



**HAL**  
open science

# Identification et caractérisation d'ARN régulateurs impliqués dans la réponse au stress et la virulence chez **Enterococcus faecalis**

Marine Salze

► **To cite this version:**

Marine Salze. Identification et caractérisation d'ARN régulateurs impliqués dans la réponse au stress et la virulence chez *Enterococcus faecalis*. Santé. Normandie Université, 2019. Français. NNT : 2019NORMC403 . tel-03221792

**HAL Id: tel-03221792**

**<https://theses.hal.science/tel-03221792v1>**

Submitted on 10 May 2021

**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.



Normandie Université

## THÈSE

Pour obtenir le diplôme de doctorat

Spécialité ASPECTS MOLECULAIRES ET CELLULAIRES DE LA BIOLOGIE

Préparée au sein de l'Université de Caen Normandie

**Identification et caractérisation d'ARN régulateurs impliqués dans la réponse au stress et la virulence chez *Enterococcus faecalis***

Présentée et soutenue par  
**Marine SALZE**

**Thèse soutenue publiquement le 09/05/2019  
devant le jury composé de**

Mme HILDE DE REUSE	Directeur de recherche, Institut Pasteur	Rapporteur du jury
Mme PASCALE ROMBY	Directeur de recherche au CNRS, Université de Strasbourg	Rapporteur du jury
Mme NATHALIE CONNIL	Maître de conférences HDR, Université Rouen Normandie	Membre du jury
Mme CECILE MULLER- PUJOL	Maître de conférences, Université Caen Normandie	Membre du jury
M. HARALD PUTZER	Directeur de recherche, Institut de biologie physico-chimique	Président du jury
M. ALAIN RINCE	Professeur des universités, Université Caen Normandie	Directeur de thèse

**Thèse dirigée par ALAIN RINCE, UR risques microbiens (Caen)**



UNIVERSITÉ  
CAEN  
NORMANDIE



Normandie de Biologie Intégrative,  
Santé, Environnement



*Un scientifique dans son laboratoire est non seulement un technicien : il est aussi un enfant placé devant des phénomènes naturels qui l'impressionnent comme des contes de fées.*

- Marie Curie

## Remerciements

Je souhaite en premier lieu remercier le Docteur Hilde De Reuse et le Docteur Pascale Romby pour avoir accepté d'évaluer cette thèse et pour leur investissement en tant que rapporteurs. Je souhaiterais également remercier le Docteur Harald Putzer et le Docteur Nathalie Connil pour leur présence dans le jury et pour avoir accepté d'examiner mon travail de thèse.

Je remercie également le Professeur Axel Hartke et le Professeur Alain Rincé qui m'ont accueillie au sein du laboratoire U2RM.

Je tiens à remercier le Professeur Alain Rincé qui a dirigé ces travaux. Merci pour ta patience, pour le temps que tu as toujours réussi à me donner, et pour ta gentillesse. Et merci encore de m'avoir proposé ce sujet au moment où je m'étais résignée à ne pas faire de thèse. J'ai beaucoup appris à tes côtés, merci encore.

Je souhaite adresser également tous mes remerciements au Docteur Cécile Muller pour avoir co-encadré cette thèse. Merci d'avoir guidé et soutenu pendant ces (plus de) trois ans la boule de stress que je suis, malgré les bévues et les chutes de confiance. Merci pour tout ce que tu m'as appris. Je te souhaite le meilleur pour la suite et bonne chance pour ton HDR.

Je voudrais remercier toute l'équipe d'enseignants-chercheurs et de chercheurs du laboratoire, pour leur aide et leurs conseils en toute circonstance, mais aussi pour avoir donné le goût de la recherche à une petite étudiante de M1. Merci également à Benoit Bernay et Julien Pontin pour les manip de spectrométrie de masse. Merci également à Didier Goux pour les expériences de microscopie électronique. Merci aussi à Eliette Bisson, Nicolas Verneuil, Abdellah Benachour et Jean-Christophe Giard qui m'ont apporté leur aide précieuse pour les corrections de ce manuscrit.



Un merci tout particulièrement aux techniciennes, Isabelle et Anne-Cécile, que j'ai fait trimer à la fin de ma thèse dans la course aux derniers résultats. Merci pour votre aide indispensable.

Je souhaite tous particulièrement remercier les thésards et ATER du laboratoire, Cindy, Loïc, Maxime (merci pour l'éthanol dans le tampon A4 !), Hakim, Delphine et Valentin. Merci pour les rires, le soutien, la Door of Shame Fame, les soirées, les jeux de mots nuls, les discussions éclairées et avisées du vendredi soir après 17h.

Et surtout merci à tous pour cette atmosphère joyeuse qui a toujours régné autant que je me souviens dans ce laboratoire, et qui a été un incroyable moteur pour la réussite de cette thèse.

J'adresse une tendresse toute particulière à ma famille. Maman, Papa, merci pour le soutien, les sessions skype, les passages-éclairés. Tout cela n'avait pas de prix. Merci à mon frère Guillaume, soutien indéfectible malgré la distance, and thank you to my sister-in-law, Haruka, always with a nice word. Une tendresse pour mes magnifiques petites nièces, Lilwenn et Manon.

Merci à Julien, Fiona, Magali, Di. Et surtout Camille. Ça fait 15 ans qu'on forme notre équipe de toquées, merci à toi d'être toujours là.

Je vais terminer ces remerciements par Simon, qui a partagé ma vie et ma thèse malgré son éloignement, et continue encore. Tu es arrivé dans ma vie au milieu de mon doctorat, tu l'as chamboulée pour le mieux, merci pour tout.

Merci à tous

# Sommaire

Liste des abréviations .....	5
Liste des figures .....	7
Liste des tableaux .....	9
<i>Introduction</i> .....	<i>11</i>
I. Contexte de l'étude .....	12
II. Le genre <i>Enterococcus</i> .....	13
1. Historique .....	13
2. Caractéristiques du genre <i>Enterococcus</i> .....	14
3. <i>Enterococcus faecalis</i> .....	15
a) Environnement et applications .....	15
b) Données génomiques.....	16
c) Opportunisme et virulence .....	16
d) Métabolisme.....	18
III. La régulation transcriptionnelle et post-transcriptionnelle par les ARN régulateurs.....	21
1. Généralités sur la régulation de l'expression des gènes .....	21
2. Les petits ARN .....	23
a) Les petits ARN codés en <i>cis</i> .....	23
b) Les petits ARN codés en <i>trans</i> .....	24
c) Cas particuliers .....	26
3. Les ARN liant les protéines .....	26
4. Les CRISPR .....	28
5. Les 5'UTR (UnTranslated Region) régulatrices .....	28
a) Régulation par blocage ou déclenchement de la traduction .....	29
b) Régulation par terminaison prématurée de la transcription .....	30
6. Les 3'UTR régulatrices.....	31
IV. Les protéines impliquées dans le métabolisme des ARN .....	32
1. Les protéines chaperonnes à ARN .....	32
2. Le dégradosome .....	33
a) Les RNases .....	34

b) Les hélicases à motif DEAD.....	35
V. Les ARNs régulateurs chez <i>E. faecalis</i> .....	37
VI. Contexte et objectifs .....	41
<i>Matériel et Méthodes</i> .....	43
<i>Résultats</i> .....	53
Chapitre 1 : Identification de petits ARN putatifs induits in vivo et en conditions de stress.....	54
Publication 1 : Identification of the general stress stimulon related to colonization in <i>Enterococcus faecalis</i> .....	54
Chapitre 2 : Etude du rôle et de l'action d'un petit ARN régulateur agissant en <i>trans</i> .....	86
Chapitre 3 : Etude des mécanismes de régulation d'une 5'UTR possédant un terminateur de transcription.....	101
1. Contexte : découverte de la 5'UTR.....	101
2. Publication 2 : Characterization of a Bgl antiterminator in <i>Enterococcus faecalis</i> .....	104
3. Résultats complémentaires .....	134
Chapitre 4 : Caractérisation d'une hélicase à motif DEAD régulée par une région 5' non traduite (5'UTR).....	138
1. Publication 3 : Study of RNA metabolism actors in <i>Enterococcus faecalis</i> ...	138
2. Résultats complémentaires.....	168
<i>Discussion générale et perspectives</i> .....	171
<i>Références</i> .....	180
<b>Annexe 1</b> : Résultats complémentaires de la publication 1 « Identification of the general stress stimulon related to colonization in <i>Enterococcus faecalis</i> ».....	190
<b>Annexe 2</b> : Résultats complémentaires de la publication 2 « Characterization of a Bgl antiterminator in <i>Enterococcus faecalis</i> » .....	210
<b>Annexe 3</b> : Résultats complémentaires de la publication 3 « Study of RNA metabolism actors in <i>Enterococcus faecalis</i> ».....	211

