faecium* BM4147 confers resistance to teicoplanin. *Gene* 154, 87-92. doi:10.1016/0378-1119(94)00851-i.
- Arthur, M., Molinas, C., Depardieu, F., et Courvalin, P. (1993). Characterization of Tn1546, a Tn3-related transposon conferring glycopeptide resistance by synthesis of depsipeptide peptidoglycan precursors in *Enterococcus faecium* BM4147. *J Bacteriol* 175, 117-127.
- Ayobami, O., Willrich, N., Reuss, A., Eckmanns, T., et Markwart, R. (2020). The ongoing challenge of vancomycin-resistant *Enterococcus faecium* and *Enterococcus faecalis* in Europe: an epidemiological analysis of bloodstream infections. *Emerg Microbes Infect* 9, 1180-1193. doi:10.1080/22221751.2020.1769500.
- Babitzke, P., et Romeo, T. (2007). CsrB sRNA family: sequestration of RNA-binding regulatory proteins. *Current Opinion in Microbiology* 10, 156-163. doi:10.1016/j.mib.2007.03.007.
- Bauer, G., Berens, C., Projan, S. J., et Hillen, W. (2004). Comparison of tetracycline and tigecycline binding to ribosomes mapped by dimethylsulphate and drug-directed Fe<sup>2+</sup> cleavage of 16S rRNA. *J Antimicrob Chemother* 53, 592-599. doi:10.1093/jac/dkh125.
- Beabout, K., Hammerstrom, T. G., Perez, A. M., Magalhães, B. F., Prater, A. G., Clements, T. P., et al. (2015a). The ribosomal S10 protein is a general target for decreased tigecycline susceptibility. *Antimicrob Agents Chemother* 59, 5561-5566. doi:10.1128/AAC.00547-15.
- Beabout, K., Hammerstrom, T. G., Wang, T. T., Bhatti, M., Christie, P. J., Saxer, G., et al. (2015b). Rampant parasexuality evolves in a hospital pathogen during antibiotic selection. *Molecular Biology and Evolution* 32, 2585. doi:10.1093/molbev/msv133.
- Becker, B., et Cooper, M. A. (2013). Aminoglycoside antibiotics in the 21st century. *ACS Chem. Biol.* 8, 105-115. doi:10.1021/cb3005116.

- Bender, J. K., Cattoir, V., Hegstad, K., Sadowy, E., Coque, T. M., Westh, H., *et al.* (2018). Update on prevalence and mechanisms of resistance to linezolid, tigecycline and daptomycin in enterococci in Europe: Towards a common nomenclature. *Drug Resist. Updat.* 40, 25-39. doi:10.1016/j.drup.2018.10.002.
- Bhattacharya, M., Parakh, A., et Narang, M. (2009). Tigecycline. *Journal of Postgraduate Medicine* 55, 65. doi:10.4103/0022-3859.48443.
- Boehm, A. B., et Sassoubre, L. M. (2014). "Enterococci as indicators of environmental fecal contamination," in *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*, eds. M. S. Gilmore, D. B. Clewell, Y. Ike, and N. Shankar (Boston: Massachusetts Eye and Ear Infirmary).
- Bonacina, J., Suárez, N., Hormigo, R., Fadda, S., Lechner, M., et Saavedra, L. (2017). A genomic view of food-related and probiotic *Enterococcus* strains. *DNA Research* 24, 11-24. doi:10.1093/dnares/dsw043.
- Borovinskaya, M. A., Shoji, S., Holton, J. M., Fredrick, K., et Cate, J. H. D. (2007). A steric block in translation caused by the antibiotic spectinomycin. *ACS Chem Biol* 2, 545-552. doi:10.1021/cb700100n.
- Boto, L. (2010). Horizontal gene transfer in evolution: facts and challenges. *Proceedings of the Royal Society B: Biological Sciences* 277, 819. doi:10.1098/rspb.2009.1679.
- Boukthir, S., Dejoies, L., Zouari, A., Collet, A., Potrel, S., Auger, G., *et al.* (2020). In vitro activity of eravacycline and mechanisms of resistance in enterococci. *Int J Antimicrob Agents* 56, 106215. doi:10.1016/j.ijantimicag.2020.106215.
- Brantl, S. (2007). Regulatory mechanisms employed by *cis*-encoded antisense RNAs. *Current Opinion in Microbiology* 10, 102-109. doi:10.1016/j.mib.2007.03.012.
- Brodersen, D. E., Clemons, W. M., Carter, A. P., Morgan-Warren, R. J., Wimberly, B. T., et Ramakrishnan, V. (2000). The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. *Cell* 103, 1143-1154. doi:10.1016/S0092-8674(00)00216-6.
- Burdett, V. (1990). Nucleotide sequence of the *tet*(M) gene of Tn916. *Nucleic Acids Res* 18, 6137.
- Burdett, V. (1991). Purification and characterization of Tet(M), a protein that renders ribosomes resistant to tetracycline. *J. Biol. Chem.* 266, 2872-2877.
- Burdett, V. (1996). Tet(M)-promoted release of tetracycline from ribosomes is GTP dependent. *J Bacteriol* 178, 3246-3251.
- Byappanahalli, M. N., Nevers, M. B., Korajkic, A., Staley, Z. R., et Harwood, V. J. (2012). Enterococci in the environment. *Microbiol Mol Biol Rev* 76, 685-706. doi:10.1128/MMBR.00023-12.
- Caldelari, I., Chao, Y., Romby, P., et Vogel, J. (2013). RNA-mediated regulation in pathogenic bacteria. *Cold Spring Harb Perspect Med* 3. doi:10.1101/cshperspect.a010298.
- Canani, R. B., Cirillo, P., Terrin, G., Cesarano, L., Spagnuolo, M. I., Vincenzo, A. D., *et al.* (2007). Probiotics for treatment of acute diarrhoea in children: randomised clinical trial of five different preparations. *BMJ* 335, 340. doi:10.1136/bmj.39272.581736.55.
- Carbona, S. L., Sauvageot, N., Giard, J.-C., Benachour, A., Posteraro, B., Auffray, Y., *et al.* (2007). Comparative study of the physiological roles of three peroxidases (NADH peroxidase, Alkyl hydroperoxide reductase and Thiol peroxidase) in oxidative stress response, survival inside macrophages and virulence of *Enterococcus faecalis*. *Molecular Microbiology* 66, 1148-1163. doi:https://doi.org/10.1111/j.1365-2958.2007.05987.x.
- Carrier, C., et Courvalin, P. (1990). Emergence of 4',4"-aminoglycoside nucleotidyltransferase in enterococci. *Antimicrob Agents Chemother* 34, 1565-1569. doi:10.1128/aac.34.8.1565.
- Carver, T., Harris, S. R., Berriman, M., Parkhill, J., et McQuillan, J. A. (2012). Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* 28, 464-469. doi:10.1093/bioinformatics/btr703.

- Cattoir, V. (2004). Pompes d'efflux et résistance aux antibiotiques chez les bactéries. *Pathologie Biologie* 52, 607-616. doi:10.1016/j.patbio.2004.09.001.
- Cattoir, V., et Giard, J.-C. (2014). Antibiotic resistance in *Enterococcus faecium* clinical isolates. *Expert Review of Anti-infective Therapy* 12, 239-248. doi:10.1586/14787210.2014.870886.
- Cattoir, V., Isnard, C., Cosquer, T., Odhiambo, A., Bucquet, F., Guérin, F., et al. (2015). Genomic analysis of reduced susceptibility to tigecycline in *Enterococcus faecium*. *Antimicrob Agents Chemother* 59, 239-244. doi:10.1128/AAC.04174-14.
- Cattoir, V., et Leclercq, R. (2010). [Enterococci resistant to glycopeptides]. *Med Sci (Paris)* 26, 936-942. doi:10.1051/medsci/20102611936.
- Cattoir, V., et Leclercq, R. (2013). Twenty-five years of shared life with vancomycin-resistant enterococci: is it time to divorce? *J Antimicrob Chemother* 68, 731-742. doi:10.1093/jac/dks469.
- Celiberto, L. S., Bedani, R., DeJani, N. N., Ivo de Medeiros, A., Sampaio Zuanon, J. A., Spolidorio, L. C., et al. (2017). Effect of a probiotic beverage consumption (*Enterococcus faecium* CRL 183 and *Bifidobacterium longum* ATCC 15707) in rats with chemically induced colitis. *PLoS One* 12, e0175935. doi:10.1371/journal.pone.0175935.
- Celli, J., et Trieu-Cuot, P. (1998). Circularization of Tn916 is required for expression of the transposon-encoded transfer functions: characterization of long tetracycline-inducible transcripts reading through the attachment site. *Molecular Microbiology* 28, 103-117. doi:10.1046/j.1365-2958.1998.00778.x.
- Chancey, S. T., Bai, X., Kumar, N., Drabek, E. F., Daugherty, S. C., Colon, T., et al. (2015). Transcriptional attenuation controls macrolide inducible efflux and resistance in *Streptococcus pneumoniae* and in other Gram-positive bacteria containing *mef/mel(msr(D))* elements. *PLoS ONE* 10. doi:10.1371/journal.pone.0116254.
- Chancey, S. T., Zhou, X., Zähler, D., et Stephens, D. S. (2011). Induction of efflux-mediated macrolide resistance in *Streptococcus pneumoniae*. *Antimicrob Agents Chemother* 55, 3413-3422. doi:10.1128/AAC.00060-11.
- Chander, P., Halbig, K. M., Miller, J. K., Fields, C. J., Bonner, H. K. S., Grabner, G. K., et al. (2005). Structure of the nucleotide complex of PyrR, the *pyr* attenuation protein from *Bacillus caldolyticus*, suggests dual regulation by pyrimidine and purine nucleotides. *J Bacteriol* 187, 1773-1782. doi:10.1128/JB.187.5.1773-1782.2005.
- Chopra, I., et Roberts, M. (2001). Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiol Mol Biol Rev* 65, 232-260. doi:10.1128/MMBR.65.2.232-260.2001.
- Chukwudi, C. U., et Good, L. (2016). Interaction of the tetracyclines with double-stranded RNAs of random base sequence: new perspectives on the target and mechanism of action. *J. Antibiot.* 69, 622-630. doi:10.1038/ja.2015.145.
- Ciric, L., Jasni, A., Vries, L. E. de, Agersø, Y., Mullany, P., et Roberts, A. P. (2013). The Tn916/Tn1545 Family of conjugative transposons. *Landes Bioscience*.
- Clewell, D. B., Weaver, K. E., Dunny, G. M., Coque, T. M., Francia, M. V., et Hayes, F. (2014). "Extrachromosomal and mobile elements in Enterococci: transmission, maintenance, and epidemiology," in *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*, eds. M. S. Gilmore, D. B. Clewell, Y. Ike, and N. Shankar (Boston: Massachusetts Eye and Ear Infirmary).
- Cocito, C., Di Giambattista, M., Nyssen, E., et Vannuffel, P. (1997). Inhibition of protein synthesis by streptogramins and related antibiotics. *J Antimicrob Chemother* 39 Suppl A, 7-13. doi:10.1093/jac/39.suppl\_1.7.
- Collins, J. A., Irnov, I., Baker, S., et Winkler, W. C. (2007). Mechanism of mRNA destabilization by the *glmS* ribozyme. *Genes Dev.* 21, 3356-3368. doi:10.1101/gad.1605307.
- Colomb-Cotin, M., Lacoste, J., Brun-Buisson, C., Jarlier, V., Coignard, B., et Vaux, S. (2016). Estimating the morbidity and mortality associated with infections due to multidrug-resistant

- bacteria (MDRB), France, 2012. *Antimicrob Resist Infect Control* 5, 56. doi:10.1186/s13756-016-0154-z.
- Corvaisier, S., Bordeau, V., et Felden, B. (2003). Inhibition of transfer messenger RNA aminoacylation and *trans*-translation by aminoglycoside antibiotics. *J. Biol. Chem.* 278, 14788-14797. doi:10.1074/jbc.M212830200.
- Costa, Y., Galimand, M., Leclercq, R., Duval, J., et Courvalin, P. (1993). Characterization of the chromosomal *aac*(6')-II gene specific for *Enterococcus faecium*. *Antimicrob. Agents Chemother.* 37, 1896-1903.
- Courvalin, P. (2006). Vancomycin resistance in Gram-positive cocci. *Clin Infect Dis.* 42, S25-S34. doi:10.1086/491711.
- Daghrir, R., et Drogui, P. (2013). Tetracycline antibiotics in the environment: a review. *Environ Chem Lett* 11, 209-227. doi:10.1007/s10311-013-0404-8.
- Daillère, R., Vétizou, M., Waldschmitt, N., Yamazaki, T., Isnard, C., Poirier-Colame, V., et al. (2016). *Enterococcus hirae* and *Barnesiella intestinihominis* facilitate cyclophosphamide-induced therapeutic immunomodulatory effects. *Immunity* 45, 931-943. doi:10.1016/j.immuni.2016.09.009.
- Dantley, K. A., Dannelly, H. K., et Burdett, V. (1998). Binding Interaction between Tet(M) and the Ribosome: Requirements for Binding. *J Bacteriol* 180, 4089-4092.
- Debono, M., Abbott, B. J., Molloy, R. M., Fukuda, D. S., Hunt, A. H., Daupert, V. M., et al. (1988). Enzymatic and chemical modifications of lipopeptide antibiotic A21978C: the synthesis and evaluation of daptomycin (LY146032). *J Antibiot (Tokyo)* 41, 1093-1105. doi:10.7164/antibiotics.41.1093.
- Dejoies, L., Le Neindre, K., Reissier, S., Felden, B., et Cattoir, V. (2021). Distinct expression profiles of regulatory RNAs in the response to biocides in *Staphylococcus aureus* and *Enterococcus faecium*. *Sci Rep* 11, 6892. doi:10.1038/s41598-021-86376-y.
- Delorme, C., Ehrlich, S. D., et Renault, P. (1999). Regulation of expression of the *Lactococcus lactis* histidine operon. *J. Bacteriol.* 181, 2026-2037.
- Depardieu, F., Podglajen, I., Leclercq, R., Collatz, E., et Courvalin, P. (2007). Modes and modulations of antibiotic resistance gene expression. *Clinical Microbiology Reviews* 20, 79. doi:10.1128/CMR.00015-06.
- Derbise, A., Dyke, K. G., et el Solh, N. (1996). Characterization of a *Staphylococcus aureus* transposon, Tn5405, located within Tn5404 and carrying the aminoglycoside resistance genes, *aphA*-3 and *aadE*. *Plasmid* 35, 174-188. doi:10.1006/plas.1996.0020.
- Dersch, P., Khan, M. A., Mühlen, S., et Görke, B. (2017). Roles of regulatory RNAs for antibiotic resistance in bacteria and their potential value as novel drug targets. *Frontiers in Microbiology* 8. doi:10.3389/fmicb.2017.00803.
- Diaz, L., Kiratisin, P., Mendes, R. E., Panesso, D., Singh, K. V., et Arias, C. A. (2012). Transferable plasmid-mediated resistance to linezolid due to *cfr* in a human clinical isolate of *Enterococcus faecalis*. *Antimicrob Agents Chemother* 56, 3917-3922. doi:10.1128/AAC.00419-12.
- Diaz, L., Tran, T. T., Munita, J. M., Miller, W. R., Rincon, S., Carvajal, L. P., et al. (2014). Whole-genome analyses of *Enterococcus faecium* isolates with diverse daptomycin MICs. *Antimicrob Agents Chemother* 58, 4527-4534. doi:10.1128/AAC.02686-14.
- Doherty, E. A., et Doudna, J. A. (2000). Ribozyme structures and mechanisms. *Annu. Rev. Biochem.* 69, 597-615. doi:10.1146/annurev.biochem.69.1.597.
- Dönhöfer, A., Franckenberg, S., Wickles, S., Berninghausen, O., Beckmann, R., et Wilson, D. N. (2012). Structural basis for TetM-mediated tetracycline resistance. *Proc Natl Acad Sci U S A* 109, 16900-16905. doi:10.1073/pnas.1208037109.
- Dubnau, D. (1984). Translational attenuation: the regulation of bacterial resistance to the macrolide-lincosamide-streptogramin B antibiotics. *CRC Crit Rev Biochem* 16, 103-132. doi:10.3109/10409238409102300.

- Eliopoulos, G. M., et Roberts, M. C. (2003). Tetracycline therapy: Update. *Clinical Infectious Diseases* 36, 462-467. doi:10.1086/367622.
- Felden, B., et Cattoir, V. (2018). Bacterial adaptation to antibiotics through regulatory RNAs. *Antimicrob. Agents Chemother.* doi:10.1128/AAC.02503-17.
- Ferretti, J. J., Gilmore, K. S., et Courvalin, P. (1986). Nucleotide sequence analysis of the gene specifying the bifunctional 6'-aminoglycoside acetyltransferase 2"-aminoglycoside phosphotransferase enzyme in *Streptococcus faecalis* and identification and cloning of gene regions specifying the two activities. *J Bacteriol* 167, 631-638.
- Fiedler, S., Bender, J. K., Klare, I., Halbedel, S., Grohmann, E., Szewzyk, U., et al. (2016). Tigecycline resistance in clinical isolates of *Enterococcus faecium* is mediated by an upregulation of plasmid-encoded tetracycline determinants *tet(L)* and *tet(M)*. *J Antimicrob Chemother* 71, 871-881. doi:10.1093/jac/dkv420.
- Flannagan, S. E., Zitzow, L. A., Su, Y. A., et Clewell, D. B. (1994). Nucleotide sequence of the 18-kb conjugative transposon Tn916 from *Enterococcus faecalis*. *Plasmid* 32, 350-354. doi:10.1006/plas.1994.1077.
- Fontana, R., Aldegheri, M., Ligozzi, M., Lopez, H., Sucari, A., et Satta, G. (1994). Overproduction of a low-affinity penicillin-binding protein and high-level ampicillin resistance in *Enterococcus faecium*. *Antimicrob Agents Chemother* 38, 1980.
- Foulquié Moreno, M. R., Sarantinopoulos, P., Tsakalidou, E., et De Vuyst, L. (2006). The role and application of enterococci in food and health. *International Journal of Food Microbiology* 106, 1-24. doi:10.1016/j.ijfoodmicro.2005.06.026.
- Fouquier d'Hérouël, A., Wessner, F., Halpern, D., Ly-Vu, J., Kennedy, S. P., Serror, P., et al. (2011). A simple and efficient method to search for selected primary transcripts: non-coding and antisense RNAs in the human pathogen *Enterococcus faecalis*. *Nucleic Acids Research* 39, e46. doi:10.1093/nar/gkr012.
- Franke, A. E., et Clewell, D. B. (1981). Evidence for a chromosome-borne resistance transposon (Tn916) in *Streptococcus faecalis* that is capable of "conjugal" transfer in the absence of a conjugative plasmid. *J. Bacteriol.* 145, 494-502.
- Gaca, A. O., et Lemos, J. A. (2019). Adaptation to adversity: the intermingling of stress tolerance and pathogenesis in enterococci. *Microbiol Mol Biol Rev* 83. doi:10.1128/MMBR.00008-19.
- Gallagher, J. C. (2019). Omadacycline: a modernized tetracycline. *Clin Infect Dis* 69, S1-S5. doi:10.1093/cid/ciz394.
- García-Solache, M., et Rice, L. B. (2019). The *Enterococcus*: a model of adaptability to its environment. *Clin Microbiol Rev* 32. doi:10.1128/CMR.00058-18.
- Garnier, F., Taourit, S., Glaser, P., Courvalin, P., et Galimand, M. (2000). Characterization of transposon Tn1549, conferring *VanB*-type resistance in *Enterococcus* spp. *Microbiology (Reading)* oyamada, 1481-1489. doi:10.1099/00221287-146-6-1481.
- Gaugué, I., Oberto, J., et Plumbridge, J. (2014). Regulation of amino sugar utilization in *Bacillus subtilis* by the GntR family regulators, NagR and GamR. *Molecular Microbiology* 92, 100-115. doi:https://doi.org/10.1111/mmi.12544.
- Gerdeman, M. S., Henkin, T. M., et Hines, J. V. (2002). In vitro structure-function studies of the *Bacillus subtilis* *tyrS* mRNA antiterminator: evidence for factor-independent tRNA acceptor stem binding specificity. *Nucleic Acids Res.* 30, 1065-1072. doi:10.1093/nar/30.4.1065.
- Gerdes, K., et Wagner, E. G. H. (2007). RNA antitoxins. *Current Opinion in Microbiology* 10, 117-124. doi:10.1016/j.mib.2007.03.003.
- Ghim, S. Y., et Neuhard, J. (1994). The pyrimidine biosynthesis operon of the thermophile *Bacillus caldolyticus* includes genes for uracil phosphoribosyltransferase and uracil permease. *J Bacteriol* 176, 3698-3707.
- Ghim, S.-Y., Kim, C. C., Bonner, E. R., D'Elia, J. N., Grabner, G. K., et Switzer, R. L. (1999). The *Enterococcus faecalis* *pyr* operon is regulated by autogenous transcriptional attenuation at a single site in the 5' leader. *J Bacteriol* 181, 1324-1329.

- Giovanetti, E., Brenciani, A., Bacciaglia, A., Montanari, M. P., et Varaldo, P. E. (2007). Expressed and unexpressed *tet(M)* genes and the *erm(B)*-carrying Tn1116 transposon in *Streptococcus pyogenes* and *Streptococcus pneumoniae*. *Antimicrob Agents Chemother* 51, 4535; author reply 4535-4536. doi:10.1128/AAC.01155-07.
- Gollnick, P., et Babitzke, P. (2002). Transcription attenuation. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* 1577, 240-250. doi:10.1016/S0167-4781(02)00455-4.
- Goossens, H., Ferech, M., Vander Stichele, R., Elseviers, M., et ESAC Project Group (2005). Outpatient antibiotic use in Europe and association with resistance: a cross-national database study. *Lancet* 365, 579-587. doi:10.1016/S0140-6736(05)17907-0.
- Grady, R., et Hayes, F. (2003). Axe-Txe, a broad-spectrum proteic toxin-antitoxin system specified by a multidrug-resistant, clinical isolate of *Enterococcus faecium*. *Molecular Microbiology* 47, 1419-1432. doi:10.1046/j.1365-2958.2003.03387.x.
- Greenfield, T. J., Franch, T., Gerdes, K., et Weaver, K. E. (2001). Antisense RNA regulation of the *par* post-segregational killing system: structural analysis and mechanism of binding of the antisense RNA, RNAI and its target, RNAI. *Molecular Microbiology* 42, 527-537. doi:10.1046/j.1365-2958.2001.02663.x.
- Grigg, J. C., et Ke, A. (2013). Structural determinants for geometry and information decoding of tRNA by T Box Leader RNA. *Structure* 21. doi:10.1016/j.str.2013.09.001.
- Grossman, T. H. (2016). Tetracycline antibiotics and resistance. *Cold Spring Harbor Perspectives in Medicine* 6. doi:10.1101/cshperspect.a025387.
- Grundy, F. J., et Henkin, T. M. (1993). tRNA as a positive regulator of transcription antitermination in *B. subtilis*. *Cell* 74, 475-482. doi:10.1016/0092-8674(93)80049-k.
- Grundy, F. J., et Henkin, T. M. (1994). Conservation of a transcription antitermination mechanism in aminoacyl-tRNA synthetase and amino acid biosynthesis genes in Gram-positive bacteria. *J. Mol. Biol.* 235, 798-804. doi:10.1006/jmbi.1994.1038.
- Guillaume, G., Ledent, V., Moens, W., et Collard, J.-M. (2004). Phylogeny of efflux-mediated tetracycline resistance genes and related proteins revisited. *Microb Drug Resist* 10, 11-26. doi:10.1089/107662904323047754.
- Gutiérrez-Preciado, A., Henkin, T. M., Grundy, F. J., Yanofsky, C., et Merino, E. (2009). Biochemical features and functional implications of the RNA-based T-box regulatory mechanism. *Microbiol. Mol. Biol. Rev.* 73, 36-61. doi:10.1128/MMBR.00026-08.
- Hanczvikkel, A., et Tóth, Á. (2018). Quantitative study about the role of environmental conditions in the survival capability of multidrug-resistant bacteria. *J Infect Public Health* 11, 801-806. doi:10.1016/j.jiph.2018.05.001.
- Hansen, J. L., Ippolito, J. A., Ban, N., Nissen, P., Moore, P. B., et Steitz, T. A. (2002). The structures of four macrolide antibiotics bound to the large ribosomal subunit. *Molecular Cell* 10, 117-128. doi:10.1016/S1097-2765(02)00570-1.
- Hansen, J. L., Moore, P. B., et Steitz, T. A. (2003). Structures of five antibiotics bound at the peptidyl transferase center of the large ribosomal subunit. *J Mol Biol* 330, 1061-1075. doi:10.1016/s0022-2836(03)00668-5.
- HCSP (2013). Prévention de la transmission croisée des « Bactéries Hautement Résistantes aux antibiotiques émergentes » (BHRe). Paris: Haut Conseil de la Santé Publique.
- HCSP (2019). Actualisation des recommandations relatives aux BHRe. Paris: Haut Conseil de la Santé Publique.
- Hegde, S. S., Vetting, M. W., Mitchenall, L. A., Maxwell, A., et Blanchard, J. S. (2011). Structural and biochemical analysis of the pentapeptide repeat protein EfsQnr, a potent DNA gyrase inhibitor. *Antimicrob Agents Chemother* 55, 110-117. doi:10.1128/AAC.01158-10.
- Hegstad, K., Mikalsen, T., Coque, T. M., Werner, G., et Sundsfjord, A. (2010). Mobile genetic elements and their contribution to the emergence of antimicrobial resistant *Enterococcus faecalis* and *Enterococcus faecium*. *Clinical Microbiology and Infection* 16, 541-554. doi:10.1111/j.1469-0691.2010.03226.x.

- Heidary, M., Khosravi, A. D., Khoshnood, S., Nasiri, M. J., Soleimani, S., et Goudarzi, M. (2018). Daptomycin. *J Antimicrob Chemother* 73, 1-11. doi:10.1093/jac/dkx349.
- Henkin, T. M., Glass, B. L., et Grundy, F. J. (1992). Analysis of the *Bacillus subtilis* *tyrS* gene: conservation of a regulatory sequence in multiple tRNA synthetase genes. *J. Bacteriol.* 174, 1299-1306. doi:10.1128/jb.174.4.1299-1306.1992.
- Ho, S. W., Jung, D., Calhoun, J. R., Lear, J. D., Okon, M., Scott, W. R. P., et al. (2008). Effect of divalent cations on the structure of the antibiotic daptomycin. *Eur Biophys J* 37, 421-433. doi:10.1007/s00249-007-0227-2.
- Holzappel, W., Arini, A., Aeschbacher, M., Coppolecchia, R., et Pot, B. (2018). *Enterococcus faecium* SF68 as a model for efficacy and safety evaluation of pharmaceutical probiotics. *Benef Microbes* 9, 375-388. doi:10.3920/BM2017.0148.
- Hooper, D. C., et Jacoby, G. A. (2015). Mechanisms of drug resistance: quinolone resistance. *Ann N Y Acad Sci* 1354, 12-31. doi:10.1111/nyas.12830.
- Horvath, P., et Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327, 167-170. doi:10.1126/science.1179555.
- Hummel, A., Holzappel, W. H., et Franz, C. M. A. P. (2007). Characterisation and transfer of antibiotic resistance genes from enterococci isolated from food. *Syst Appl Microbiol* 30, 1-7. doi:10.1016/j.syapm.2006.02.004.
- Hutchinson, D. K. (2003). Oxazolidinone antibacterial agents: a critical review. *Curr Top Med Chem* 3, 1021-1042. doi:10.2174/1568026033452195.
- Innocenti, N., Golumbeanu, M., Fouquier d'Hérouël, A., Lacoux, C., Bonnin, R. A., Kennedy, S. P., et al. (2015). Whole-genome mapping of 5' RNA ends in bacteria by tagged sequencing: a comprehensive view in *Enterococcus faecalis*. *RNA* 21, 1018-1030. doi:10.1261/rna.048470.114.
- Inouye, M., et Phadtare, S. (2004). Cold shock response and adaptation at near-freezing temperature in microorganisms. *Sci. STKE*, pe26-pe26. doi:10.1126/stke.2372004pe26.
- Isnard, C. (2017). *Enterococcus* spp. : entre pathogènes opportunistes et probiotiques. *Thèse de science de docteur en Biologie, Université de Normandie*.
- Jaworski, D. D., Flannagan, S. E., et Clewell, D. B. (1996). Analyses of *traA*, *int-Tn*, and *xis-Tn* mutations in the conjugative transposon Tn916 in *Enterococcus faecalis*. *Plasmid* 36, 201-208. doi:10.1006/plas.1996.0047.
- Jett, B. D., Huycke, M. M., et Gilmore, M. S. (1994). Virulence of enterococci. *Clin. Microbiol. Rev.* 7, 462-478. doi:10.1128/CMR.7.4.462.
- Joanne M. Willey, Linda M. Sherwood, et Christopher J. Woolverton (2013). *Microbiologie de Prescott*. 4e édition. De Boeck Supérieur.
- Johnson, C. M., et Grossman, A. D. (2015). Integrative and Conjugative Elements (ICEs): What they do and how they work. *Annu Rev Genet* 49, 577-601. doi:10.1146/annurev-genet-112414-055018.
- Kajfasz, J. K., Mendoza, J. E., Gaca, A. O., Miller, J. H., Koselny, K. A., Giambiagi-deMarval, M., et al. (2012). The Spx regulator modulates stress responses and virulence in *Enterococcus faecalis*. *Infection and Immunity* 80, 2265-2275. doi:10.1128/IAI.00026-12.
- Kalamorz, F., Reichenbach, B., März, W., Rak, B., et Görke, B. (2007). Feedback control of glucosamine-6-phosphate synthase GlmS expression depends on the small RNA GlmZ and involves the novel protein Yhbj in *Escherichia coli*. *Molecular Microbiology* 65, 1518-1533. doi:https://doi.org/10.1111/j.1365-2958.2007.05888.x.
- Kawada-Matsuo, M., Oogai, Y., et Komatsuzawa, H. (2017). Sugar allocation to metabolic pathways is tightly regulated and affects the virulence of *Streptococcus mutans*. *Genes* 8, 11. doi:10.3390/genes8010011.
- Kenig, M., et Abraham, E. P. (1976). Antimicrobial activities and antagonists of bacilysin and anticapsin. *J Gen Microbiol* 94, 37-45. doi:10.1099/00221287-94-1-37.