## Liste des abréviations

aa : acides aminés

ADN : Acide Désoxyribonucléique

ARN : Acide Ribonucléique

ATP : Adénosine Triphosphate

CMB : Concentration Minimale Bactéricide

CMI : Concentration Minimale Inhibitrice

CRISPR : Clustured Regularly Interspaced Short Palindromic Repeats (Courtes Répétitions Palindromiques Groupées et Régulièrement Espacées)

DO : Densité Optique

GM17 : Milieu M17 glucosé (0,5%)

GUS :  $\beta$ -glucuronidase

iTRAQ : Isobaric Tag for Relative and Absolute Quantification (Étiquettes Isobares pour la Quantification Relative et Absolue)

kb : kilobases

Mb : Megabases

NAG : N-Acetyl-glucosamine

OMS : Organisation Mondiale de la Santé

ORF : Open Reading Frame (Cadre Ouvert de Lecture)

pb : paires de bases

PCR : Polymerase Chain Reaction (Réaction de Polymérisation en Chaîne)

RT-PCR : Rétro-Transcription puis PCR

PEP : Phosphoénolpyruvate

PTS : Phospho-Transferase System (Système Phosphotransférase)

RBS : Ribosome Binding Site (Site de Liaison du Ribosome)

RM17 : Milieu M17 supplémenté en ribose

RNase : Ribonucléase

sARN : small RNA ou petit ARN régulateur

SD : Shine-Dalgarno

UFC : Unité formant colonies

UTR : Untranslated region (Région non traduite)

VRE : Vancomycin Resistant Enterococci (Entérocoques Résistants à la Vancomycine)

## Liste des figures

**Figure 1** : Classification de l'Organisation Mondiale de la Santé des bactéries pathogènes pour lesquelles la recherche de nouveaux antibiotiques est prioritaire.

**Figure 2** : *E. faecalis*, souche V19, vue en microscopie électronique à balayage.

**Figure 3** : Les systèmes PTS chez les bactéries.

**Figure 4** : Les différents niveaux de régulation de l'expression des gènes par les ARN régulateurs.

**Figure 5** : Mode d'action des sARN codés en *cis* régulant l'expression d'un gène isolé ou d'un opéron.

**Figure 6** : Modes d'action des sARN codés en *trans*.

**Figure 7** : Les différents modes d'actions des ARN régulateurs liant les protéines.

**Figure 8** : Exemples d'action régulatrice par des 5'UTR.

**Figure 9** : Schéma explicatif de la régulation de type terminateur/anti-terminateur de l'opéron *bgl* chez *E. coli*.

**Figure 10** : Schéma des composants principaux du dégradosome de *Bacillus subtilis*.

**Figure 11** : Les différents domaines et séquences conservées des hélicases à motif DEAD.

**Figure 12** : Schéma explicatif de la régulation du système *eut*.

**Figure 13** : Conformation et associations entre les motifs complémentaires du locus RNAI/RNAII sur le plasmide pAD1.

**Figure 14** : Environnement génomique du petit ARN SRC65.

**Figure 15** : Northern blot mettant en évidence le sARN SRC65.

**Figure 16** : Alignement des sARN potentiels SRC65 et SRC90, en utilisant le programme MultAlin.

**Figure 17** : Northern blot mettant en évidence le sARN SRC90.

**Figure 18** : Environnement génomique du petit ARN SRC90.

**Figure 19** : Courbes de croissance des souches V19,  $\Delta$ SRC65, pAGEnt et SRC65+ à 37°C en RM17 supplémenté en agmatine 40 mM.

**Figure 20** : Résultats de l'analyse de l'abondance des ARNm codés par les gènes *ef0160*, *ef2221* et *ef3251* par RT-qPCR, réalisée avec les ARN des souches  $\Delta$ SRC65 et SRC65+.

**Figure 21** : RT-PCR effectuée avec des oligonucléotides localisés sur SRC85 et sur *ef1515*.

**Figure 22** : Gel retard permettant d'étudier la fixation de CcpA sur 5'1515 en présence de son co-facteur HprP

**Figure 23** : Survie de larves *Galleria mellonella* infectées par *E. faecalis* V19 ou par le mutant  $\Delta$ *ef1518*.

**Figure 24** : Résultats des dosages de la  $\beta$ -glucuronidase avec les constructions C4 et C5, transformées dans les souches V19 et  $\Delta$ *cshC*.

**Figure 25** : Observation en microscopie électronique à balayage de la souche V19 et de la souche  $\Delta$ 5'*cshA*.

**Figure 26** : Schéma récapitulatif des résultats concernant les trois ARN régulateurs étudiés.

## Liste des tableaux

**Tableau 1** : Principaux facteurs de virulence connus chez *Enterococcus faecalis*, avec leurs rôles référencés lors du processus d'infection et leur pathologie associée.

**Tableau 2** : Tableau récapitulatif des sARN caractérisés chez *E. faecalis*.

**Tableau 3** : Liste des oligonucléotides utilisés dans les parties hors articles.

**Tableau 4** : Facteurs de variation de l'expression de SRC65 et SCR90 chez la souche V19 d'*E. faecalis* soumise à des conditions de stress, et comparés aux cellules en phase exponentielle de croissance en milieu BHI non stressant.

**Tableau 5** : Cibles potentielles de SRC65 identifiées à l'aide du logiciel TargetRNA2.

**Tableau 6** : Liste des protéines détectées lors des expériences d'iTRAq avec un facteur d'induction ou de répression ( $\Delta$ SRC65/V19) d'au moins 2, et pour lesquelles le test statistique (t-test) montre une  $p$ -value inférieure à 0,05.

**Tableau 7** : Cibles potentielles de SRC90 identifiées à l'aide du logiciel TargetRNA2.



# *Introduction*

## I. Contexte de l'étude

Depuis plusieurs années, une attention croissante est portée aux micro-organismes responsables d'infections en milieu hospitalier. Celles-ci sont définies par le comité américain consultatif des pratiques pour la maîtrise des infections liées aux soins (HICPAC United States Healthcare Infection Control Practices Advisory Committee) comme toute infection contractée dans un établissement de soins, y compris les hospitalisations à domicile (Babady, 2016). Les patients immunodéprimés sont particulièrement exposés à ces maladies qui peuvent être mortelles pour ces personnes. En effet, 2 millions de cas dont près de 80 000 décès sont recensés chaque année aux États-Unis, et 750 000 cas dont 7 000 à 10 000 décès en France (Brock *et al.*, 2007).

Ainsi, sur la période de 2001 à 2017, le genre *Enterococcus* représentait 9,4% des micro-organismes les plus fréquemment retrouvés en France (Maugat *et al.*, 2018). Les infections à entérocoques sont majoritairement causées par deux représentants du genre *Enterococcus*, qui compte pourtant 47 espèces (Van Tyne et Gilmore, 2014) : *Enterococcus faecium* (environ 40% des cas aux États-Unis) et notre bactérie d'intérêt, *E. faecalis* (environ 60% des cas aux États-Unis) (Gao *et al.*, 2017). Sur la période de 2011 à 2014, 7,4% des signalements d'infections nosocomiales aux États-unis impliquaient *E. faecalis* (Weiner *et al.*, 2016).

Ces infections attirent de plus en plus l'attention des soignants et des scientifiques, notamment à cause de l'acquisition croissante de résistances aux traitements et plus particulièrement aux antibiotiques. En effet, l'Organisation Mondiale de la Santé (OMS), a dressé le 27 février 2017 une liste de bactéries pathogènes pour lesquelles il est urgent de trouver de nouveaux antibiotiques (Figure 1). Ces microorganismes sont classés en trois catégories représentatives du niveau de priorité : Critique (priorité 1), Elevée (priorité 2) et Moyenne (priorité 3) (Tacconelli *et al.*, 2018). Parmi ces organismes se trouvent les entérocoques résistants à la vancomycine (priorité 2), en majorité *E. faecium*. Ces souches sont appelées VRE, pour

« Vancomycin Resistant Enterococci ». Environ 80% des isolats cliniques d'*E. faecium* sont des VRE, contre 8 à 9% pour *E. faecalis* (Miller *et al.*, 2016).

Priorité 1 : Critique	Priorité 2 : Elevée	Priorité 3 : Moyenne
<ul style="list-style-type: none"> <li>• <b><i>Acinetobacter baumannii</i></b>, résistance aux carbapénèmes</li> <li>• <b><i>Pseudomonas aeruginosa</i></b>, résistance aux carbapénèmes</li> <li>• <b><i>Enterobacteriaceae</i></b>, résistance aux carbapénèmes, résistance aux céphalosporines</li> </ul>	<ul style="list-style-type: none"> <li>• <b><i>Enterococcus faecium</i></b>, résistance à la vancomycine</li> <li>• <b><i>Staphylococcus aureus</i></b>, résistance à la méthicilline, résistance intermédiaire ou complète à la vancomycine</li> <li>• <b><i>Helicobacter pylori</i></b>, résistance à la clarithromycine</li> <li>• <b><i>Campylobacter spp.</i></b>, résistance aux fluoroquinolones</li> <li>• <b><i>Salmonella spp.</i></b>, résistance aux fluoroquinolones</li> <li>• <b><i>Neisseria gonorrhoeae</i></b>, résistance aux céphalosporines, résistance aux fluoroquinolones</li> </ul>	<ul style="list-style-type: none"> <li>• <b><i>Streptococcus pneumoniae</i></b>, insensible à la pénicilline</li> <li>• <b><i>Haemophilus influenzae</i></b>, résistance à l'ampicilline</li> <li>• <b><i>Shigella spp.</i></b>, résistance aux fluoroquinolones</li> </ul>

Figure 1 : Classification de l'Organisation Mondiale de la Santé des bactéries pathogènes pour lesquelles la recherche de nouveaux antibiotiques est prioritaire. (Adapté de Tacconelli *et al.*, 2018).

## II. Le genre *Enterococcus*

### 1. Historique

Les entérocoques ont été identifiés pour la première fois en 1899 par Thiercelin qui leur donna le nom d'« entérocoques » en référence à leur origine intestinale (Thiercelin, 1899), puis la même année par MacCallum et Hastings (1899) sous le nom de *Micrococcus zymogenes* lorsqu'ils isolèrent une souche d'une endocardite. Cette même bactérie fut ensuite détectée comme responsable d'une endocardite par Andrewes et Horder (1906), qui lui donnèrent le nom de *Streptococcus faecalis*. De ce fait, les autres entérocoques actuels ont tout d'abord été classés

dans le genre *Streptococcus* avant d'être transférés en 1984 dans le genre *Enterococcus* nouvellement créé, sur la base de la séquence de leur ARN 16S (Schleifer et Kilpper-Bälz, 1984). Cependant les bactéries du genre *Enterococcus* sont toujours répertoriées dans le groupe D de la classification de Lancefield, la classification sérologique des streptocoques (Lancefield, 1933; Murray, 1990).

## 2. Caractéristiques du genre *Enterococcus*

Les entérocoques sont des bactéries lactiques du phylum des Firmicutes, à Gram positif et à faible pourcentage de GC (35% dans le cas d'*E. faecalis*). Elles appartiennent à la famille des *Enterococcaceae* (Gilmore *et al.*, 2014). Ces bactéries sont de forme ovoïde, d'environ 0,6 à 2,5  $\mu\text{m}$  de diamètre (Thiercelin, 1899) (Figure 2), isolées ou groupées en paire, anaérobies facultatives et non sporulantes (Murray, 1990), et ont une température optimale de croissance d'environ 35°C. Elles sont chimioorganotrophes, leur métabolisme est homofermentaire et elles sont considérées comme catalase négative, bien qu'elles possèdent un gène fonctionnel puisque l'activité catalase est détectable dans certaines conditions, comme en présence d'hème (Frankenberg *et al.*, 2002).

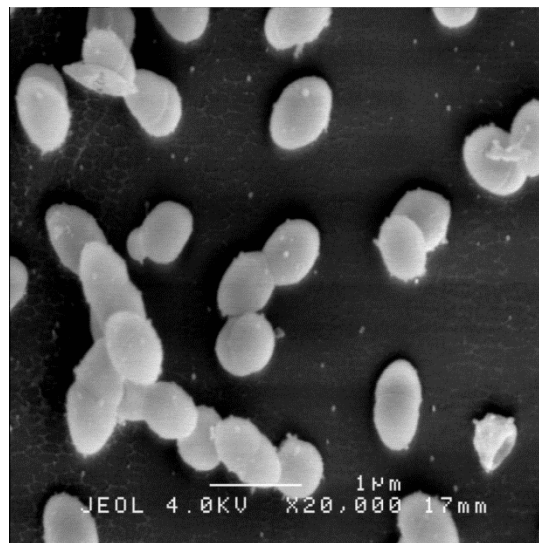


Figure 2 : *E. faecalis*, souche V19, vue en microscopie électronique à balayage (source : cette étude).

Les entérocoques se distinguent des streptocoques notamment par leur grande capacité de résistance à de nombreux stress. Ainsi, les entérocoques sont capables de croître à des températures comprises entre 10°C et 45°C, et de survivre à un choc chaud à 60°C pendant 30 minutes. Elles tolèrent également de grandes concentrations en sels (jusqu'à 28% de NaCl), des variations de pH importantes (pH 4,8 à 9,6), un stress oxydatif, et une exposition aux rayonnements ultra-violet (UV à 240 nm) (Sherman, 1937; Hartke *et al.*, 1998; Yuen et Ausubel, 2014). En conditions de carence prolongée en glucose (plus de 24h), leur résistance à certains stress augmente, c'est le cas notamment pour les stress thermique, oxydatif et acide (Giard *et al.*, 1996).

Chez l'homme adulte, le genre *Enterococcus* représente environ 1% de la flore intestinale (Gilmore *et al.*, 2014), qui est un écosystème bactérien très complexe et bénéfique pour l'organisme.