- Kery, M. B., Feldman, M., Livny, J., et Tjaden, B. (2014). TargetRNA2: identifying targets of small regulatory RNAs in bacteria. *Nucleic Acids Res* 42, W124-W129. doi:10.1093/nar/gku317.
- Khan, M. A., Göpel, Y., Milewski, S., et Görke, B. (2016). Two small RNAs conserved in enterobacteriaceae provide intrinsic resistance to antibiotics targeting the cell wall biosynthesis enzyme glucosamine-6-phosphate synthase. *Front Microbiol* 7, 908. doi:10.3389/fmicb.2016.00908.
- Kim, E. B., et Marco, M. L. (2014). Nonclinical and clinical *Enterococcus faecium* strains, but not *Enterococcus faecalis* strains, have distinct structural and functional genomic features. *Appl Environ Microbiol* 80, 154-165. doi:10.1128/AEM.03108-13.
- Klein, G. (2003). Taxonomy, ecology and antibiotic resistance of enterococci from food and the gastro-intestinal tract. *International Journal of Food Microbiology* 88, 123-131. doi:10.1016/S0168-1605(03)00175-2.
- Klevens, R. M., Edwards, J. R., Richards, C. L., Horan, T. C., Gaynes, R. P., Pollock, D. A., et al. (2007). Estimating health care-associated infections and deaths in U.S. hospitals, 2002. *Public Health Rep* 122, 160-166. doi:10.1177/003335490712200205.
- Kohler, V., Vaishampayan, A., et Grohmann, E. (2018). Broad-host-range Inc18 plasmids: Occurrence, spread and transfer mechanisms. *Plasmid* 99. doi:10.1016/j.plasmid.2018.06.001.
- Kolter, R., et Yanofsky, C. (1982). Attenuation in amino acid biosynthetic operons. *Annu Rev Genet* 16, 113-134. doi:10.1146/annurev.ge.16.120182.000553.
- Kreuzer, K. D., et Henkin, T. M. (2018). The T-Box Riboswitch: tRNA as an effector to modulate gene regulation. *Microbiol Spectr* 6. doi:10.1128/microbiolspec.RWR-0028-2018.
- Kristich, C. J., Rice, L. B., et Arias, C. A. (2014). "Enterococcal infection—treatment and antibiotic resistance," in *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*, eds. M. S. Gilmore, D. B. Clewell, Y. Ike, and N. Shankar (Boston: Massachusetts Eye and Ear Infirmary).
- Lalaouna, D., Eyraud, A., Chabelskaya, S., Felden, B., et Massé, E. (2014). Regulatory RNAs involved in bacterial antibiotic resistance. *PLoS Pathog.* 10, e1004299. doi:10.1371/journal.ppat.1004299.
- Lalaouna, D., et Massé, E. (2015). Identification of sRNA interacting with a transcript of interest using MS2-affinity purification coupled with RNA sequencing (MAPS) technology. *Genom Data* 5, 136-138. doi:10.1016/j.gdata.2015.05.033.
- Lam, M. M. C., Seemann, T., Bulach, D. M., Gladman, S. L., Chen, H., Haring, V., et al. (2012). Comparative analysis of the first complete *Enterococcus faecium* genome. *J. Bacteriol.* 194, 2334-2341. doi:10.1128/JB.00259-12.
- Lancefield, R. C. (1933). A serological differentiation of human and other groups of hemolytic streptococci. *J Exp Med* 57, 571-595.
- Laport, M. S., da Silva, M. R., Silva, C. C., do Carmo de Freire Bastos, M., et Giambiagi-deMarval, M. (2003). Heat-resistance and heat-shock response in the nosocomial pathogen *Enterococcus faecium*. *Curr Microbiol* 46, 0313-0317. doi:10.1007/s00284-002-3828-0.
- Laurberg, M., Asahara, H., Korostelev, A., Zhu, J., Trakhanov, S., et Noller, H. F. (2008). Structural basis for translation termination on the 70S ribosome. *Nature* 454, 852-857. doi:10.1038/nature07115.
- Lautenbach, E., Bilker, W. B., et Brennan, P. J. (1999). Enterococcal bacteremia: risk factors for vancomycin resistance and predictors of mortality. *Infect Control Hosp Epidemiol* 20, 318-323. doi:10.1086/501624.
- Lazaris, A., Coleman, D. C., Kearns, A. M., Pichon, B., Kinnevey, P. M., Earls, M. R., et al. (2017). Novel multiresistance *cfr* plasmids in linezolid-resistant methicillin-resistant *Staphylococcus epidermidis* and vancomycin-resistant *Enterococcus faecium* (VRE) from a hospital outbreak: co-location of *cfr* and *optrA* in VRE. *J Antimicrob Chemother* 72, 3252-3257. doi:10.1093/jac/dkx292.



- Le Neindre, K. (2018). Nouvel ARN exprimé par *Enterococcus faecium* et son implication dans la résistance bactérienne. *Thèse d'exercice de docteur en Pharmacie, Université Rennes 1*.
- Le Neindre, K., Dejoies, L., Reissier, S., Felden, B., et Cattoir, V. (2021). A *cis*-RNA-mediated expression regulation of the *tet(M)* resistance gene in *Enterococcus faecium*. *bioRxiv*, 2021.03.24.436905. doi:10.1101/2021.03.24.436905.
- Leavis, H., Top, J., Shankar, N., Borgen, K., Bonten, M., van Embden, J., *et al.* (2004). A novel putative enterococcal pathogenicity island linked to the *esp* virulence gene of *Enterococcus faecium* and associated with epidemicity. *J Bacteriol* 186, 672-682. doi:10.1128/JB.186.3.672-682.2004.
- Lebreton, F., Manson, A. L., Saavedra, J. T., Straub, T. J., Earl, A. M., et Gilmore, M. S. (2017). Tracing the enterococci from paleozoic origins to the hospital. *Cell* 169, 849-861.e13. doi:10.1016/j.cell.2017.04.027.
- Lebreton, F., Schaik, W. van, McGuire, A. M., Godfrey, P., Griggs, A., Mazumdar, V., *et al.* (2013). Emergence of epidemic multidrug-resistant *Enterococcus faecium* from animal and commensal strains. *mBio* 4, e00534-13. doi:10.1128/mBio.00534-13.
- Leclerc, H., Devriese, L. A., et Mossel, D. a. A. (1996). Taxonomical changes in intestinal (faecal) enterococci and streptococci: consequences on their use as indicators of faecal contamination in drinking water. *Journal of Applied Bacteriology* 81, 459-466. doi:10.1111/j.1365-2672.1996.tb03533.x.van tyn
- Leclercq, R. (2002). Mechanisms of resistance to macrolides and lincosamides: nature of the resistance elements and their clinical implications. *Clin Infect Dis* 34, 482-492. doi:10.1086/324626.
- Lee, T., Pang, S., Abraham, S., et Coombs, G. W. (2019). Antimicrobial-resistant CC17 *Enterococcus faecium*: The past, the present and the future. *J Glob Antimicrob Resist* 16. doi:10.1016/j.jgar.2018.08.016.
- Leipe, D. D., Wolf, Y. I., Koonin, E. V., et Aravind, L. (2002). Classification and evolution of P-loop GTPases and related ATPases. *Journal of Molecular Biology* 317, 41-72. doi:10.1006/jmbi.2001.5378.
- León-Sampedro, R., Novais, C., Peixe, L., Baquero, F., et Coque, T. M. (2016). Diversity and evolution of the Tn5801-*tet(M)*-Like Integrative and Conjugative Elements among *Enterococcus*, *Streptococcus*, and *Staphylococcus*. *Antimicrob. Agents Chemother.* 60, 1736-1746. doi:10.1128/AAC.01864-15.
- Li, D., Wang, Y., Schwarz, S., Cai, J., Fan, R., Li, J., *et al.* (2016). Co-location of the oxazolidinone resistance genes *optrA* and *cfr* on a multiresistance plasmid from *Staphylococcus sciuri*. *J Antimicrob Chemother* 71, 1474-1478. doi:10.1093/jac/dkw040.
- Liang, W., et Deutscher, M. P. (2016). REP sequences: Mediators of the environmental stress response? *RNA Biol* 13, 152-156. doi:10.1080/15476286.2015.1112489.
- Liang, W., Rudd, K. E., et Deutscher, M. P. (2015). A role for REP sequences in regulating translation. *Molecular Cell* 58, 431-439. doi:10.1016/j.molcel.2015.03.019.
- Lin, A. H., Murray, R. W., Vidmar, T. J., et Marotti, K. R. (1997). The oxazolidinone eperzolid binds to the 50S ribosomal subunit and competes with binding of chloramphenicol and lincomycin. *Antimicrob Agents Chemother* 41, 2127-2131. doi:10.1128/AAC.41.10.2127.
- Londei, P. (2005). Evolution of translational initiation: new insights from the archaea. *FEMS Microbiol Rev* 29, 185-200. doi:10.1016/j.femsre.2004.10.002.
- Long, K. S., Poehlsgaard, J., Kehrenberg, C., Schwarz, S., et Vester, B. (2006). The Cfr rRNA methyltransferase confers resistance to phenicols, lincosamides, oxazolidinones, pleuromutilins, and streptogramin A antibiotics. *Antimicrob Agents Chemother* 50, 2500-2505. doi:10.1128/AAC.00131-06.
- Lovett, P. S. (1990). Translational attenuation as the regulator of inducible *cat* genes. *J Bacteriol* 172, 1-6. doi:10.1128/jb.172.1.1-6.1990.

- Lovett, P. S., et Rogers, E. J. (1996). Ribosome regulation by the nascent peptide. *Microbiol Rev* 60, 366-385.
- Lu, Y., Turner, R. J., et Switzer, R. L. (1995). Roles of the three transcriptional attenuators of the *Bacillus subtilis* pyrimidine biosynthetic operon in the regulation of its expression. *J Bacteriol* 177, 1315-1325.
- Luo, X., Hsiao, H.-H., Bubunenko, M., Weber, G., Court, D. L., Gottesman, M. E., et al. (2008). Structural and functional analysis of the *E. coli* NusB-S10 transcription antitermination complex. *Molecular Cell* 32, 791-802. doi:10.1016/j.molcel.2008.10.028.
- Marchand, J. A., Pierson Smela, M. D., Jordan, T. H. H., Narasimhan, K., et Church, G. M. (2020). TBDB: a database of structurally annotated T-box riboswitch:tRNA pairs. *Nucleic Acids Res.* doi:10.1093/nar/gkaa721.
- Marshall, S. H., Donskey, C. J., Hutton-Thomas, R., Salata, R. A., et Rice, L. B. (2002). Gene dosage and linezolid resistance in *Enterococcus faecium* and *Enterococcus faecalis*. *Antimicrobial Agents and Chemotherapy* 46, 3334. doi:10.1128/AAC.46.10.3334-3336.2002.
- McBride, S. M., Fischetti, V. A., LeBlanc, D. J., Moellering, R. C., et Gilmore, M. S. (2007). Genetic diversity among *Enterococcus faecalis*. *PLoS One* 2. doi:10.1371/journal.pone.0000582.
- McCown, P. J., Corbino, K. A., Stav, S., Sherlock, M. E., et Breaker, R. R. (2017). Riboswitch diversity and distribution. *RNA* 23, 995-1011. doi:10.1261/rna.061234.117.
- McCown, P. J., Roth, A., et Breaker, R. R. (2011). An expanded collection and refined consensus model of *glmS* ribozymes. *RNA* 17, 728-736. doi:10.1261/rna.2590811.
- McCown, P. J., Winkler, W. C., et Breaker, R. R. (2012). Mechanism and distribution of *glmS* ribozymes. *Methods Mol Biol* 848, 113-129. doi:10.1007/978-1-61779-545-9\_8.
- McCoy, L. S., Xie, Y., et Tor, Y. (2011). Antibiotics that target protein synthesis. *Wiley Interdiscip Rev RNA* 2, 209-232. doi:10.1002/wrna.60.
- McGuinness, W. A., Malachowa, N., et DeLeo, F. R. (2017). Vancomycin resistance in *Staphylococcus aureus*. *Yale J Biol Med* 90, 269-281.
- Mendes, R. E., Deshpande, L. M., et Jones, R. N. (2014). Linezolid update: stable in vitro activity following more than a decade of clinical use and summary of associated resistance mechanisms. *Drug Resist Updat* 17, 1-12. doi:10.1016/j.drug.2014.04.002.
- Mengin-Lecreux, D., et van Heijenoort, J. (1994). Copurification of glucosamine-1-phosphate acetyltransferase and N-acetylglucosamine-1-phosphate uridylyltransferase activities of *Escherichia coli*: characterization of the *glmU* gene product as a bifunctional enzyme catalyzing two subsequent steps in the pathway for UDP-N-acetylglucosamine synthesis. *J Bacteriol* 176, 5788-5795.
- Michaux, C., Hansen, E. E., Jenniches, L., Gerovac, M., Barquist, L., et Vogel, J. (2020). Single-nucleotide RNA maps for the two major nosocomial pathogens *Enterococcus faecalis* and *Enterococcus faecium*. *Front Cell Infect Microbiol* 10. doi:10.3389/fcimb.2020.600325.
- Michaux, C., Hartke, A., Martini, C., Reiss, S., Albrecht, D., Budin-Verneuil, A., et al. (2014a). Involvement of *Enterococcus faecalis* small RNAs in stress response and virulence. *Infect. Immun.* 82, 3599-3611. doi:10.1128/IAI.01900-14.
- Michaux, C., Verneuil, N., Hartke, A., et Giard, J.-C. (2014b). Physiological roles of small RNA molecules. *Microbiology (Reading, Engl.)* 160, 1007-1019. doi:10.1099/mic.0.076208-0.
- Mika, F., et Hengge, R. (2013). Small regulatory RNAs in the control of motility and biofilm formation in *E. coli* and *Salmonella*. *Int J Mol Sci* 14, 4560-4579. doi:10.3390/ijms14034560.
- Miller, O. L., Hamkalo, B. A., et Thomas, C. A. (1970). Visualization of bacterial genes in action. *Science* 169, 392-395. doi:10.1126/science.169.3943.392.
- Miranda-Ríos, J., Navarro, M., et Soberón, M. (2001). A conserved RNA structure (thi box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. *Proc Natl Acad Sci U S A* 98, 9736-9741. doi:10.1073/pnas.161168098.

- Mironov, A. S., Gusarov, I., Rafikov, R., Lopez, L. E., Shatalin, K., Kreneva, R. A., *et al.* (2002). Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* 111, 747-756. doi:10.1016/S0092-8674(02)01134-0.
- Mishra, N. N., Bayer, A. S., Tran, T. T., Shamoo, Y., Mileykovskaya, E., Dowhan, W., *et al.* (2012). Daptomycin resistance in enterococci is associated with distinct alterations of cell membrane phospholipid content. *PLOS ONE* 7, e43958. doi:10.1371/journal.pone.0043958.
- Moazed, D., et Noller, H. F. (1989). Intermediate states in the movement of transfer RNA in the ribosome. *Nature* 342, 142-148. doi:10.1038/342142a0.
- Moffatt, B. A., et Ashihara, H. (2002). Purine and pyrimidine nucleotide synthesis and metabolism. *Arabidopsis Book* 1. doi:10.1199/tab.0018.
- Molloy, B. B., Lively, D. H., Gale, R. M., Gorman, M., et Boeck, L. D. (1972). A new dipeptide antibiotic from *Streptomyces collinus*, Lindenbein. *J Antibiot (Tokyo)* 25, 137-140. doi:10.7164/antibiotics.25.137.
- Montealegre, M. C., Roh, J. H., Rae, M., Davlieva, M. G., Singh, K. V., Shamoo, Y., *et al.* (2017). Differential penicillin-binding protein 5 (PBP5) levels in the *Enterococcus faecium* clades with different levels of ampicillin resistance. *Antimicrob Agents Chemother* 61. doi:10.1128/AAC.02034-16.
- Murdoch, D. R., Corey, G. R., Hoen, B., Miró, J. M., Fowler, V. G., Jr, *et al.* (2009). Clinical presentation, etiology and outcome of infective endocarditis in the 21st century: the international collaboration on endocarditis-prospective cohort study. *Archives of internal medicine* 169, 463. doi:10.1001/archinternmed.2008.603.
- Murray, B. E. (1990). The life and times of the *Enterococcus*. *Clin. Microbiol. Rev.* 3, 46-65.
- Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L., et Breaker, R. R. (2002). Genetic control by a metabolite binding mRNA. *Chemistry & Biology* 9, 1043-1049. doi:10.1016/S1074-5521(02)00224-7.
- Naville, M., Ghuillot-Gaudeffroy, A., Marchais, A., et Gautheret, D. (2011). ARNold: a web tool for the prediction of Rho-independent transcription terminators. *RNA Biol* 8, 11-13. doi:10.4161/rna.8.1.13346.
- Neely, A. N., et Maley, M. P. (2000). Survival of enterococci and staphylococci on hospital fabrics and plastic. *J Clin Microbiol* 38, 724-726.
- Nelson, A. R., Henkin, T. M., et Agris, P. F. (2006). tRNA regulation of gene expression: interactions of an mRNA 5'-UTR with a regulatory tRNA. *RNA* 12, 1254-1261. doi:10.1261/rna.29906.
- Nelson, M. L., et Levy, S. B. (2011). The history of the tetracyclines. *Annals of the New York Academy of Sciences* 1241, 17-32. doi:10.1111/j.1749-6632.2011.06354.x.
- Nguyen, F., Starosta, A. L., Arenz, S., Sohmen, D., Dönhöfer, A., et Wilson, D. N. (2014). Tetracycline antibiotics and resistance mechanisms. *Biol. Chem.* 395, 559-575. doi:10.1515/hsz-2013-0292.
- Nishimoto, Y., Kobayashi, N., Alam, M. M., Ishino, M., Uehara, N., et Watanabe, N. (2005). Analysis of the prevalence of tetracycline resistance genes in clinical isolates of *Enterococcus faecalis* and *Enterococcus faecium* in a Japanese hospital. *Microb Drug Resist* 11, 146-153. doi:10.1089/mdr.2005.11.146.
- Nissen, P., Hansen, J., Ban, N., Moore, P. B., et Steitz, T. A. (2000). The structural basis of ribosome activity in peptide bond synthesis. *Science* 289, 920-930. doi:10.1126/science.289.5481.920.
- Nissen, P., Kjeldgaard, M., Thirup, S., Clark, B. F. C., et Nyborg, J. (1996). The ternary complex of aminoacylated tRNA and EF-Tu-GTP. Recognition of a bond and a fold. *Biochimie* 78, 921-933. doi:10.1016/S0300-9084(97)86714-4.
- Noble, W. C., Virani, Z., et Cree, R. G. A. (1992). Co-transfer of vancomycin and other resistance genes from *Enterococcus faecalis* NCTC 12201 to *Staphylococcus aureus*. *FEMS Microbiol Lett* 93, 195-198. doi:10.1111/j.1574-6968.1992.tb05089.x.