### 3. *Enterococcus faecalis*

#### a) Environnement et applications

Les entérocoques sont généralement associés au tractus gastro-intestinal des mammifères, mais sont également retrouvés chez les reptiles, les oiseaux, et même les insectes (Martin et Mundt, 1972). *E. faecalis* ne fait pas exception, et elle est même considérée comme ubiquitaire (Rincé *et al.*, 2003) : elle est retrouvée dans de nombreux environnements et est notamment utilisée comme indicateur de contamination fécale (Ostrolenk *et al.*, 1947).

*E. faecalis* est aussi utilisée comme probiotique chez l'homme avec le mélange de souches Symbioflor® (Fritzenwanker *et al.*, 2016) et dans l'élevage porcin, où une souche particulière empêche le développement de *Salmonella Typhimurium* (Kim *et al.*, 2017). Elle pourrait aussi avoir des applications dans le domaine médical. En effet, certaines souches ont montré des

propriétés cytotoxiques contre des lignées cellulaires cancéreuses tout en épargnant les cellules saines (Nami *et al.*, 2014; Ran *et al.*, 2018).

#### b) Données génomiques

D'après la base de données NCBI, 23 souches d'*E. faecalis* ont été complètement séquencées. Le génome de la souche type V583 contient potentiellement 3412 gènes répartis sur le chromosome (3,22 Mb, 3257 gènes) et sur trois plasmides : pTEF1 (66 kb, 74 gènes), pTEF2 (58 kb, 62 gènes) et pTEF3 (18 kb, 19 gènes). Plus d'un quart de ce génome est composé d'éléments génétiques mobiles comme des prophages, des séquences insérées, des transposons et un îlot de pathogénicité (Paulsen *et al.*, 2003 ; Van Tyne et Gilmore, 2014).

#### c) Opportunisme et virulence

Comme la plupart des entérocoques, *E. faecalis* est le plus souvent inoffensive chez les sujets sains. Cependant, cette espèce est l'une des plus fréquemment impliquées dans les infections liées aux soins : on parle alors de pathogène opportuniste. Certaines souches peuvent provoquer des infections urinaires, mais également des septicémies, des endocardites et des méningites dans les cas les plus graves (Babady, 2016). Les infections urinaires à entérocoques sont le plus souvent acquises *via* des dispositifs médicaux comme les cathéters urinaires, notamment lors d'hospitalisations à longue durée. Des bactériémies résultent le plus souvent de ces infections, et elles-mêmes sont susceptibles d'évoluer en endocardites, forme la plus grave d'infection à entérocoques. Le genre *Enterococcus* est la cause de 5 à 15% des endocardites bactériennes, et *E. faecalis* est l'espèce la plus souvent retrouvée pour ces pathologies, avec un taux de mortalité qui s'élève à 9 à 15% (Gilmore *et al.*, 2014).

Ce caractère opportuniste est dû à de nombreux facteurs très divers et dont les principaux sont répertoriés dans le Tableau 1.

Tableau 1: Principaux facteurs d'opportuniste ou de virulence décrits chez *Enterococcus faecalis*, avec leurs rôles connus lors du processus d'infection et leur pathologie associée.

<b>Facteurs de virulence</b>	<b>Exemples</b>	<b>Rôles principaux</b>	<b>Pathologie associée</b>	<b>Références</b>
<b>Facteurs de la surface cellulaire</b>	Substance d'agrégation (Asc10 (PrgB), Asp1, Asa1)	Interaction et internalisation par les cellules de l'hôte, évasion du système immunitaire	Endocardite	(Arias et Murray, 2012; Madsen <i>et al.</i> , 2017 ; Gilmore <i>et al.</i> , 2014)
	Protéine de surface (Esp)	Formation de biofilm, associée aux souches résistantes à l'ampicilline	Infection urinaire	
	Adhésine au collagène (Ace),	Attachement aux tissus de l'hôte, bactériémie	Endocardite	
	Protéine A liant la fibronectine (EfbA),	Attachement à la matrice extracellulaire	Endocardite	
	Métalloprotéase de membrane (Eep),		Endocardite	
	Glycolipides (BgsA, BgsB),		Endocardite	
<b>Facteurs sécrétés</b>	Hémolysine, Cytolysine (Cyl), Gélatinase (GelE),	Cytotoxicité	Endocardite	(Arias et Murray, 2012; Madsen <i>et al.</i> , 2017 ; Gilmore <i>et al.</i> , 2014)
	Sérine protéase (SprE)	Survie en infection intrapéritonéale (souris)		
<b>Protéines de stress</b>	Protéine générale de stress Gls24,	Résistance aux stress	Endocardite	(Madsen <i>et al.</i> , 2017 ; Gilmore <i>et al.</i> , 2014)
	Peroxydases (Npr, Tpx)	Résistance dans les macrophages		
<b>Pili et fimbriae</b>	EbpABC	Adhérence au collagène, formation de biofilm, interaction avec les plaquettes	Endocardite	(Madsen <i>et al.</i> , 2017 ; Gilmore <i>et al.</i> , 2014)

Les gènes codant ces facteurs peuvent être situés sur le chromosome bactérien, mais aussi sur des éléments mobiles comme des plasmides ou un îlot de pathogénicité présent chez certaines

souches. On trouve ainsi des toxines sécrétées dans le milieu extérieur, comme par exemple l'hémolysine-cytolysine Cyl, capable de lyser les hématies chez l'homme, le cheval et le lapin (Arias et Murray, 2012), ou encore la gélatinase GelE, qui dégrade le collagène, le fibrinogène, la fibrine et des éléments du système du complément (Madsen *et al.*, 2017). Des protéines de surface comme les substances d'agrégations, mais également des protéines de réponses aux stress (notamment la protéine Gls24), ont été montrées comme étant impliquées dans la virulence d'*E. faecalis* (pour revue : Madsen *et al.*, 2017).

Bien que *E. faecalis* possède relativement peu de facteurs de virulence en comparaison à d'autres pathogènes, elle est redoutable par sa grande capacité de résistance intrinsèque et sa facilité d'acquisition de résistance à de nombreux antibiotiques et à d'autres agents antimicrobiens. Ces dernières années ont notamment vu émerger un nombre croissant de souches résistantes à la vancomycine, antibiotique de seconde ligne contre cette famille de bactéries (Murray, 1990; Paulsen *et al.*, 2003; Arias et Murray, 2012). *E. faecalis* a également un effet sur les cellules de l'hôte en étant capable d'induire la caspase-1 et l'inflammasome chez les cellules ostéoblastiques humaines, et à terme provoquer leur apoptose (Ran *et al.*, 2018). Ainsi, des caractéristiques comme les facteurs de virulence, l'antibiorésistance mais aussi la grande capacité à se multiplier dans un environnement hostile participent à la persistance d'*E. faecalis* en milieu hospitalier.

#### d) Métabolisme

Comme tous les entérocoques, *E. faecalis* est une bactérie lactique homofermentaire, ce qui signifie qu'elle produit de l'acide lactique comme principal produit de la fermentation. Son cytochrome bd n'est actif qu'en présence d'hème (Winstedt *et al.*, 2000) et elle ne possède pas



de porphyrine, elle est donc incapable de produire de l'énergie par voie oxydative, se reposant presque uniquement sur la voie fermentaire (Brock *et al.*, 2007). Les principales sources d'énergie sont les sucres, ou carbohydrates. *E. faecalis* est capable de métaboliser au moins 13 sucres différents *via* ses trois voies de métabolisation : Emben-Meyerhof-Parnas (glycolyse), Entner-Doudoroff et pentoses phosphate. Les deux premières permettent la synthèse d'ATP par des phosphorylations successives. La voie des pentoses phosphates induit la production de pentoses, nécessaires à la synthèse des acides nucléiques à partir du glucose. Elle permet également la production d'énergie sous la forme de NADPH (Brock *et al.*, 2007).