- Nou, X., et Kadner, R. J. (1998). Coupled changes in translation and transcription during cobalamin-dependent regulation of *btuB* expression in *Escherichia coli*. *J Bacteriol* 180, 6719-6728.
- Novichkov, P. S., Kazakov, A. E., Ravcheev, D. A., Leyn, S. A., Kovaleva, G. Y., Sutormin, R. A., et al. (2013). RegPrecise 3.0--a resource for genome-scale exploration of transcriptional regulation in bacteria. *BMC Genomics* 14, 745. doi:10.1186/1471-2164-14-745.
- Novotna, G., et Janata, J. (2006). A new evolutionary variant of the streptogramin A resistance protein, Vga(A)LC, from *Staphylococcus haemolyticus* with shifted substrate specificity towards lincosamides. *Antimicrob Agents Chemother* 50, 4070-4076. doi:10.1128/AAC.00799-06.
- O'Donovan, G. A., et Neuhard, J. (1970). Pyrimidine metabolism in microorganisms. *Bacteriol Rev* 34, 278-343.
- O'Driscoll, T., et Crank, C. W. (2015). Vancomycin-resistant enterococcal infections: epidemiology, clinical manifestations, and optimal management. *Infect Drug Resist* 8, 217-230. doi:10.2147/IDR.S54125.
- Ogle, J. M., Brodersen, D. E., Clemons, W. M., Tarry, M. J., Carter, A. P., et Ramakrishnan, V. (2001). Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* 292, 897-902. doi:10.1126/science.1060612.
- Oliveira, D. M. P. D., Forde, B. M., Kidd, T. J., Harris, P. N. A., Schembri, M. A., Beatson, S. A., et al. (2020). Antimicrobial resistance in ESKAPE pathogens. *Clinical Microbiology Reviews* 33. doi:10.1128/CMR.00181-19.
- Olsen, L. R., Vetting, M. W., et Roderick, S. L. (2007). Structure of the *E. coli* bifunctional GlmU acetyltransferase active site with substrates and products. *Protein Sci* 16, 1230-1235. doi:10.1110/ps.072779707.
- Oyamada, Y., Ito, H., Fujimoto, K., Asada, R., Niga, T., Okamoto, R., et al. (2006). Combination of known and unknown mechanisms confers high-level resistance to fluoroquinolones in *Enterococcus faecium*. *Journal of Medical Microbiology* 55, 729-736. doi:10.1099/jmm.0.46303-0.
- Palmer, K. L., et Gilmore, M. S. (2010). Multidrug-resistant enterococci lack CRISPR-cas. *mBio* 1. doi:10.1128/mBio.00227-10.
- Palmer, K. L., Kos, V. N., et Gilmore, M. S. (2010). Horizontal gene transfer and the genomics of enterococcal antibiotic resistance. *Curr Opin Microbiol* 13, 632-639. doi:10.1016/j.mib.2010.08.004.
- Paulsen, I. T., Brown, M. H., et Skurray, R. A. (1996). Proton-dependent multidrug efflux systems. *Microbiol Rev* 60, 575-608.
- Pinto, B., Pierotti, R., Canale, G., et Reali, D. (1999). Characterization of 'faecal streptococci' as indicators of faecal pollution and distribution in the environment. *Letters in Applied Microbiology* 29, 258-263. doi:10.1046/j.1472-765X.1999.00633.x.
- Pioletti, M., Schlünzen, F., Harms, J., Zarivach, R., Glühmann, M., Avila, H., et al. (2001). Crystal structures of complexes of the small ribosomal subunit with tetracycline, edeine and IF3. *EMBO J* 20, 1829-1839. doi:10.1093/emboj/20.8.1829.
- Pitman, S., et Cho, K. H. (2015). The mechanisms of virulence regulation by small noncoding RNAs in low GC Gram-positive pathogens. *Int J Mol Sci* 16, 29797-29814. doi:10.3390/ijms161226194.
- Plumbridge, J. (1995). Co-ordinated regulation of amino sugar biosynthesis and degradation: the NagC repressor acts as both an activator and a repressor for the transcription of the *glmUS* operon and requires two separated NagC binding sites. *The EMBO Journal* 14, 3958-3965. doi:10.1002/j.1460-2075.1995.tb00067.x.
- Portillo, A., Ruiz-Larrea, F., Zarazaga, M., Alonso, A., Martinez, J. L., et Torres, C. (2000). Macrolide resistance genes in *Enterococcus* spp. *Antimicrob Agents Chemother* 44, 967.

- Poulsen, S. M., Kofoed, C., et Vester, B. (2000). Inhibition of the ribosomal peptidyl transferase reaction by the mycarose moiety of the antibiotics carbomycin, spiramycin and tylosin. *Journal of Molecular Biology* 304, 471-481. doi:10.1006/jmbi.2000.4229.
- Putzer, H., et Laalami, S. (2013). *Regulation of the Expression of Aminoacyl-tRNA Synthetases and Translation Factors*. Landes Bioscience.
- Putzer, H., Laalami, S., Brakhage, A. A., Condon, C., et Grunberg-Manago, M. (1995). Aminoacyl-tRNA synthetase gene regulation in *Bacillus subtilis*: induction, repression and growth-rate regulation. *Mol. Microbiol.* 16, 709-718. doi:10.1111/j.1365-2958.1995.tb02432.x.
- Qin, X., Galloway-Peña, J. R., Sillanpaa, J., Roh, J. H., Nallapareddy, S. R., Chowdhury, S., et al. (2012). Complete genome sequence of *Enterococcus faecium* strain TX16 and comparative genomic analysis of *Enterococcus faecium* genomes. *BMC Microbiol* 12, 135. doi:10.1186/1471-2180-12-135.
- Quereda, J. J., et Cossart, P. (2017). Regulating bacterial virulence with RNA. *Annu Rev Microbiol* 71, 263-280. doi:10.1146/annurev-micro-030117-020335.
- Rajendran, V., Kalita, P., Shukla, H., Kumar, A., et Tripathi, T. (2018). Aminoacyl-tRNA synthetases: Structure, function, and drug discovery. *International Journal of Biological Macromolecules* 111, 400-414. doi:10.1016/j.ijbiomac.2017.12.157.
- Ramirez, M. S., et Tolmasky, M. E. (2010). Aminoglycoside modifying enzymes. *Drug Resist Updat* 13, 151-171. doi:10.1016/j.drup.2010.08.003.
- Ramu, H., Mankin, A., et Vazquez-Laslop, N. (2009). Programmed drug-dependent ribosome stalling. *Molecular Microbiology* 71, 811-824. doi:https://doi.org/10.1111/j.1365-2958.2008.06576.x.
- Ranaweera, I., Shrestha, U., Ranjana, K. C., Kakarla, P., Willmon, T. M., Hernandez, A. J., et al. (2015). Structural comparison of bacterial multidrug efflux pumps of the major facilitator superfamily. *Trends Cell Mol Biol* 10, 131-140.
- Reese, M. G. (2001). Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. *Comput Chem* 26, 51-56. doi:10.1016/s0097-8485(01)00099-7.
- Reichenbach, B., Maes, A., Kalamorz, F., Hajnsdorf, E., et Görke, B. (2008). The small RNA GlmY acts upstream of the sRNA GlmZ in the activation of *glmS* expression and is subject to regulation by polyadenylation in *Escherichia coli*. *Nucleic Acids Res* 36, 2570-2580. doi:10.1093/nar/gkn091.
- Reissier, S., et Cattoir, V. (2020). Streptogramins for the treatment of infections caused by Gram-positive pathogens. *Expert Rev Anti Infect Ther*, 1-13. doi:10.1080/14787210.2021.1834851.
- Reissier, S., Le Neindre, K., Bordeau, V., Dejoies, L., Le Bot, A., Felden, B., Cattoir, V., et Revest, M. (2021). The regulatory RNA *ern0160* confers a potential selective advantage to *Enterococcus faecium* for intestinal colonization. *Front. Microbiol.* 12:757227. Doi:10.3389/fmicb.2021.757227.
- Reynolds, E., Ross, J. I., et Cove, J. H. (2003). Msr(A) and related macrolide/streptogramin resistance determinants: incomplete transporters? *Int J Antimicrob Agents* 22, 228-236. doi:10.1016/s0924-8579(03)00218-8.
- Rice, L. B. (1998). Tn916 family conjugative transposons and dissemination of antimicrobial resistance determinants. *Antimicrob Agents Chemother* 42, 1871-1877.
- Rice, L. B. (2008). Federal funding for the study of antimicrobial resistance in nosocomial pathogens: no ESKAPE. *J Infect Dis.* 197, 1079-1081. doi:10.1086/533452.
- Rice, L. B., Bellais, S., Carias, L. L., Hutton-Thomas, R., Bonomo, R. A., Caspers, P., et al. (2004). Impact of specific *pbp5* mutations on expression of  $\beta$ -lactam resistance in *Enterococcus faecium*. *Antimicrob Agents Chemother* 48, 3028. doi:10.1128/AAC.48.8.3028-3032.2004.
- Rice, L. B., Carias, L. L., Donskey, C. L., et Rudin, S. D. (1998). Transferable, plasmid-mediated *vanB*-type glycopeptide resistance in *Enterococcus faecium*. *Antimicrob Agents Chemother* 42, 963-964.

- Roberts, A. P., et Mullany, P. (2009). A modular master on the move: the Tn916 family of mobile genetic elements. *Trends Microbiol* 17, 251-258. doi:10.1016/j.tim.2009.03.002.
- Roberts, A. P., et Mullany, P. (2011). Tn916-like genetic elements: a diverse group of modular mobile elements conferring antibiotic resistance. *FEMS Microbiol Rev* 35, 856-871. doi:10.1111/j.1574-6976.2011.00283.x.
- Roberts, M. C. (2008). Update on macrolide-lincosamide-streptogramin, ketolide, and oxazolidinone resistance genes. *FEMS Microbiol Lett* 282, 147-159. doi:10.1111/j.1574-6968.2008.01145.x.
- Rodionov, D. A., Mironov, A. A., et Gelfand, M. S. (2002). Conservation of the biotin regulon and the BirA regulatory signal in eubacteria and archaea. *Genome Res* 12, 1507-1516. doi:10.1101/gr.314502.
- Rodnina, M. V., Savelsbergh, A., Katunin, V. I., et Wintermeyer, W. (1997). Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome. *Nature* 385, 37-41. doi:10.1038/385037a0.
- Rodríguez-Díaz, J., Rubio-del-Campo, A., et Yebra, M. J. (2012a). Metabolic engineering of *Lactobacillus casei* for production of UDP-N-acetylglucosamine. *Biotechnology and Bioengineering* 109, 1704-1712. doi:https://doi.org/10.1002/bit.24475.
- Rodríguez-Díaz, J., Rubio-del-Campo, A., et Yebra, M. J. (2012b). Regulatory insights into the production of UDP-N-acetylglucosamine by *Lactobacillus casei*. *Bioengineered* 3, 339-342. doi:10.4161/bioe.21271.
- Rollins, S. M., Grundy, F. J., et Henkin, T. M. (1997). Analysis of *cis*-acting sequence and structural elements required for antitermination of the *Bacillus subtilis* *tyrS* gene. *Mol. Microbiol.* 25, 411-421. doi:10.1046/j.1365-2958.1997.4851839.x.
- Rosvoll, T. C. S., Pedersen, T., Sletvold, H., Johnsen, P. J., Sollid, J. E., Simonsen, G. S., et al. (2010). PCR-based plasmid typing in *Enterococcus faecium* strains reveals widely distributed pRE25-, pRUM-, pIP501- and pHTbeta-related replicons associated with glycopeptide resistance and stabilizing toxin-antitoxin systems. *FEMS Immunol Med Microbiol* 58, 254-268. doi:10.1111/j.1574-695X.2009.00633.x.
- Ruiz-Masó, J. A., Machón, C., Bordanaba-Ruiseco, L., Espinosa, M., Coll, M., et Del Solar, G. (2015). Plasmid Rolling-Circle Replication. *Microbiol Spectr* 3, PLAS-0035-2014. doi:10.1128/microbiolspec.PLAS-0035-2014.
- Rybák, M. J., Hershberger, E., Moldovan, T., et Grucz, R. G. (2000). In vitro activities of daptomycin, vancomycin, linezolid, and quinupristin-dalfopristin against staphylococci and enterococci, including vancomycin- intermediate and -resistant strains. *Antimicrob Agents Chemother* 44, 1062-1066.
- Saad, N. Y., Stamatopoulou, V., Brayé, M., Drainas, D., Stathopoulos, C., et Becker, H. D. (2013). Two-codon T-box riboswitch binding two tRNAs. *Proc. Natl. Acad. Sci. U.S.A.* 110, 12756-12761. doi:10.1073/pnas.1304307110.
- Sahm, D. F., Kissinger, J., Gilmore, M. S., Murray, P. R., Mulder, R., Solliday, J., et al. (1989). In vitro susceptibility studies of vancomycin-resistant *Enterococcus faecalis*. *Antimicrob Agents Chemother* 33, 1588-1591. doi:10.1128/aac.33.9.1588.
- Saito, S., Kakeshita, H., et Nakamura, K. (2009). Novel small RNA-encoding genes in the intergenic regions of *Bacillus subtilis*. *Gene* 428, 2-8. doi:10.1016/j.gene.2008.09.024.
- Salyers, A. A., Shoemaker, N. B., Stevens, A. M., et Li, L. Y. (1995). Conjugative transposons: an unusual and diverse set of integrated gene transfer elements. *Microbiol Rev* 59, 579-590.
- Santé Publique France (2019). Enquête nationale de prévalence des infections nosocomiales et des traitements anti-infectieux en établissements de santé, France, mai-juin 2017.
- Sassi, M., Guérin, F., Zouari, A., Beyrouthy, R., Auzou, M., Fines-Guyon, M., et al. (2019). Emergence of *optrA*-mediated linezolid resistance in enterococci from France, 2006-16. *J Antimicrob Chemother* 74, 1469-1472. doi:10.1093/jac/dkz097.
- Savinov, A., et Block, S. M. (2018). Self-cleavage of the *glmS* ribozyme core is controlled by a fragile folding element. *PNAS* 115, 11976-11981. doi:10.1073/pnas.1812122115.