Pour pouvoir métaboliser les carbohydrates, la bactérie doit d'abord les importer. Pour cela, *E. faecalis* dispose de trois types de de systèmes transport : les perméases (transport simple), les systèmes ABC (ATP-Binding Cassette) et les systèmes PTS (Phospho-Transferase System). Chez les entérocoques, les systèmes de transport des sucres les plus couramment utilisés sont les systèmes PTS. Ces systèmes sont spécifiques d'un ou plusieurs sucres qu'ils importent dans la cellule tout en les phosphorylant, les rendant ainsi immédiatement métabolisables. Ils sont composés d'au moins 5 enzymes (Figure 3) : EI, Hpr, EIIA, EIIB et EIIC. EI et Hpr ne sont pas spécifiques d'un sucre, contrairement aux trois enzymes EII. Les protéines EI, Hpr et EIIa sont cytoplasmiques, tandis qu'EIIB est liée à la membrane et EIIC est intégralement membranaire. Parfois, une sixième enzyme est présente, EIID, qui est également membranaire. L'énergie nécessaire au fonctionnement du PTS est fournie par le phosphoénolpyruvate (PEP) et est transportée jusqu'à l'enzyme EIIC via une cascade de phosphorylation (Postma *et al.*, 1993; Brock *et al.*, 2007; Gilmore *et al.*, 2014).

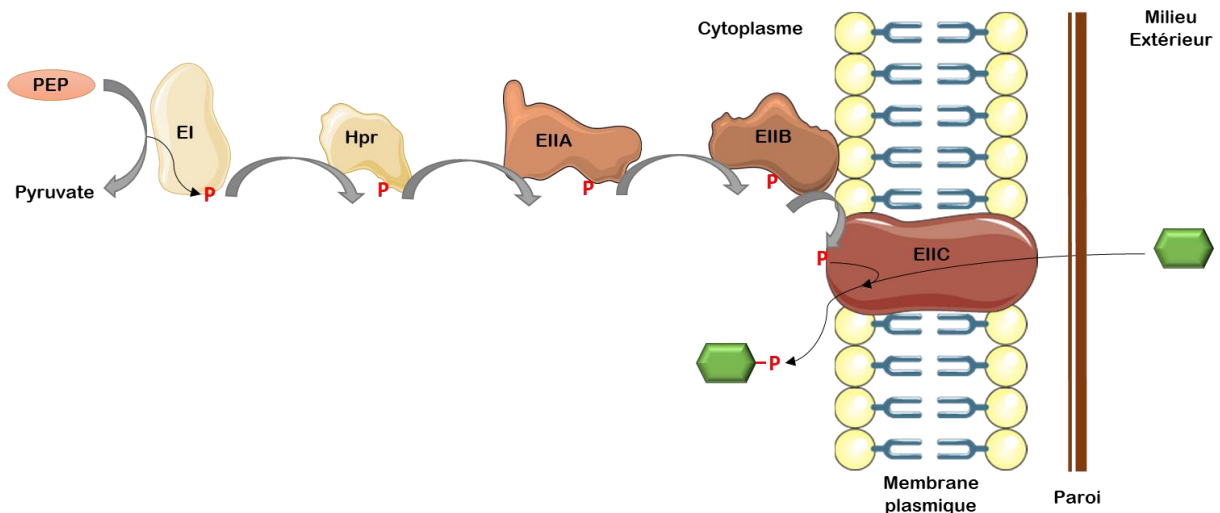


Figure 3 : Les systèmes PTS chez les bactéries. Le groupement phosphate est fourni initialement par le PEP (Phosphoénolpyruvate) qui est converti en pyruvate lors de la phosphorylation de l'enzyme EI. Le groupement phosphate est ensuite transféré à Hpr, puis à EIIA et enfin à EIIB et EIIC qui vont phosphoryler le sucre (en vert) en le transportant à l'intérieur de la bactérie.

Chez *E. faecalis*, 97 gènes annotés comme codant des protéines de systèmes PTS ont été référencés sur la base de données KEGG (Kanehisa *et al.*, 2017). Ils transportent des sucres très variés, que ce soient des monosaccharides ou des sucres plus complexes.

Les systèmes PTS d'une bactérie ne sont généralement pas tous exprimés de façon simultanée. En effet, la bactérie adapte l'expression des gènes codants ces systèmes de transport en fonction des sources nutritives présentes dans le milieu. Cette modulation de l'expression des gènes PTS permet d'éviter de synthétiser des transporteurs pour des carbohydrates non disponibles mais aussi de privilégier l'utilisation des sucres les plus facilement métabolisables. Elle fait appel à des mécanismes d'induction de l'expression des gènes PTS en présence du sucre correspondant et à des mécanismes qui au contraire permettent de réprimer des gènes PTS en présence d'autres sucres facilement assimilables (comme le glucose) (Meyer *et al.*, 2011). Le principal mécanisme de répression est appelé « répression catabolique » et implique une protéine clé appelée CcpA chez la plupart des bactéries à Gram positif, dont *E. faecalis*. Cette protéine, active en présence de glucose, va se fixer sur une séquence spécifique appelée boîte

*cre* (catabolite responsive element), située dans la région promotrice des gènes à réprimer. CcpA utilise comme co-facteur la protéine Hpr phosphorylée sur la sérine 46 (Marciniak *et al.*, 2012). En formant un complexe, les deux protéines acquièrent une capacité de fixation à l'ADN (Jones *et al.*, 1997). Ce système de régulation permet à la bactérie d'exprimer les gènes nécessaires à l'assimilation du sucre le plus facilement métabolisable, et ainsi d'économiser une grande quantité d'énergie. C'est un bon exemple d'adaptation régulée par une protéine. Il existe cependant d'autres acteurs de la régulation de l'expression des gènes comme les ARN régulateurs.

### III. La régulation transcriptionnelle et post-transcriptionnelle par les ARN régulateurs

#### 1. Généralités sur la régulation de l'expression des gènes

Chez une bactérie, les gènes dits « de ménage » ou domestiques sont exprimés en permanence quelles que soient les conditions extérieures. Mais d'autres gènes impliqués dans des voies métaboliques voient leur expression modulée. Ces gènes nécessitent d'être exprimés uniquement lorsque les conditions l'exigent, comme par exemple un stress ou la présence d'un sucre pour son PTS correspondant. Cela passe en premier lieu par la captation d'un signal (de nature physique, chimique, ou encore biologique) qui aboutit à une modulation de l'expression des gènes et ainsi permettra de répondre à une condition spécifique.

Cette régulation est souvent exercée par des protéines, comme par exemple les régulateurs Fur d'*Escherichia coli* (Troxell et Hassan, 2013) et *E. faecalis* (Latorre *et al.*, 2018), TstR chez *Lactobacillus brevis* (Pagliai *et al.*, 2014), au encore CcpA, évoqué précédemment. Sont également très connus les systèmes à deux composantes, constitué d'un senseur membranaire et d'un régulateur de transcription cytoplasmique (Monedero *et al.*, 2017). Les facteurs sigma

sont également très étudiés, étant des acteurs centraux de la transcription et de l'expression de nombreux gènes (essentiels, impliqués dans la réponse au stress, dans le quorum-sensing, la synthèse des flagelles et des endospores,...) (Davis *et al.*, 2017).

Cependant, il existe d'autres moyens de modulation de l'expression des gènes comme les ARN régulateurs. Ces ARN régulateurs permettent de modifier l'expression des gènes de manière fine et précise et constituent un mode de régulation rapide et économe en énergie comparativement aux régulateurs de nature protéique qui nécessitent une étape de traduction. Il en existe différents types : les petits ARN (sARN), les ARN liant les protéines, les CRISPR, et les 5'UTR (UnTranslated Region). Leur action régulatrice peut s'effectuer à plusieurs niveaux : transcriptionnel, post-transcriptionnel et post-traductionnel (Figure 4).

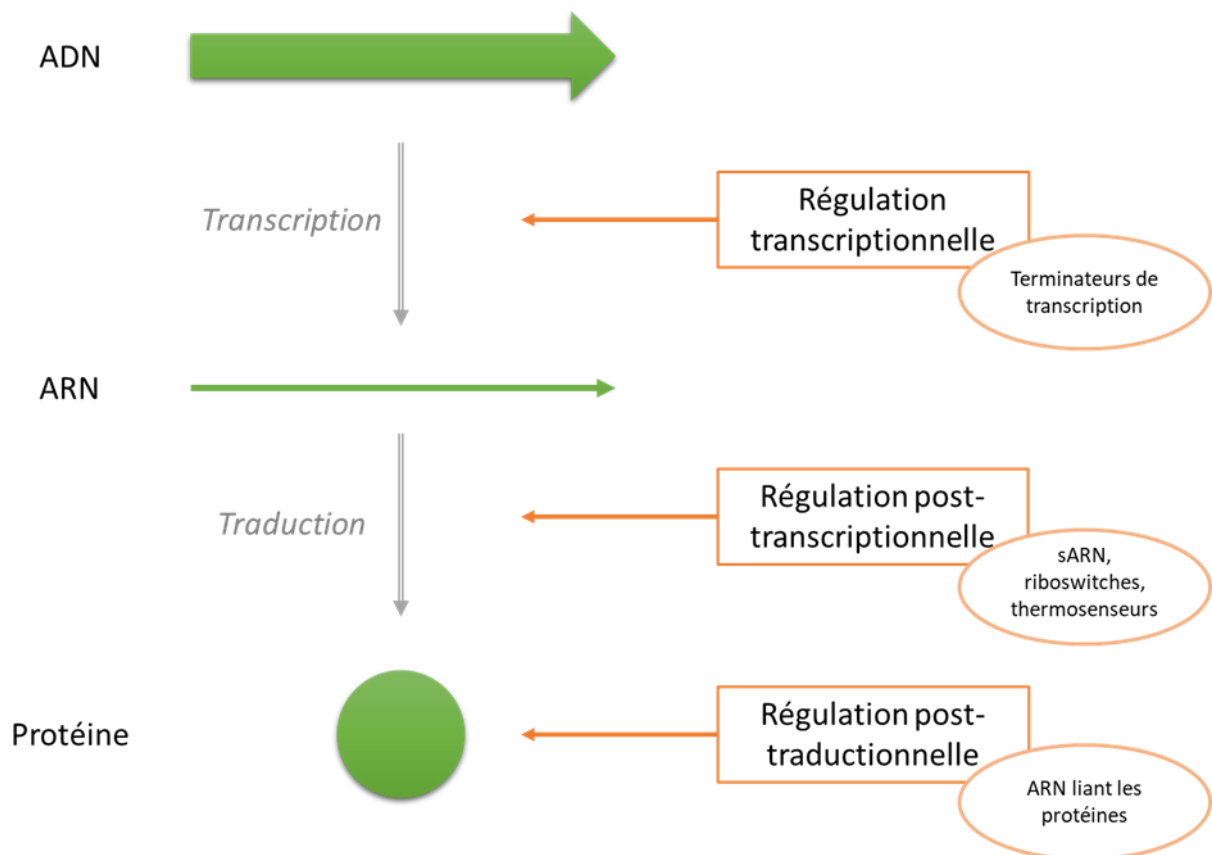


Figure 4 : Les différents niveaux de régulation de l'expression des gènes par les ARN régulateurs.

## 2. Les petits ARN

Les petits ARN, ou sARN, ont une taille variant généralement de 50 à 500 nucléotides (Gottesman et Storz, 2011), et agissent en se liant directement à leur cible. Ils ont été découverts pour la première fois en 1984 chez *E. coli* (Mizuno *et al.*, 1984). Chez de nombreuses espèces, les sARN ont été montrés comme étant impliqués dans des mécanismes extrêmement variés, allant du métabolisme du carbone, des acides aminés ou du fer, à la formation de biofilms, le quorum sensing, la réponse au stress et la virulence (Michaux *et al.*, 2014b). Il en existe deux catégories principales : les sARN codés en *cis*, et les sARN codés en *trans*.

### a) Les petits ARN codés en *cis*

Les sARN codés en *cis*, dits antisens, sont codés sur le brin opposé de leur ARNm cible. Ils sont ainsi capables de s'y lier par complémentarité parfaite de plus de 75 nucléotides (variable selon le sARN) et par conséquent d'inhiber leur expression (blocage de la séquence RBS, inhibition de la traduction et dégradation de l'ARNm cible) (Figure 5, A) ou de l'induire (fixation sur des structures secondaires inhibitrices). Ils sont le plus souvent impliqués dans la régulation du nombre de copies d'éléments mobiles (plasmides, transposons,...), ou du taux de protéines potentiellement toxiques pour la bactérie, comme par exemple IstR et OhsC chez *E. coli* (Waters et Storz, 2009; Gottesman et Storz, 2011; Michaux *et al.*, 2014b).

D'autres sARN antisens ont un rôle dans la régulation de l'expression de gènes faisant partie de structures opéroniques (Figure 5, B). Par exemple, chez *E. coli*, le sARN antisens *gadY* est complémentaire de l'extrémité 3' du premier gène de l'opéron *gadXW*, impliqué dans la résistance au stress acide (Tramonti *et al.*, 2008). La fixation de *gadY* sur ce site de liaison de l'ARNm provoque une augmentation de la transcription de *gadX* et réprime celle de *gadW*, probablement via un clivage de l'ARNm entre les deux ORF. Le sARN RNA $\beta$  constitue un autre exemple : il est localisé sur le plasmide de virulence pJM1 chez *Vibrio anguillarum* (Stork

et al., 2007) et provoque une terminaison prématurée de la transcription de l'opéron ITB (pour Iron-Transport-Biosynthesis). Il couvre une région intergénique entière de l'opéron, et agit comme un terminateur de transcription.