- Schedlbauer, A., Kaminishi, T., Ochoa-Lizarralde, B., Dhimole, N., Zhou, S., López-Alonso, J. P., *et al.* (2015). Structural characterization of an alternative mode of tigecycline binding to the bacterial ribosome. *Antimicrob Agents Chemother* 59, 2849-2854. doi:10.1128/AAC.04895-14.
- Schleifer, K. H., et Kilpper-Bälz, R. (1984). Transfer of *Streptococcus faecalis* and *Streptococcus faecium* to the genus *Enterococcus* nom. rev. as *Enterococcus faecalis* comb. nov. and *Enterococcus faecium* comb. nov. *International Journal of Systematic and Evolutionary Microbiology* 34, 31-34. doi:10.1099/00207713-34-1-31.
- Schmeing, T. M., et Ramakrishnan, V. (2009). What recent ribosome structures have revealed about the mechanism of translation. *Nature* 461, 1234-1242. doi:10.1038/nature08403.
- Schoenfelder, S. M. K., Marincola, G., Geiger, T., Goerke, C., Wolz, C., et Ziebuhr, W. (2013). Methionine biosynthesis in *Staphylococcus aureus* is tightly controlled by a hierarchical network involving an initiator tRNA-specific T-box riboswitch. *PLoS Pathog.* 9, e1003606. doi:10.1371/journal.ppat.1003606.
- Schwarz, F. V., Perreten, V., et Teuber, M. (2001). Sequence of the 50-kb conjugative multiresistance plasmid pRE25 from *Enterococcus faecalis* RE25. *Plasmid* 46, 170-187. doi:10.1006/plas.2001.1544.
- Scornec, H., Bellanger, X., Guilloteau, H., Groshenry, G., et Merlin, C. (2017). Inducibility of Tn916 conjugative transfer in *Enterococcus faecalis* by subinhibitory concentrations of ribosome-targeting antibiotics. *J Antimicrob Chemother* 72, 2722-2728. doi:10.1093/jac/dkx202.
- Scott, J. R., Bringel, F., Marra, D., Alstine, G. V., et Rudy, C. K. (1994). Conjugative transposition of Tn916: preferred targets and evidence for conjugative transfer of a single strand and for a double-stranded circular intermediate. *Molecular Microbiology* 11, 1099-1108. doi:https://doi.org/10.1111/j.1365-2958.1994.tb00386.x.
- Seier-Petersen, M. A., Jasni, A., Aarestrup, F. M., Vigre, H., Mullany, P., Roberts, A. P., *et al.* (2014). Effect of subinhibitory concentrations of four commonly used biocides on the conjugative transfer of Tn916 in *Bacillus subtilis*. *J Antimicrob Chemother* 69, 343-348. doi:10.1093/jac/dkt370.
- Seliverstov, A. V., Putzer, H., Gelfand, M. S., et Lyubetsky, V. A. (2005). Comparative analysis of RNA regulatory elements of amino acid metabolism genes in Actinobacteria. *BMC Microbiol* 5, 54. doi:10.1186/1471-2180-5-54.
- Selton-Suty, C., Hoen, B., Grentzinger, A., Houplon, P., Maignan, M., Juillière, Y., *et al.* (1997). Clinical and bacteriological characteristics of infective endocarditis in the elderly. *Heart* 77, 260-263.
- Serganov, A., et Nudler, E. (2013). A decade of riboswitches. *Cell* 152, 17-24. doi:10.1016/j.cell.2012.12.024.
- Sherman, J. M. (1937). The streptococci. *Bacteriological Reviews* 1, 3.
- Shioya, K., Michaux, C., Kuenne, C., Hain, T., Verneuil, N., Budin-Verneuil, A., *et al.* (2011). Genome-Wide identification of small RNAs in the opportunistic pathogen *Enterococcus faecalis* V583. *PLoS ONE* 6. doi:10.1371/journal.pone.0023948.
- Sifaoui, F., Arthur, M., Rice, L., et Gutmann, L. (2001). Role of penicillin-binding protein 5 in expression of ampicillin resistance and peptidoglycan structure in *Enterococcus faecium*. *Antimicrobial Agents and Chemotherapy* 45, 2594. doi:10.1128/AAC.45.9.2594-2597.2001.
- Siguier, P., Gournayre, E., et Chandler, M. (2014). Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol Rev* 38, 865-891. doi:10.1111/1574-6976.12067.
- Sinel, C., Augagneur, Y., Sassi, M., Bronsard, J., Cacaci, M., Guérin, F., *et al.* (2017a). Small RNAs in vancomycin-resistant *Enterococcus faecium* involved in daptomycin response and resistance. *Scientific Reports* 7. doi:10.1038/s41598-017-11265-2.
- Sinel, C., Cacaci, M., Meignen, P., Guérin, F., Davies, B. W., Sanguinetti, M., *et al.* (2017b). Subinhibitory concentrations of ciprofloxacin enhance antimicrobial resistance and

- pathogenicity of *Enterococcus faecium*. *Antimicrob. Agents Chemother.* doi:10.1128/AAC.02763-16.
- Sinel, C., Cosquer, T., Auzou, M., Goux, D., Giard, J.-C., et Cattoir, V. (2016). Sequential steps of daptomycin resistance in *Enterococcus faecium* and reversion to hypersusceptibility through IS-mediated inactivation of the *liaFSR* operon. *J Antimicrob Chemother* 71, 2793-2797. doi:10.1093/jac/dkw229.
- Su, Y. A., et Clewell, D. B. (1993). Characterization of the left 4 kb of conjugative transposon Tn916: determinants involved in excision. *Plasmid* 30, 234-250. doi:10.1006/plas.1993.1055.
- Su, Y. A., He, P., et Clewell, D. B. (1992). Characterization of the *tet(M)* determinant of Tn916: evidence for regulation by transcription attenuation. *Antimicrob. Agents Chemother.* 36, 769.
- Sutcliffe, J. A., O'Brien, W., Fyfe, C., et Grossman, T. H. (2013). Antibacterial activity of eravacycline (TP-434), a novel fluorocycline, against hospital and community pathogens. *Antimicrob. Agents Chemother.* 57, 5548-5558. doi:10.1128/AAC.01288-13.
- Switzer, R. L., Turner, R. J., et Lu, Y. (1999). Regulation of the *Bacillus subtilis* pyrimidine biosynthetic operon by transcriptional attenuation: control of gene expression by an mRNA-binding protein. *Prog. Nucleic Acid Res. Mol. Biol.* 62, 329-367. doi:10.1016/s0079-6603(08)60512-7.
- Taylor, S. D., et Palmer, M. (2016). The action mechanism of daptomycin. *Bioorganic & Medicinal Chemistry* 24, 6253-6268. doi:10.1016/j.bmc.2016.05.052.
- Tenson, T., Lovmar, M., et Ehrenberg, M. (2003). The mechanism of action of macrolides, lincosamides and streptogramin B reveals the nascent peptide exit path in the ribosome. *J Mol Biol* 330, 1005-1014. doi:10.1016/s0022-2836(03)00662-4.
- Thaker, M., Spanogiannopoulos, P., et Wright, G. D. (2010). The tetracycline resistome. *Cell. Mol. Life Sci.* 67, 419-431. doi:10.1007/s00018-009-0172-6.
- Top, J., Willems, R., Blok, H., de Regt, M., Jalink, K., Troelstra, A., et al. (2007). Ecological replacement of *Enterococcus faecalis* by multiresistant clonal complex 17 *Enterococcus faecium*. *Clin Microbiol Infect* 13, 316-319. doi:10.1111/j.1469-0691.2006.01631.x.
- Tran, J. H., et Jacoby, G. A. (2002). Mechanism of plasmid-mediated quinolone resistance. *Proc. Natl. Acad. Sci. U.S.A.* 99, 5638-5642. doi:10.1073/pnas.082092899.
- Treviño, J., Perez, N., et Sumby, P. (2010). The 4.5S RNA component of the signal recognition particle is required for group A *Streptococcus* virulence. *Microbiology (Reading)* 156, 1342-1350. doi:10.1099/mic.0.036558-0.
- Trieu-Cuot, P., Poyart-Salmeron, C., Carlier, C., et Courvalin, P. (1993). Sequence requirements for target activity in site-specific recombination mediated by the *Int* protein of transposon Tn1545. *Molecular Microbiology* 8, 179-185. doi:https://doi.org/10.1111/j.1365-2958.1993.tb01214.x.
- Turnbough, C. L., et Switzer, R. L. (2008). Regulation of pyrimidine biosynthetic gene expression in bacteria: repression without repressors. *Microbiol. Mol. Biol. Rev.* 72, 266-300, table of contents. doi:10.1128/MMBR.00001-08.
- Turner, R. J., Bonner, E. R., Grabner, G. K., et Switzer, R. L. (1998). Purification and Characterization of *Bacillus subtilis* PyrR, a bifunctional *pyr* mRNA-binding attenuation protein/uracil phosphoribosyltransferase. *Journal of Biological Chemistry* 273, 5932-5938. doi:10.1074/jbc.273.10.5932.
- Turner, R. J., Lu, Y., et Switzer, R. L. (1994). Regulation of the *Bacillus subtilis* pyrimidine biosynthetic (*pyr*) gene cluster by an autogenous transcriptional attenuation mechanism. *J Bacteriol* 176, 3708-3722.
- Van Melder, L., et Saavedra De Bast, M. (2009). Bacterial toxin-antitoxin systems: more than selfish entities? *PLoS Genet* 5, e1000437. doi:10.1371/journal.pgen.1000437.
- van Schaik, W., Top, J., Riley, D. R., Boekhorst, J., Vrijenhoek, J. E., Schapendonk, C. M., et al. (2010). Pyrosequencing-based comparative genome analysis of the nosocomial pathogen



- Enterococcus faecium* and identification of a large transferable pathogenicity island. *BMC Genomics* 11, 239. doi:10.1186/1471-2164-11-239.
- Van Tyne, D., et Gilmore, M. S. (2014). Friend turned foe: evolution of enterococcal virulence and antibiotic resistance. *Annu Rev Microbiol* 68, 337-356. doi:10.1146/annurev-micro-091213-113003.
- Vázquez-Laslop, N., et Mankin, A. S. (2018). How macrolide antibiotics work. *Trends Biochem Sci* 43, 668-684. doi:10.1016/j.tibs.2018.06.011.
- Verneuil, N., Mazé, A., Sanguinetti, M., Laplace, J.-M., Benachour, A., Auffray, Y., et al. (2006). Implication of (Mn)superoxide dismutase of *Enterococcus faecalis* in oxidative stress responses and survival inside macrophages. *Microbiology (Reading)* 152, 2579-2589. doi:10.1099/mic.0.28922-0.
- Vester, B., et Douthwaite, S. (2001). Macrolide resistance conferred by base substitutions in 23S rRNA. *Antimicrob Agents Chemother* 45, 1-12. doi:10.1128/aac.45.1.1-12.2001.
- Vimberg, V., Zieglerová, L., Buriánková, K., Branny, P., et Balíková Novotná, G. (2020). VanZ reduces the binding of lipoglycopeptide antibiotics to *Staphylococcus aureus* and *Streptococcus pneumoniae* cells. *Front. Microbiol.* 11. doi:10.3389/fmicb.2020.00566.
- Vincent, J.-L., Rello, J., Marshall, J., Silva, E., Anzueto, A., Martin, C. D., et al. (2009). International study of the prevalence and outcomes of infection in intensive care units. *JAMA* 302, 2323-2329. doi:10.1001/jama.2009.1754.
- Vitreschak, A. G., Lyubetskaya, E. V., Shirshin, M. A., Gelfand, M. S., et Lyubetsky, V. A. (2004). Attenuation regulation of amino acid biosynthetic operons in proteobacteria: comparative genomics analysis. *FEMS Microbiology Letters* 234, 357-370. doi:https://doi.org/10.1111/j.1574-6968.2004.tb09555.x.
- Vitreschak, A. G., Mironov, A. A., Lyubetsky, V. A., et Gelfand, M. S. (2008). Comparative genomic analysis of T-box regulatory systems in bacteria. *RNA* 14, 717-735. doi:10.1261/rna.819308.
- Wang, H., et Mullany, P. (2000). The large resolvase TndX is required and sufficient for integration and excision of derivatives of the novel conjugative transposon Tn5397. *J Bacteriol* 182, 6577-6583.
- Wang, H., Roberts, A. P., Lyras, D., Rood, J. I., Wilks, M., et Mullany, P. (2000). Characterization of the ends and target sites of the novel conjugative transposon Tn5397 from *Clostridium difficile*: excision and circularization is mediated by the large resolvase, TndX. *J Bacteriol* 182, 3775-3783. doi:10.1128/jb.182.13.3775-3783.2000.
- Wang, H., Smith, M. C. M., et Mullany, P. (2006). The conjugative transposon Tn5397 has a strong preference for integration into its *Clostridium difficile* target site. *J Bacteriol* 188, 4871-4878. doi:10.1128/JB.00210-06.
- Wang, J., et Nikonowicz, E. P. (2011). Solution structure of the K-turn and Specifier Loop domains from the *Bacillus subtilis* tyrS T-box leader RNA. *J Mol Biol* 408, 99-117. doi:10.1016/j.jmb.2011.02.014.
- Wang, Y., Lv, Y., Cai, J., Schwarz, S., Cui, L., Hu, Z., et al. (2015). A novel gene, *optrA*, that confers transferable resistance to oxazolidinones and phenicols and its presence in *Enterococcus faecalis* and *Enterococcus faecium* of human and animal origin. *J Antimicrob Chemother* 70, 2182-2190. doi:10.1093/jac/dkv116.
- Wardal, E., Sadowy, E., et Hryniewicz, W. (2010). Complex nature of enterococcal pheromone-responsive plasmids. *Pol J Microbiol* 59, 79-87.
- Waters, L. S., et Storz, G. (2009). Regulatory RNAs in Bacteria. *Cell* 136, 615-628. doi:10.1016/j.cell.2009.01.043.
- Weaver, K. E., Jensen, K. D., Colwell, A., et Sriram, S. I. (1996). Functional analysis of the *Enterococcus faecalis* plasmid pAD1-encoded stability determinant *par*. *Mol Microbiol* 20, 53-63. doi:10.1111/j.1365-2958.1996.tb02488.x.
- Weaver, K. E., Reddy, S. G., Brinkman, C. L., Patel, S., Bayles, K. W., et Endres, J. L. (2009). Identification and characterization of a family of toxin-antitoxin systems related to the

- Enterococcus faecalis* plasmid pAD1 par addiction module. *Microbiology (Reading)* 155, 2930-2940. doi:10.1099/mic.0.030932-0.
- Weaver, K. E., Rice, L. B., et Churchward, G. (2002). "Plasmids and Transposons," in *The Enterococci* (John Wiley & Sons, Ltd), 219-263. doi:10.1128/9781555817923.ch6.
- Weaver, K. E., Weaver, D. M., Wells, C. L., Waters, C. M., Gardner, M. E., et Ehli, E. A. (2003). *Enterococcus faecalis* plasmid pAD1-encoded Fst toxin affects membrane permeability and alters cellular responses to lantibiotics. *J Bacteriol* 185, 2169-2177. doi:10.1128/jb.185.7.2169-2177.2003.
- Wei, Y., et Bechhofer, D. H. (2002). Tetracycline induces stabilization of mRNA in *Bacillus subtilis*. *Journal of Bacteriology* 184, 889-894. doi:10.1128/jb.184.4.889-894.2002.
- Weiner, L. M., Webb, A. K., Limbago, B., Dudeck, M. A., Patel, J., Kallen, A. J., et al. (2016). Antimicrobial-resistant pathogens associated with healthcare-associated infections: summary of data reported to the national healthcare safety network at the centers for disease control and prevention, 2011-2014. *Infect Control Hosp Epidemiol* 37, 1288-1301. doi:10.1017/ice.2016.174.
- Wendt, C., Wiesenthal, B., Dietz, E., et Rüden, H. (1998). Survival of vancomycin-resistant and vancomycin-susceptible enterococci on dry surfaces. *J Clin Microbiol* 36, 3734-3736.
- Werner, G., Hildebrandt, B., et Witte, W. (2001). Aminoglycoside-streptothricin resistance gene cluster *aadE-sat4-aphA-3* disseminated among multiresistant isolates of *Enterococcus faecium*. *Antimicrob. Agents Chemother.* 45, 3267-3269. doi:10.1128/AAC.45.11.3267-3269.2001.
- Willems, R. J. L., et Bonten, M. J. M. (2007). Glycopeptide-resistant enterococci: deciphering virulence, resistance and epidemicity. *Curr. Opin. Infect. Dis.* 20, 384-390. doi:10.1097/QCO.0b013e32818be63d.
- Wilson, D. N., Schlutzen, F., Harms, J. M., Starosta, A. L., Connell, S. R., et Fucini, P. (2008). The oxazolidinone antibiotics perturb the ribosomal peptidyl-transferase center and effect tRNA positioning. *Proc Natl Acad Sci U S A* 105, 13339-13344. doi:10.1073/pnas.0804276105.
- Winkler, W. C., Grundy, F. J., Murphy, B. A., et Henkin, T. M. (2001). The GA motif: an RNA element common to bacterial antitermination systems, rRNA, and eukaryotic RNAs. *RNA* 7, 1165-1172. doi:10.1017/s1355838201002370.
- Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A., et Breaker, R. R. (2004). Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* 428, 281-286. doi:10.1038/nature02362.
- Wozniak, R. A. F., et Waldor, M. K. (2010). Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nature Reviews Microbiology* 8, 552-563. doi:10.1038/nrmicro2382.
- Wright, P. R., Georg, J., Mann, M., Sorescu, D. A., Richter, A. S., Lott, S., et al. (2014). CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains. *Nucleic Acids Res.* 42, W119-123. doi:10.1093/nar/gku359.
- Yanofsky, C. (1981). Attenuation in the control of expression of bacterial operons. *Nature* 289, 751-758. doi:10.1038/289751a0.
- Yousef, M. R., Grundy, F. J., et Henkin, T. M. (2003). tRNA requirements for glyQS antitermination: a new twist on tRNA. *RNA* 9, 1148-1156. doi:10.1261/rna.5540203.
- Zaher, H. S., et Green, R. (2009). Quality control by the ribosome following peptide bond formation. *Nature* 457, 161-166. doi:10.1038/nature07582.
- Zhang, H., et Switzer, R. L. (2003). Transcriptional pausing in the *Bacillus subtilis* *pyr* operon in vitro: a role in transcriptional attenuation? *J Bacteriol* 185, 4764-4771. doi:10.1128/JB.185.16.4764-4771.2003.
- Zhang, J., et Ferré-D'Amaré, A. R. (2013). Co-crystal structure of a T-box riboswitch stem I domain in complex with its cognate tRNA. *Nature* 500, 363-366. doi:10.1038/nature12440.

- Zhong, Z., Kwok, L.-Y., Hou, Q., Sun, Y., Li, W., Zhang, H., *et al.* (2019). Comparative genomic analysis revealed great plasticity and environmental adaptation of the genomes of *Enterococcus faecium*. *BMC Genomics* 20, 602. doi:10.1186/s12864-019-5975-8.
- Zielenkiewicz, U., et Cegłowski, P. (2005). The toxin-antitoxin system of the streptococcal plasmid pSM19035. *J Bacteriol* 187, 6094-6105. doi:10.1128/JB.187.17.6094-6105.2005.
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* 31, 3406.

## Annexes

# Matériels et méthodes supplémentaires

Tableau supplémentaire 1. Oligonucléotides utilisés non décrits dans les articles

Nom	Séquence nucléotidique (5'-3')	Cible et expérience
NBern0120	CTCAAGCAGACAAGCCCTTC	NB sondes
NBern0160	ATACATGCAAAGCACGGTGT	
NBern0280	TCGCCTTTATTGATTTTACTTCTTG	
NBern0980	GGTTTTAAACTGTCCACCTTCAC	
NBern1260	TCTCATTTAAACGATGTATTCCTAA	
NBern1300	TGCATACATACAATTTGGAACATT	
NBern2020	GGGAGTCGACGTATTTTTTAAGGCTC	
qPCRsRNA0120F	ATCACCTGCCGGATGCAG	Amorces qPCR
qPCRsRNA0120R	TACTCAAGCAGACAAGCCCT	
qPCRsRNA0160F	ACCTCGGAAGGATTCAAATAAACA	
qPCRsRNA0160R	CAACACCGTGCTTTGCATGT	
qPCRsRNA2050F	GAAATTCATACACGGTACCACCT	
qPCRsRNA2050R	CGCGTACGAGCGTTATTCG	
E1162cont00130F	GGAAGATTGTAAATAACAATATTGGTAC	E1162 vérification
E1162contr00022R	GTTTCTCTTAATATCCTGATAAACCG	
sRNA0120F1	GCAAGACTTTTGTGGAGAAGGGC	5'3'RACE <i>ern0120</i>
sRNA0120F2	CTGCTTGAGTAATCAAGTACCG	
sRNA0120R1	GACACCCCTGCATCCGGCAGGTG	
sRNA0120R2	GCATCCGGCAGGTGATAAAGATAG	
sRNA0160F1	CACCGTGCTTTGCATGTATG	5'3'RACE <i>ern0160</i>
sRNA0160F2	CTGTTTATTTGAATCCTTCCGAGG	
sRNA0160R1	GGTTTGTACTCTACACGCATG	
sRNA0160R2	GTACTTCTAACAACCTTTGTTGC	
KOsRNA0160F1	TTGACCCGGGCTATCTGACACTACATGTAG	Délétion et <i>trans-</i> complémentation <i>ern0160</i>
KOsRNA0160F2	CTCTCGATAAGCTCGATCTATATAATAGCATCCTTTG	
KOsRNA0160R1	CGAGCTTATCGAGAGCTCGGCTTTTTTC	
KOsRNA0160R2	ACACCCGGGTTTGTTTATTCATATCTTGC	
KOsRNA0160vF	GTGCTTATATTTGCAAACGCTACCAAAGG	
KOsRNA0160vR	GGAGATGCGATGTTCTGTGAAATAAGG	
pCsRNA0160F2	GGTGGTGGATCCGGTTATGAAAAGCCGCTAG	
pCsRNA0160R2	GGTGGTGAATTCCTATTCCTCCTTTAAATCTTCGTAC	
F1sRNA0280	GTAATAATCAATAAAGGCGACATC	5'3'RACE <i>ern0280</i>
F2sRNA0280	GACATCAACCTGAAAAAAGCCC	
R1sRNA0280	GTTTCGGTTATTCAACCAATAAC	
R2sRNA0280	CTGAATAAAGAGACACAACTCC	
sRNA1260F2	GAATCGTCAGATAGCCTCCTC	5'3'RACE <i>ern1260</i>
sARN1260F3	GAGTTTTTCGAGAAGGCTATCG	
sRNA1260R1	GGGTGACTATTACTAACTATTAC	
sRNA1260R2	TAAACGATGTATTCCTAACCTC	
sRNA1300F1	GTTTCCAAATTGTATGTATGCAG	5'3'RACE <i>ern1300</i>
sRNA1300F2	GTATGCAGACCAAAAGCCACATTG	
sRNA1300R1	GTTAAATAACTTATATCTCACCG	
sRNA1300R2	CGTAACTAATGAATGCTCAATATCG	
sRNA2210F1	GGCAGGAGACCACCGTTACC	5'3'RACE <i>ern2210</i>
sRNA2210F2	GTTACCAAATAGATAGGCCCG	
sRNA2210R1	CACAAGTCACACTACTCATA	
sRNA2410F1	GCTAAGTGGCTGAGCAGTAAAATAC	5'3'RACE <i>ern2410</i>
sRNA2410R1	GGCTTTTGTCTGCAATACAAAGTTC	
sRNA2410R2	AACTGTCGTTTAACGTTTGTATCC	

Tableau supplémentaire 2. Liste des génomes chromosomiques utilisés selon les différents ARN pour les prédictions bio-informatiques par CopraRNA.

Les souches ont été sélectionnées si une homologie de 100 % entre l'ARN et le génome chromosomique est retrouvée et que la référence soit disponible sur la base de données de CopraRNA. Bien que théoriquement disponible, la souche Aus0004 n'a pas pu être incluse à cause d'une anomalie technique du logiciel.

Souche	ncbi ref	ST	R (ResFinder 4.1)	0030	0120	0160	1260	2020	2050	2410
DO/TX16	NC_017960	18	<i>msr(C), aac(6')-li, pbp5</i>		✓	✓	✓			
Aus0085	NC_021994	203	<i>msr(C), parC, gyrA, pbp5, aac(6')-li, tet(M), vanB</i>		✓	✓				
VRE001	NZ_CP018071	203	<i>parC, gyrA, aac(6')-li, pbp5, msr(C), tet(M), vanB</i>	✓				✓		
V1836	NZ_CP044264	787	<i>parC, gyrA, aac(6')-li, pbp5, msr(C), dfrG, tet(M)</i>	✓	✓		✓		✓	✓
KUHS13	NZ_AP022341	17	<i>msr(C), aac(6')-li, pbp5, dfrG, tet(M), tet(L), parC, gyrA</i>	✓	✓	✓		✓	✓	✓
E4402	NZ_LR135174	78	<i>msr(C), aac(6')-li, pbp5, tet(M)</i>	✓	✓		✓	✓		
13-022	NZ_CP025392	17	<i>gyrA, parC, erm(T), msr(C), aac(6')-li, aph(2'')-la, aac(6')-aph(2''), dfrG, tet(M), tet(L), pbp5, vanB</i>	✓				✓		
VRE1/HOU503	NZ_CP040706	280	<i>msr(C), parC, gyrA, pbp5, aac(6')-li, dfrG, tet(M), tet(L)</i>		✓	✓	✓		✓	✓
QU50	NZ_AP019394	311/507	<i>msr(C), pbp5, gyrA, parC, aac(6')-li</i>		✓	✓				
ME3	NZ_CP043865	80	<i>gyrA, parC, aac(6')-li, pbp5, msr(C), erm(T), dfrG, tet(M), tet(L)</i>		✓	✓	✓	✓	✓	✓
515	NZ_CP041257	549	<i>parC, gyrA, msr(C), aac(6')-li, pbp5</i>		✓	✓			✓	✓
NM213	NZ_CP034949	1054	<i>msr(C), aac(6')-li, pbp5, gyrA</i>		✓					
V24	NZ_CP036151	80	<i>msr(C), aac(6')-li, pbp5, dfrG, gyrA, parC</i>			✓		✓		
VRE	NZ_CP046077	?	<i>parC, gyrA, aac(6')-li, pbp5, dfrG, msr(C)</i>			✓		✓		
FB-1	NZ_CP040907	19	<i>msr(C), pbp5, gyrA, aac(6')-li</i>			✓				

# Cibles prédites pour *ern0120* et *ern0160*

Liste des cibles prédites par Target RNA2 d'*ern0120* (p-value <0,05).

Rang	P-value	COG	Code	Gène	Annotation
1	0.000	L	COG1039	rnhC	ribonuclease HIII
2	0.000	L	COG2826	tra8	Integrase
3	0.001	-	-	-	hypothetical protein
4	0.001	S	COG4483	yqgQ	hypothetical protein
5	0.001	M	COG0668	mscS	transporter, small conductance mechanosensitive ion channel family protein
6	0.001	R	COG1380	lrgA	LrgA family protein
7	0.002	G	COG0800	eda	keto-hydroxyglutarate-aldolase/keto-deoxy- phosphogluconate aldolase
8	0.002	K	COG1813	aMBF1	helix-turn-helix domain-containing protein
9	0.002	H	COG1893	panE	2-dehydropantoate 2-reductase
10	0.003	MG	COG0451	wcaG	NAD dependent epimerase/dehydratase family protein
11	0.004	S	COG3760	proX	YbaK/prolyl-tRNA synthetase-associated domain protein
12	0.005	L	COG3328	IS285	ISEf1, transposase
13	0.005	TK	COG0745	ompR	DNA-binding response regulator
14	0.005	C	COG0022	acoB	transketolase
15	0.006	E	COG0253	dapF	diaminopimelate epimerase
16	0.006	H	COG1057	nadD	nicotinic acid mononucleotide adenyltransferase
17	0.009	F	COG1001	ade	adenine deaminase
18	0.009	I	COG0416	plsX	fatty acid/phospholipid synthesis protein PlsX
19	0.010	NU	COG2804	pulE	type II/IV secretion system protein
20	0.010	L	COG0582	xerC	site-specific tyrosine recombinase XerC-family
21	0.011	V	COG1131	ccmA	ABC transporter ATP-binding protein
22	0.011	K	COG1595	rpoE	RNA polymerase sigma-70 factor, region 4
23	0.011	-	-	-	hypothetical protein
24	0.011	KE	COG1167	mocR	GntR family transcriptional regulator
25	0.011	R	COG3973	helD	DNA helicase IV
26	0.012	M	COG1083	neuA	cytidyltransferase
27	0.013	V	COG1131	ccmA	bacitracin transport ATP-binding protein BcrA
28	0.014	F	COG0519	guaA2	GMP synthase
29	0.014	M	COG2148	wcaJ	sugar transferase
30	0.014	M	COG4932	ydcQ	von Willebrand factor type A domain-containing protein (EbpA/EmpA)
31	0.015	G	COG1593	dctQ	TRAP dicarboxylate transporter subunit DctM
32	0.015	S	COG4652	-	hypothetical protein
33	0.016	-	-	-	hypothetical protein
34	0.017	G	COG2723	bglB	6-phospho-beta-galactosidase
35	0.018	J	COG0594	rnpA	Ribonuclease P protein component

Rang	P-value	COG	Code	Gène	Annotation
36	0.020	-	-	-	hypothetical protein
37	0.021	E	COG2423	OCDMu	ornithine cyclodeaminase/mu-crystallin family protein
38	0.023	KT	COG3835	cdaR	Sugar diacid utilization regulator
39	0.024	S	COG4863	yycH	hypothetical protein
40	0.025	D	COG2919	ftsB	septum formation initiator
41	0.025	-	-	-	hypothetical protein
42	0.025	T	COG1956	GAF	GAF domain-containing protein
43	0.025	R	COG0446	fadH2	pyridine nucleotide-disulfide oxidoreductase
44	0.026	E	COG0031	cysK	cysteine synthase A
45	0.026	G	COG4724	ypcD	mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase
46	0.026	-	-	-	hypothetical protein
47	0.028	P	COG2239	mgtE	magnesium transporter
48	0.030	R	COG3641	pfoR	membrane protein
49	0.030	K	COG1476	-	helix-turn-helix domain-containing protein
50	0.032	S	COG1284	yitT	hypothetical protein
51	0.034	G	COG0153	galK	galactokinase
52	0.035	-	-	-	M protein trans-acting positive regulator
53	0.035	R	COG2244	rfbX	Polysaccharide biosynthesis protein
54	0.036	S	COG1806	ppsR	Putative phosphotransferase
55	0.037	L	COG0270	dcm	C-5 cytosine-specific DNA methylase
56	0.038	G	COG2814	araJ	MFS transporter, sugar porter family protein
57	0.038	I	COG0671	pgpB	undecaprenyl-diphosphatase
58	0.039	R	COG0693	thiJ	DJ-1/PfpI family protein
59	0.042	V	COG1680	ampC	Beta-lactamase
60	0.042	F	COG0461	pyrE	orotate phosphoribosyltransferase
61	0.045	H	COG0262	folA	Dihydrofolate reductase
62	0.046	L	COG1195	recF	recombination protein F, RecF
63	0.046	S	COG4856	ybbR	hypothetical protein
64	0.047	K	COG3617	BRO	anti-repressor protein
65	0.047	J	COG0092	rpsC	30S ribosomal protein S3
66	0.047	-	-	-	DNA primase/polymerase
67	0.047	P	COG0803	znuA	periplasmic solute binding protein
68	0.048	-	-	-	hypothetical protein
69	0.048	V	COG1132	mdlB	ABC transporter ATP-binding protein
70	0.049	R	COG0110	wbbJ	maltose O-acetyltransferase



Liste des cibles prédites par copraRNA d'*ern0120* (p-value <0,05).