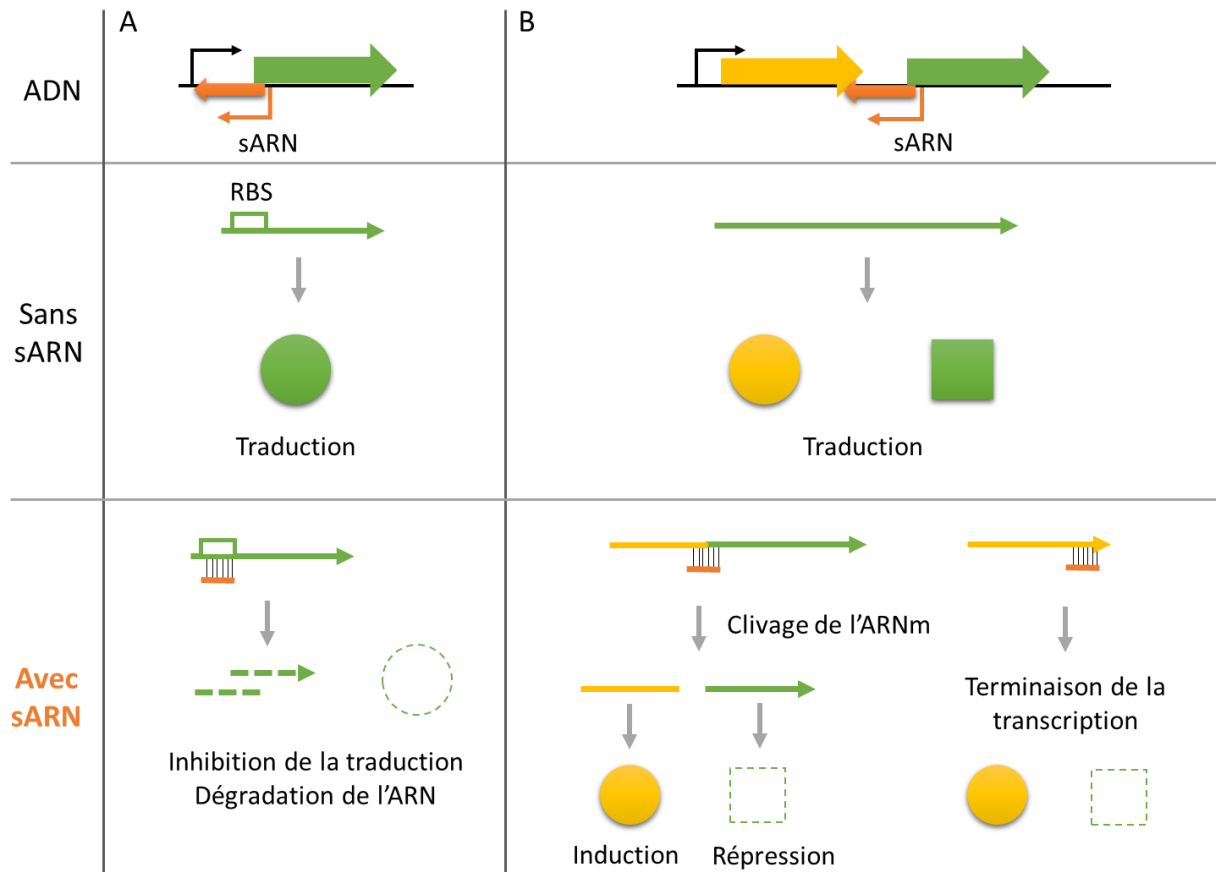


Figure 5: Modes d'action des sARN codés en cis régulant l'expression d'un gène isolé (A) ou d'un opéron (B). Les sARN se lient à leur cible par complémentarité parfaite de bases, occasionnant une répression ou une induction de l'expression des gènes cibles par clivage, dégradation de l'ARNm par des RNases ou terminaison prématurée de la transcription. RBS : Ribosome Binding Site (adapté de Waters et Storz, 2009)

#### b) Les petits ARN codés en trans

La deuxième catégorie de sARN désigne des ARN codés dans une région génomique éloignée de leur ARNm cible (codés « en trans »). Contrairement aux sARN antisens, ils partagent généralement une complémentarité limitée avec leur ARNm cible, et leur zone de liaison varie généralement de 10 à 25 bases environ (Waters et Storz, 2009). Beaucoup de ces sARN agissent

en se liant au niveau du site de liaison du ribosome RBS (Ribosome Binding Site), bloquant la traduction (Figure 6, A). D'autres sARN peuvent au contraire se lier au niveau d'une zone normalement capable de former une structure secondaire, et agissent en empêchant ce repliement, ce qui libère le site RBS et déclenche la traduction (Figure 6, B) (Gottesman et Storz, 2011). Parfois, la liaison se produit hors du site RBS ; la régulation de l'ARNm cible se fait alors par une augmentation de la sensibilité aux RNases, qui dégradent le complexe sARN-ARNm, comme c'est le cas pour les sARN SgrS et RyhB de *E. coli*, dont l'action est conjointe à la RNase E (Morita *et al.*, 2005).

Beaucoup de sARN de ce type sont impliqués dans la virulence. Par exemple, le sARN VR-RNA chez *Clostridium perfringens* régule positivement des gènes codant pour les toxines  $\alpha$  et  $\kappa$ , et est lui-même régulé par un système à deux composants, VirR/VirS (VR-RNA pour « VirR regulated RNA ») (Shimizu *et al.*, 2002).

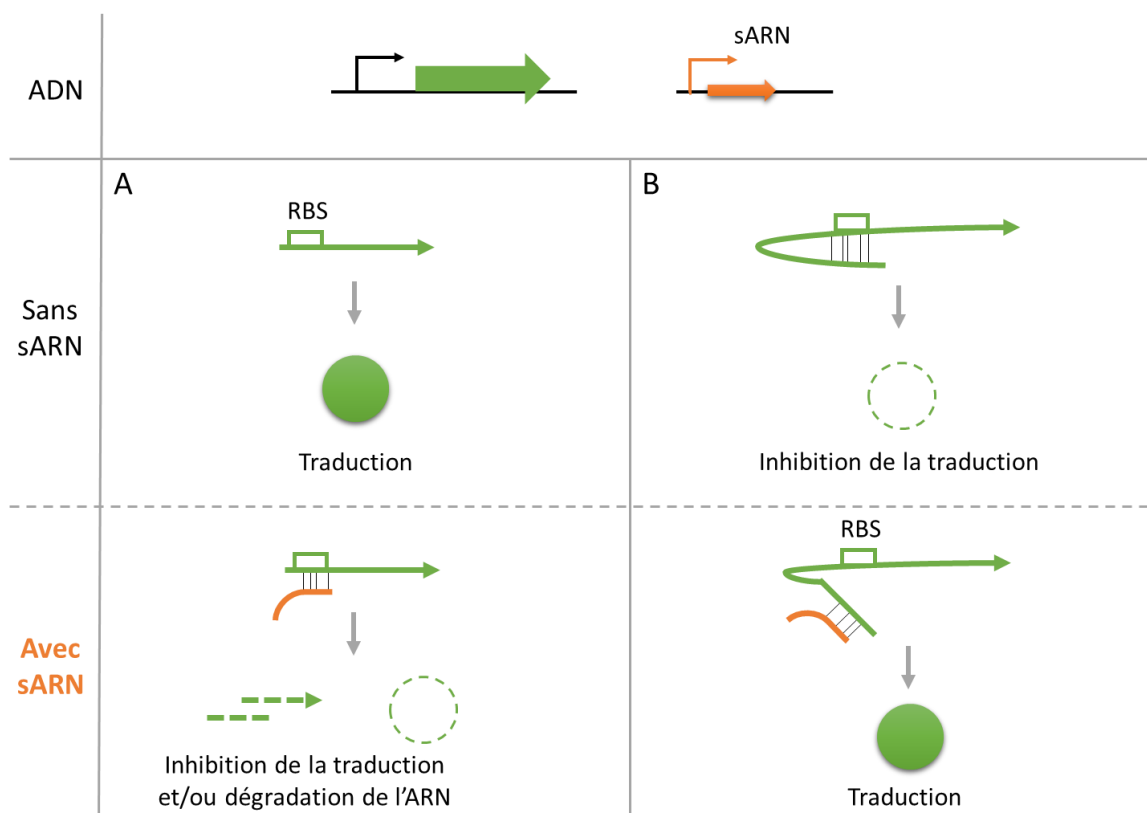


Figure 6 : Modes d'action des sARN codés en trans. Ils peuvent agir négativement (A) en se liant à l'ARNm au niveau du motif RBS ou positivement (B) en dépliant une structure secondaire qui bloque le motif RBS (adapté de Waters et Storz, 2009).

### c) Cas particuliers

Il existe des cas particuliers de sARN codés en *cis*, mais qui agissent également sur d'autres ARNm que leur cible complémentaire. Chez *Staphylococcus aureus*, RNAIII est connu comme étant un ARN antisens capable de réprimer toute une classe d'ARNm de gènes impliqués dans la virulence. Sa liaison avec son ARNm cible (codé en *cis*) crée un long ARN double brin qui est ensuite pris en charge par les RNases, en particulier la RNase III, ce qui conduit à sa dégradation. Il se fixe également par complémentarité imparfaite de base à l'ARNm du gène *rot* (codé en *trans*), ce dernier codant un régulateur transcriptionnel impliqué lui aussi dans la virulence (Boisset *et al.*, 2007).