Rang	FDR	P-value	Gène	Annotation
1	0,0003	1,28E-07	-	
2	0,0212	1,60E-05	nrdG	anaerobic ribonucleotide reductase small subunit
3	0,0396	4,48E-05	glpF	glycerol MIP family major intrinsic protein channel protein
4	0,1069	0,000161	-	
5	0,2357	0,000445	-	redox-sensing transcriptional regulator
6	0,4689	0,001063	gloA3	lactoylglutathione lyase
7	0,5495	0,001652	xfpA	D-xylulose 5-phosphate/D-fructose 6-phosphate phosphoketolase
8	0,5495	0,001694	-	MerR family transcriptional regulator
9	0,5495	0,002171	fucO	lactaldehyde reductase
10	0,5495	0,002206	trmB	tRNA (guanine-N(7))-methyltransferase
11	0,5495	0,002372	ispE	4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase
12	0,5495	0,00249	ppnK	NAD(+) kinase
13	0,5523	0,002834	-	lysozyme
14	0,5523	0,003867	metG2	methionine--tRNA ligase
15	0,5523	0,003938	aroC	chorismate synthase
16	0,5523	0,004151	-	
17	0,5523	0,004304	pcaC	4-carboxymuconolactone decarboxylase
18	0,5523	0,004334	sdhB	L-serine ammonia-lyase subunit beta
19	0,5523	0,004354	-	P-ATPase superfamily P-type ATPase cation transporter
20	0,5523	0,004361	-	ABC superfamily ATP binding cassette transporter ABC protein
21	0,5523	0,00438	-	
22	0,5583	0,004638	prgN	replication control protein PrgN
23	0,5737	0,005393	coaA	pantothenate kinase
24	0,5737	0,005478	-	glycosyl hydrolase
25	0,5737	0,005525	-	
26	0,5737	0,005633	-	
27	0,5750	0,005863	ziaA	P-ATPase superfamily P-type ATPase cation transporter
28	0,6043	0,006552	rnhC	ribonuclease HIII
29	0,6043	0,006689	-	alpha-L-rhamnosidase
30	0,6043	0,006894	scpA	segregation and condensation protein A
31	0,6043	0,007074	aspC	aspartate transaminase
32	0,6360	0,007686	amiC	N-acetylmuramoyl-L-alanine amidase
33	0,6452	0,00843	-	
34	0,6452	0,008472	rfbP	undecaprenyl-phosphate galactosephosphotransferase
35	0,6452	0,008911	-	transcriptional regulator
36	0,6452	0,009014	-	ATP-dependent DNA helicase
37	0,6452	0,009016	tufA2	elongation factor EF1A
38	0,6566	0,009422	fliG	flagellar motor switch protein
39	0,6709	0,009881	-	PTS system trehalose transporter subunit IIABC
40	0,6894	0,010647	mgtE	MgtE family magnesium (Mg2+)/cobalt (Co2+) transporter-E
41	0,6894	0,010991	ypcH	carbohydrate ABC superfamily ATP binding cassette transporter membrane protein
42	0,6894	0,011319	dltD	D-alanine transfer protein DltD
43	0,6894	0,011481	-	bacteriocin-associated integral membrane protein
44	0,6894	0,011481	-	bacteriocin-associated integral membrane protein
45	0,6894	0,011845	srtA2	sortase SrtA2
46	0,6894	0,011977	-	IS204/IS1001/IS1096/IS1165 family transposase
47	0,6898	0,012244	recJ	recombination protein J
48	0,6922	0,012631	-	amino acid ABC superfamily ATP binding cassette transporter membrane protein
49	0,6922	0,012808	-	TetR/AcrR family transcriptional regulator
50	0,7215	0,014487	ilvE	branched-chain-amino-acid transaminase
51	0,7215	0,014523	-	sucrose phosphorylase
52	0,7215	0,014561	lysA	diaminopimelate decarboxylase
53	0,7215	0,014999	panE2	2-dehydropanoate 2-reductase
54	0,7215	0,015146	-	
55	0,7215	0,015293	-	pectinesterase

Rang	FDR	P-value	Gène	Annotation
56	0,7215	0,015518	-	sugar isomerase
57	0,7215	0,015531	-	IS204/IS1001/IS1096/IS1165 family transposase
58	0,7273	0,016862	pbp1B	penicillin-binding protein 1b
59	0,7273	0,017115	ppiB	peptidyl-prolyl cis-trans isomerase
60	0,7273	0,01721	-	
61	0,7273	0,017342	-	bacteriophage protein
62	0,7273	0,017596	rpiR	RpiR family transcriptional regulator
63	0,7273	0,017683	-	cell wall surface anchor family protein
64	0,7273	0,017833	-	ABC superfamily ATP binding cassette transporter membrane protein
65	0,7273	0,018058	-	phosphoketolase
66	0,7273	0,018128	lacE	PTS system lactose-N N'-diacetylchitobiose-beta-glucoside transporter subunit IIBC
67	0,7316	0,018776	rcfA	CRP/FNR family transcriptional regulator
68	0,7316	0,0193	-	transcriptional regulator
69	0,7316	0,020143	-	sugar ABC transporter ATP-binding protein
70	0,7316	0,020332	-	
71	0,7316	0,020452	-	LysM family protein
72	0,7316	0,020453	purH	bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase
73	0,7316	0,020662	uxuB	fructuronate reductase
74	0,7316	0,02114	sppA	S49 family peptidase
75	0,7316	0,021384	-	IS204/IS1001/IS1096/IS1165 family transposase
76	0,7316	0,021422	-	
77	0,7316	0,022109	-	ABC superfamily ATP binding cassette transporter ABC protein
78	0,7316	0,02245	-	
79	0,7316	0,022602	fms14	LPXTG family cell surface protein Fms14
80	0,7316	0,022743	-	
81	0,7316	0,022778	gutN	PTS system glucitol/sorbitol transporter subunit IIBC
82	0,7316	0,02279	-	
83	0,7316	0,023074	-	plasmid replication protein
84	0,7316	0,02423	-	
85	0,7316	0,02471	-	M20/M25/M40 family peptidase
86	0,7316	0,024815	rfbP2	undecaprenyl-phosphate galactose phosphotransferase
87	0,7316	0,024882	wecD	GNAT family acetyltransferase
88	0,7316	0,025177	acpP2	acyl carrier protein
89	0,7316	0,025224	-	GNAT family acetyltransferase
90	0,7316	0,025276	agrR3	arginine repressor
91	0,7316	0,026697	-	
92	0,7316	0,026757	-	BglG family transcriptional antiterminator
93	0,7316	0,027283	-	bacteriophage protein
94	0,7316	0,027455	-	
95	0,7316	0,027653	-	addiction module toxin
96	0,7316	0,027656	-	ABC superfamily ATP binding cassette transporter ABC protein
97	0,7316	0,02774	-	M3B subfamily peptidase
98	0,7316	0,027904	fbp	fructose-bisphosphatase
99	0,7316	0,028194	folA	dihydrofolate reductase
100	0,7316	0,028355	dppD	oligopeptide ABC superfamily ATP binding cassette transporter ABC protein
101	0,7316	0,028485	-	transcriptional regulator
102	0,7316	0,028565	-	
103	0,7316	0,028609	-	
104	0,7316	0,029671	-	A24A family peptidase
105	0,7316	0,029885	-	
106	0,7316	0,030005	-	ABC superfamily ATP binding cassette transporter ABC protein
107	0,7316	0,030601	alsD	acetolactate decarboxylase
108	0,7316	0,030945	xpaC	5-bromo-4-chloroindolyl phosphate hydrolysis protein
109	0,7316	0,031701	msrB	peptide-methionine (R)-S-oxide reductase
110	0,7316	0,031841	-	topoisomerase-primase (TOPRIM) domain-containing protein
111	0,7316	0,031949	-	
112	0,7316	0,032344	rpmA	50S ribosomal protein L27
113	0,7316	0,03266	-	M20A family peptidase
114	0,7316	0,032743	-	GNAT family acetyltransferase
115	0,7316	0,033918	-	PTS system mannose/fructose/sorbose transporter subunit IIA

Rang	FDR	P-value	Gène	Annotation
116	0,7316	0,03403	-	oxidoreductase
117	0,7316	0,034204	-	cof family protein
118	0,7316	0,034512	purR3	LacI family transcriptional regulator
119	0,7316	0,03469	sgaB	PTS system transporter subunit IIB
120	0,7316	0,0351	-	IS256 family transposase
121	0,7316	0,035278	ascB	6-phospho-beta-glucosidase
122	0,7316	0,03545	panB	3-methyl-2-oxobutanoate hydroxymethyltransferase
123	0,7316	0,035556	-	PTS system mannose/fructose/sorbose transporter subunit IID
124	0,7316	0,035561	murA	UDP-N-acetylglucosamine 1-carboxyvinyltransferase
125	0,7316	0,035731	-	transposase
126	0,7316	0,036524	-	
127	0,7316	0,036547	malA2	maltodextrose utilization protein MalA
128	0,7316	0,037085	-	lipoate--protein ligase A
129	0,7316	0,037248	uppP	undecaprenyl-diphosphatase
130	0,7316	0,037285	-	
131	0,7316	0,037294	-	PadR family transcriptional regulator
132	0,7316	0,037652	fhuB	iron (Fe3+) ABC superfamily ATP binding cassette transporter membrane protein
133	0,7316	0,038005	-	
134	0,7316	0,038284	comGA	competence protein ComGA
135	0,7316	0,038427	-	
136	0,7316	0,038524	proY	APC family amino acid transporter proline permease
137	0,7316	0,038999	-	
138	0,7316	0,039187	-	GNAT family acetyltransferase
139	0,7316	0,039264	-	VanZ/RDD domain-containing protein
140	0,7316	0,040099	terL	bacteriophage terminase large subunit
141	0,7316	0,040179	-	
142	0,7316	0,040286	abfA2	alpha-N-arabinofuranosidase
143	0,7316	0,040762	-	
144	0,7316	0,040771	pyrP	NCS2 family uracil:cation symporter-2
145	0,7316	0,04089	-	
146	0,7316	0,04102	dexB	glucan 1 6-alpha-glucosidase
147	0,7316	0,04119	-	IS4 family transposase
148	0,7316	0,04119	-	IS4 family transposase
149	0,7316	0,04211	gatB	aspartyl/glutamyl-tRNA amidotransferase subunit B
150	0,7316	0,043047	glpQ	glycerophosphodiester phosphodiesterase
151	0,7316	0,043061	lepB	S26A subfamily signal peptidase I
152	0,7316	0,043113	-	conjugative transposon protein
153	0,7316	0,043466	-	PTS system lactose/cellobiose transporter enzyme I
154	0,7316	0,04348	mutM	DNA-formamidopyrimidine glycosylase
155	0,7316	0,044146	-	LysR family transcriptional regulator
156	0,7316	0,044248	-	
157	0,7316	0,044502	glnQ2	glutamate ABC superfamily ATP binding cassette transporter ABC protein
158	0,7316	0,044716	-	cystathionine beta-synthase (CBS) domain-containing protein
159	0,7316	0,044852	ffh	Sec family Type I general secretory pathway protein signal recognition particle protein
160	0,7316	0,044944	-	
161	0,7316	0,045063	ackA	acetate kinase
162	0,7316	0,045211	arsC3	ArsC family transcriptional regulator
163	0,7316	0,045296	-	mannitol dehydrogenase
164	0,7316	0,04531	-	
165	0,7346	0,045774	map	M24 family methionyl aminopeptidase
166	0,7435	0,046609	-	AAA-superfamily ATP binding protein
167	0,7438	0,047352	-	transposase IS3/IS911
168	0,7438	0,047367	-	GntR family transcriptional regulator
169	0,7438	0,047619	mdIA	multidrug ABC superfamily ATP binding cassette transporter ABC/membrane protein
170	0,7438	0,047754	recO	recombination protein RecO
171	0,7452	0,048392	-	resolvase
172	0,7452	0,048656	recU	recombination protein RecU
173	0,7452	0,048684	tetL	MFS family major facilitator transporter tetracycline:cation symporter TetL
174	0,7486	0,049191	-	
175	0,7490	0,049499	-	class C beta-lactamase

Liste des cibles prédites par Target RNA2 d'*ern0160* (p-value <0,05).

Rang	P-value	COG	Code	Gène	Annotation
1	0.000	S	COG4129	ygaE	membrane protein
2	0.002	R	COG4552	eis	GNAT family acetyltransferase
3	0.004	G	COG3345	galA	alpha-galactosidase
4	0.004	S	COG5523	ybfA	integral membrane protein
5	0.004	L	COG2826	tra8	Integrase
6	0.005	S	COG0344	plsY	acyl-phosphate glycerol 3-phosphate acyltransferase
7	0.006	V	COG1136	lolD	ABC transporter ATP-binding protein
8	0.006	P	COG0025	nhaP	Na <sup>+</sup> /H <sup>+</sup> antiporter
9	0.006	-	-	-	hypothetical protein
10	0.007	M	COG3764	srtA	sortase family protein
11	0.009	K	COG0195	nusA	transcription termination factor NusA
12	0.009	G	COG4668	mtlA2	PTS system mannitol-specific transporter subunit IIA
13	0.010	L	COG0692	ung	uracil-DNA glycosylase
14	0.010	R	COG0628	perM	hypothetical protein
15	0.011	PH	COG1120	fepC	ferrichrome transport ATP-binding protein FhuC
16	0.012	G	COG0588	gpmA	phosphoglycerate mutase
17	0.012	-	-	-	hypothetical protein
18	0.012	-	-	-	hypothetical protein
19	0.013	R	COG4821	ywjG	hypothetical protein
20	0.013	-	-	-	hypothetical protein
21	0.014	-	-	-	hypothetical protein
22	0.015	-	-	-	hypothetical protein
23	0.016	L	COG4974	xerD	site-specific tyrosine recombinase XerC-family
24	0.016	M	COG0472	rfe	glycoside hydrolase family protein
25	0.017	F	COG1428	dck	deoxynucleoside kinase
26	0.018	P	COG2217	zntA	cadmium-translocating P-type ATPase
27	0.018	L	COG2801	tnpA	ISEfa8, transposase
28	0.018	-	-	-	hypothetical protein
29	0.019	R	COG0491	gloB	Metallo-beta-lactamase superfamily protein
30	0.020	V	COG1131	ccmA	bacitracin transport ATP-binding protein BcrA
31	0.021	M	COG1686	dacC	D-alanyl-D-alanine carboxypeptidase
32	0.022	S	COG2855	yeiH	membrane protein
33	0.022	G	COG1455	celB	PTS system lactose/cellobiose-specific transporter subunit IIC
34	0.025	-	-	-	hypothetical protein
35	0.025	K	COG0583	lysR	LysR family transcriptional regulator
36	0.026	S	COG4684	-	integral membrane protein
37	0.027	T	COG4753	yesN	AraC family transcriptional regulator
38	0.028	-	-	-	hypothetical protein
39	0.031	L	COG3935	dnaD	DnaD and phage-associated domain protein
40	0.031	-	-	-	polysaccharide deacetylase

Rang	P-value	COG	Code	Gène	Annotation
41	0.031	-	-	-	hypothetical protein
42	0.032	V	COG1136	lolD	ABC transporter ATP-binding protein/permease
43	0.032	C	COG1038	pycA	pyruvate carboxylase
44	0.032	C	COG0281	sfcA	malic enzyme, NAD binding domain protein
45	0.033	J	COG0828	rpsU	30S ribosomal protein S21
46	0.033	-	-	-	cell wall surface anchor family protein
47	0.035	S	COG4939	tpp15	FMN-binding domain-containing protein
48	0.035	S	COG5617	-	integral membrane protein
49	0.036	J	COG0093	rplN	50S ribosomal protein L14
50	0.037	G	COG1263	ptsG1	PTS system beta-glucoside-specific transporter subunit IIABC
51	0.038	R	COG2153	elaA	GNAT family acetyltransferase
52	0.038	-	-	-	hypothetical protein
53	0.039	P	COG1393	arsC	transcriptional regulator Spx
54	0.040	G	COG3534	abfA	alpha-N-arabinofuranosidase
55	0.041	C	COG0022	acoB	transketolase
56	0.041	J	COG0097	rplF	50S ribosomal protein L6
57	0.042	M	COG1732	osmF	glycine betaine/carnitine/choline ABC transporter
58	0.042	L	COG0350	adaB	methylated-DNA--protein-cysteine methyltransferase
59	0.043	-	-	-	hypothetical protein
60	0.044	G	COG0406	phoE	phosphoglycerate mutase family protein
61	0.047	S	COG3589	-	alpha3-beta1 integrin-binding family protein
62	0.047	U	COG0681	lepB	signal peptidase I
63	0.047	J	COG1098	yabR	S1 RNA binding domain-containing protein
64	0.050	R	COG1853	rutF	hypothetical protein

Liste des cibles prédites par copraRNA d'*ern0160* (p-value <0,05).

Rang	FDR	P-value	Gène	Annotation
1	0	0	-	P-ATPase superfamily P-type ATPase cation transporter
2	0,096	7,2E-05	-	lipoate--protein ligase A
3	0,266	0,00045	padR	PadR family transcriptional regulator
4	0,266	0,0005	rluD3	pseudouridylate synthase RluD
5	0,266	0,00051	srtC3	sortase
6	0,266	0,0007	-	LysM family protein
7	0,266	0,0007	-	LysM domain surface protein
8	0,425	0,00129	-	
9	0,425	0,00143	-	hydrolase
10	0,447	0,00168	recU	recombination protein RecU
11	0,453	0,00187	-	ATP-dependent Zn protease
12	0,507	0,00244	-	trehalose/maltose hydrolase
13	0,507	0,00247	-	
14	0,508	0,00273	-	MarR family transcriptional regulator
15	0,508	0,00286	arsC3	ArsC family transcriptional regulator
16	0,529	0,00317	-	PTS system transporter subunit IIA
17	0,607	0,00408	-	PTS system glycerone kinase transporter subunit IIA
18	0,607	0,0041	-	transcriptional regulator
19	0,617	0,00489	brnQ2	LIVCS family branched chain amino acid:cation symporter
20	0,617	0,00506	-	metal-dependent hydrolase
21	0,617	0,00534	-	family I glycosyl hydrolase
22	0,617	0,00548	-	LysR family transcriptional regulator
23	0,617	0,00561	-	GNAT family acetyltransferase
24	0,617	0,00578	-	chitin binding protein
25	0,617	0,00613	atpE	H(+)-transporting ATPase F(0) C subunit
26	0,617	0,0063	birA	bifunctional biotin--[acetyl-CoA-carboxylase] ligase/biotin operon transcriptional regulator
27	0,617	0,00638	rliB	GntR family transcriptional regulator
28	0,617	0,00706	-	
29	0,617	0,0071	clpX	S14 family endopeptidase ClpX
30	0,617	0,00719	-	intracellular protease/amidase
31	0,617	0,00723	-	ATP-binding protein
32	0,617	0,0074	-	
33	0,633	0,00792	dus	tRNA-dihydrouridine synthase
34	0,633	0,00807	lacF2	lactose ABC superfamily ATP binding cassette transporter membrane protein
35	0,653	0,00858	-	
36	0,67	0,00929	-	
37	0,67	0,0093	opuAA	glycine/betaine/carnitine/choline ABC superfamily ATP binding cassette transporter ABC protein
38	0,67	0,00955	trpS	tryptophan--tRNA ligase
39	0,68	0,00994	crr	PTS system glucose/glucoside transporter subunit IIA
40	0,695	0,01047	-	redox-sensing transcriptional regulator
41	0,695	0,01083	-	flavoprotein
42	0,695	0,01171	-	GNAT family acetyltransferase
43	0,695	0,01178	-	ABC superfamily ATP binding cassette transporter membrane protein
44	0,695	0,01182	-	AraC family transcriptional regulator
45	0,695	0,01191	-	
46	0,695	0,01216	-	integral membrane protein
47	0,695	0,01301	lipA	lipase
48	0,695	0,01308	tig	cell division trigger factor
49	0,695	0,01314	secG	Sec family Type I general secretory protein SecG
50	0,695	0,0132	dltD	D-alanine transfer protein DltD
51	0,695	0,01329	-	phosphosugar isomerase
52	0,723	0,01409	gdh	glucose-1-dehydrogenase
53	0,727	0,01501	-	MerR family transcriptional regulator
54	0,727	0,01503	-	response regulator
55	0,727	0,01557	nrdI	ribonucleotide reductase
56	0,727	0,0161	-	family I glycosyl hydrolase

Rang	FDR	P-value	Gène	Annotation
57	0,727	0,01636	-	ABC superfamily ATP binding cassette transporter membrane protein
58	0,727	0,01703	srlB	PTS system glucitol/sorbitol transporter subunit IIA
59	0,727	0,01706	rpoN	DNA-directed RNA polymerase sigma subunit RpoN
60	0,727	0,01734	pad	phenolic acid decarboxylase
61	0,727	0,01745	trkG	Trk family potassium (K+) transporter membrane protein
62	0,727	0,01759	dnaA	DNA-directed DNA replication initiator protein
63	0,727	0,01777	lysA	diaminopimelate decarboxylase
64	0,727	0,01785	-	Sir2 family transcriptional regulator
65	0,727	0,01819	-	bacteriocin-associated integral membrane protein
66	0,727	0,01819	-	bacteriocin-associated integral membrane protein
67	0,727	0,01857	-	
68	0,727	0,01862	sgbH	3-dehydro-L-gulonate-6-phosphate decarboxylase
69	0,727	0,01901	-	glycosyltransferase
70	0,727	0,01965	-	sugar ABC transporter ATP-binding protein
71	0,727	0,01993	-	ABC superfamily ATP binding cassette transporter membrane protein
72	0,727	0,02084	-	
73	0,727	0,02107	-	LysR family transcriptional regulator
74	0,727	0,02107	pepA	M42 family glutamyl aminopeptidase
75	0,727	0,02111	-	alkaline phosphatase
76	0,727	0,02124	-	SGNH subfamily hydrolase
77	0,727	0,02129	rpiA	ribose-5-phosphate isomerase
78	0,727	0,02157	pbp2B	penicillin-binding protein
79	0,727	0,02176	-	conjugative transposon protein
80	0,727	0,02179	lpsB	hexapeptide transferase
81	0,732	0,02225	-	
82	0,732	0,02412	-	major facilitator superfamily protein
83	0,732	0,02419	-	
84	0,732	0,02453	recJ	recombination protein J
85	0,732	0,02458	-	A24A family peptidase
86	0,732	0,02465	-	AbrB family transcriptional regulator
87	0,732	0,02516	-	
88	0,732	0,02541	-	ABC superfamily ATP binding cassette transporter ABC protein
89	0,732	0,02546	-	exonuclease
90	0,732	0,02591	-	
91	0,732	0,02621	-	
92	0,732	0,02681	pbp5	penicillin binding protein 2 prime
93	0,732	0,02715	pheA	prephenate dehydratase
94	0,732	0,02715	asp	alkaline shock protein
95	0,732	0,02777	-	
96	0,732	0,02788	-	rhodanese domain sulfurtransferase
97	0,732	0,02831	glpK	glycerol kinase
98	0,732	0,02838	acyP	acylphosphatase
99	0,732	0,02856	relA	GTP diphosphokinase
100	0,732	0,02883	-	
101	0,732	0,02884	gpmA2	phosphoglycerate mutase
102	0,732	0,02921	-	DNA-damage-inducible protein
103	0,732	0,02922	fms11	LPXTG family cell surface protein Fms11
104	0,732	0,02922	fms11	LPXTG family cell surface protein Fms11
105	0,732	0,02922	fms11	LPXTG family cell surface protein Fms11
106	0,732	0,0293	abi	abortive infection protein
107	0,732	0,02937	abfA	alpha-L-arabinofuranosidase
108	0,737	0,03077	-	transcriptional regulator
109	0,737	0,03087	-	nicotinamidase/isochorismatase
110	0,737	0,03131	-	
111	0,737	0,03171	yugI	general stress protein 13
112	0,737	0,03213	-	alpha/beta hydrolase
113	0,737	0,0322	gapA	glyceraldehyde-3-phosphate dehydrogenase
114	0,737	0,03269	-	
115	0,737	0,03369	gadC2	APC family amino acid-polyamine-organocation transporter
116	0,737	0,03369	gadC2	APC family amino acid-polyamine-organocation transporter
117	0,737	0,03376	msrA4	peptide-methionine-(S)-S-oxide reductase

Rang	FDR	P-value	Gène	Annotation
118	0,737	0,03418	cyaA	adenylate cyclase
119	0,737	0,03439	gntR	RpiR family transcriptional regulator
120	0,737	0,03441	-	DeoR family transcriptional regulator
121	0,737	0,03598	-	RNA-binding heat shock protein
122	0,737	0,03685	tagH	teichoic acid ABC superfamily ATP binding cassette transporter ABC protein
123	0,737	0,03694	fms14	LPXTG family cell surface protein Fms14
124	0,737	0,03699	-	
125	0,737	0,03735	atpB	H(+)-transporting ATPase F(0) A subunit
126	0,737	0,03746	pyrK	dihydroorotate oxidase electron transfer subunit
127	0,737	0,03755	-	permease
128	0,737	0,0377	-	
129	0,737	0,03772	ypdA	ABC superfamily ATP binding cassette transporter permease protein
130	0,737	0,03848	-	HAD-superfamily hydrolase
131	0,737	0,0385	-	zinc ABC superfamily ATP binding cassette transporter ABC protein
132	0,737	0,03858	ardA	antirestriction protein ArdA
133	0,737	0,0388	rhaU	L-rhamnose 1-epimerase
134	0,737	0,03926	comYD	competence protein ComYD
135	0,737	0,03935	proB	glutamate 5-kinase
136	0,737	0,03946	-	lactoylglutathione lyase
137	0,737	0,03969	-	TetR/AcrR family transcriptional regulator
138	0,737	0,03973	-	
139	0,737	0,03996	sacT	BglG family transcriptional antiterminator
140	0,737	0,04037	-	ABC superfamily ATP binding cassette transporter ABC protein
141	0,737	0,04045	bglG	transcriptional antiterminator
142	0,737	0,04081	-	deoxynucleoside kinase
143	0,737	0,04098	-	
144	0,737	0,04156	-	PTS system lactose-N N'-diacetylchitobiose-beta-glucoside transporter subunit IIA
145	0,737	0,04161	-	phosphatidic acid phosphatase
146	0,737	0,04163	folD	bifunctional methylenetetrahydrofolate dehydrogenase (NADP(+))/methenyltetrahydrofolate cyclohydrolase
147	0,737	0,04177	-	integrase
148	0,737	0,04195	sfcA	malate dehydrogenase (oxaloacetate-decarboxylating) (NADP(+))
149	0,737	0,04233	scrA	PTS system sucrose transporter subunit IIBC
150	0,737	0,04287	-	integral membrane protein
151	0,737	0,04313	pcnB	polynucleotide adenyllyltransferase
152	0,737	0,04327	-	lysyl-tRNA synthetase
153	0,737	0,0433	psr	LytR family transcriptional regulator
154	0,737	0,0433	-	
155	0,737	0,04331	-	sensor histidine kinase
156	0,737	0,04355	-	ABC superfamily ATP binding cassette transporter membrane protein
157	0,737	0,04435	azlD2	LIV-E family branched chain amino acid exporter AzlD
158	0,737	0,04444	-	membrane protein
159	0,737	0,04455	-	endonuclease/exonuclease/phosphatase family protein
160	0,737	0,04476	-	
161	0,737	0,0448	rcfA	CRP/FNR family transcriptional regulator
162	0,737	0,04499	feuC	iron (Fe3+) ABC superfamily ATP binding cassette transporter ABC protein
163	0,737	0,04518	-	MscS family small conductance mechanosensitive ion channel protein
164	0,737	0,0453	proY	APC family amino acid transporter proline permease
165	0,742	0,04588	-	CAAX amino terminal protease family protein
166	0,742	0,04619	secA	Sec family Type I general secretory pathway preprotein translocase subunit SecA
167	0,744	0,04709	-	
168	0,744	0,04725	-	membrane protein
169	0,744	0,04751	sugE	small multidrug resistance protein
170	0,744	0,04773	-	MFS family major facilitator transporter C4-dicarboxylate porter
171	0,744	0,04775	-	lipoprotein
172	0,744	0,04824	-	membrane protein
173	0,744	0,04824	-	membrane protein
174	0,745	0,04863	arcT	aminotransferase
175	0,754	0,04989	-	
176	0,754	0,05	-	ABC superfamily ATP binding cassette transporter ABC protein





**Titre :** Etude fonctionnelle de nouveaux ARN régulateurs exprimés chez *Enterococcus faecium*.

**Mots clés :** *Enterococcus faecium*, Bactériologie, Régulation, Antibio-résistance, Riboswitch

**Résumé :** *Enterococcus faecium* est un pathogène opportuniste multi-résistant aux antibiotiques, adapté à l'environnement hospitalier et responsable de nombreuses épidémies hospitalières. L'étude de la réponse au stress antibiotique a récemment permis la mise en évidence des premiers petits ARN exprimés chez *E. faecium*. Ce travail de thèse s'inscrit dans la détermination des fonctions de certains de ces ARN, notamment en lien avec l'antibio-résistance. Nous avons pu identifier ces ARN soit comme des régulateurs localisés dans la région 5' non traduite (5'UTR) du gène régulé soit comme des ARN régulateurs *bona fide*.

Une investigation fonctionnelle de deux de ces régulateurs situés en 5'UTR a été effectuée par des approches génétiques, transcriptomiques et phénotypiques.

Le premier, *ern0030*, est impliqué dans la régulation du gène de résistance aux tétracyclines *tet(M)*. Nos travaux apportent de nouvelles données expérimentales suggérant l'existence d'un mécanisme de régulation alternatif à celui d'atténuation transcriptionnelle initialement décrit.

Le second, *ern2050*, est un régulateur de type T-box riboswitch impliqué dans la régulation des deux aminoacyl-synthétases HisS et AspS. Nous avons validé expérimentalement ce premier T-box riboswitch chez *E. faecium*. Cette investigation a également permis de mettre en place un milieu chimiquement défini ayant permis la détermination des acides aminés essentiels chez l'espèce *E. faecium* et de décrire l'impact de l'absence de régulation de l'opéron *hisS-aspS* sur la croissance bactérienne dans un milieu pauvre en aspartate. Les données transcriptomiques obtenues ont permis de révéler un autre T-box riboswitch spécifique de l'ARN<sup>tHis</sup>, *ern1900*.

**Title:** Functional study of novel small RNAs expressed by *Enterococcus faecium*

**Keywords:** *Enterococcus faecium*, Bacteriology, Regulation, antibiotic resistance, Riboswitch

**Abstract:** *Enterococcus faecium* is a hospital-adapted multi-drug-resistant pathogen, responsible for numerous hospital outbreaks. Recently, the study of antibiotic stress adaptation allowed to discover the first small RNAs expressed in *E. faecium*. The aim of this thesis was to investigate the functions of some of these novel sRNAs, especially in connection with antibiotic resistance. We were able to identify these RNAs as *cis*-regulatory RNA leaders or regulatory RNAs.

A functional study of two RNA leaders was performed using a panel of genetic, transcriptomic and phenotypic approaches.

The first RNA, *ern0030*, is implicated in the regulation of *tet(M)*, gene conferring tetracycline resistance. Our works showed new experimental data suggesting an alternative regulatory mechanism different from that of transcriptional attenuation initially described.

The second RNA, *ern2050*, is a T-box riboswitch involved in the regulation of expression of two aminoacyl synthetases (HisS and AspS). We experimentally validated this first T-box riboswitch in *E. faecium*. Additionally, we developed a novel chemically-defined medium, which allowed us to determine the amino acid auxotrophy of *E. faecium*. We demonstrated that the regulation of the *hisS-aspS* operon expression was primordial for bacterial growth in aspartate-depleted medium. Global transcriptomic data also revealed another T-box riboswitch, *ern1900*, specific for tRNA<sup>His</sup>.