### 3. Les ARN liant les protéines

Certains ARN sont capables de moduler l'activité des protéines en s'y liant directement. Il s'agit alors d'une régulation au niveau post-traductionnel (Figure 4). Leur action consiste le plus souvent à séquestrer une protéine cible agissant sur l'ADN ou l'ARN en ayant une séquence mimant celle de la cible de la protéine (Figure 7).



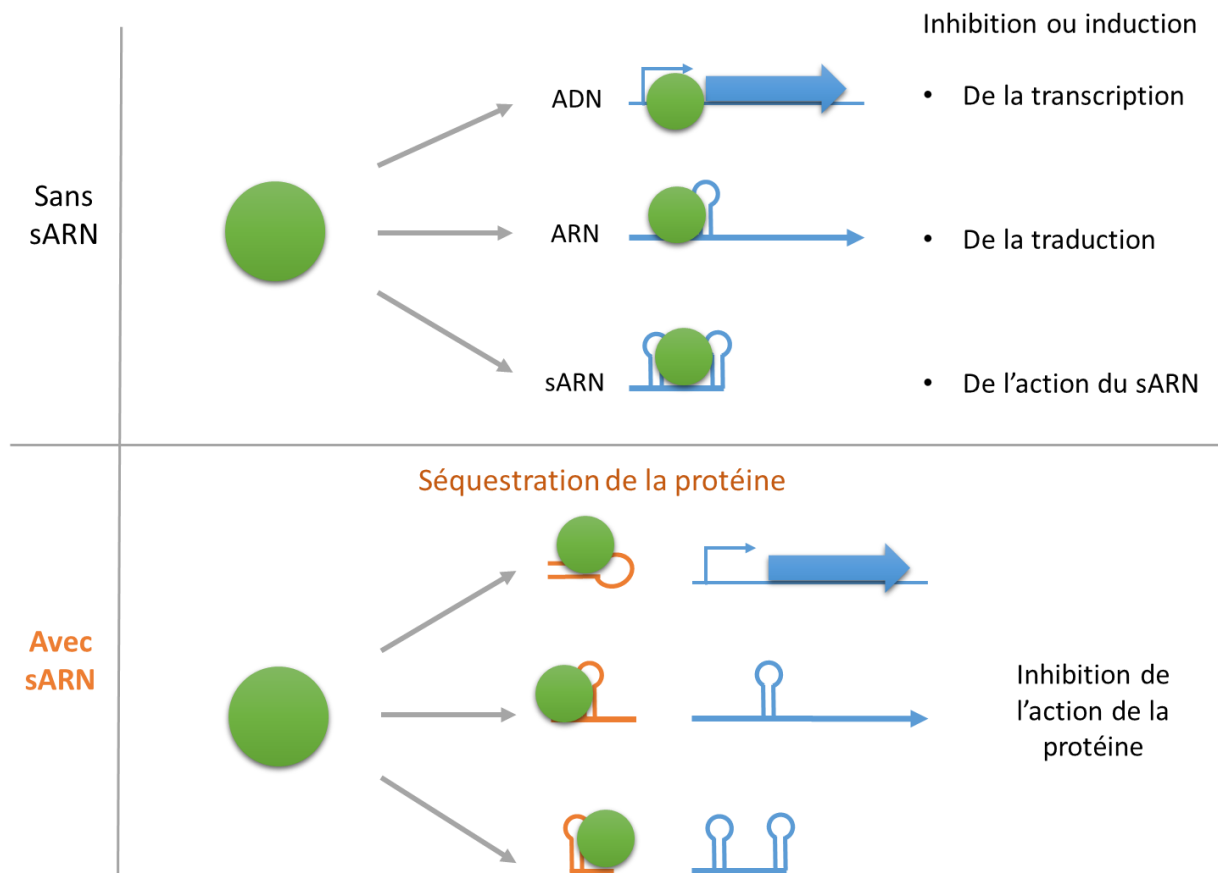


Figure 7 : Les différents modes d'actions des ARN régulateurs liant les protéines. Ces ARN agissent en séquestrant des protéines dont l'action repose sur la liaison à l'ADN ou à l'ARN. En mimant le site de liaison de la cible, ils séquestrent la protéine et l'empêchent d'agir sur sa cible.

Le plus connu de ces ARN est l'ARN 6S de *E. coli*, un ARN codé par le gène *ssrS*, et qui régule la transcription par interaction avec le site actif de l'ARN polymérase- $\sigma^{70}$  au niveau d'une structure secondaire imitant la structure ADN cible (Oliva *et al.*, 2015).

D'autres sARN, *csrB* et *csrC*, modulent l'activité de la protéine CsrA chez *E. coli*. Cette protéine liant l'ARN a un rôle dans la régulation du métabolisme du carbone et la motilité bactérienne. Les sARN *csrB* et *csrC* possèdent des sites imitant la séquence nucléotidique du site de liaison de CsrA présente sur l'ARNm cible. Ils se lient ainsi à cette protéine et inhibent son activité de liaison à l'ARNm (Waters et Storz, 2009).

L'ARN *rsmZ* constitue un autre exemple, il inhibe les effets de la protéine RsmA, un régulateur négatif de la synthèse de facteurs de virulence tels que la pyocyanine et des molécules du

quorum sensing, mais aussi un régulateur positif du swarming (mouvement coordonné d'une population de bactéries sur une surface semi-solide) ainsi que de la production de lipases et de rhamnolipides chez *Pseudomonas aeruginosa* (Heurlier *et al.*, 2004).

#### 4. Les CRISPR

Les CRISPR (Clustred Regularly Interspaced Short Palindromic Repeats) sont composés de régions constantes : des courtes séquences palindromiques (24 à 47 paires de bases) répétées de 2 à 249 fois sur et de séquences variables complémentaires des acides nucléiques de bactériophages et de plasmides entre autres. Les CRISPR sont un moyen de défense contre l'ADN exogène dont les bactériophages : en se fixant sur leur ADN, elles empêchent la réplication et/ou la transcription de ceux-ci. Elles sont aussi impliquées dans des mécanismes qui empêchent la conjugaison de plasmides (Waters et Storz, 2009). Des gènes Cas (CRISPR-associated) flanquent les CRISPR et codent des protéines essentielles à leur fonctionnement qui permettent la liaison avec leur ADN ou ARN cible. L'association entre les CRISPR et la protéine Cas-9 constitue un puissant outil de biologie moléculaire, grâce à sa capacité de clivage ciblé de l'ADN (Charpentier, 2015; Luo *et al.*, 2016).

#### 5. Les 5'UTR (UnTranslated Region) régulatrices

Les 5'UTR sont des portions d'ARN localisées à l'extrémité 5' de certains ARNm. Ces régions sont transcrites avec l'ARNm, mais ne sont traduites en protéines. Elles contiennent notamment la séquence de Shine-Dalgarno (SD, ou RBS) et ont une fonction régulatrice.

a) Régulation par blocage ou déclenchement de la traduction

Cette régulation s'opère souvent *via* la formation de structures secondaires, qui, en fonction de leur conformation, peuvent bloquer ou déclencher la traduction (Ruiz de los Mozos *et al.*, 2013) ( Figure 8). La régulation par ces 5'UTR peut aussi impliquer des protéines de liaison à l'ARN ou des sARN, comme par exemple la 5'UTR du gène *chiR* (codant un facteur de transcription induisant les gènes codant des chitinases) chez *Serratia marcescens* qui est la cible du sARN répresseur ChiX (Suzuki *et al.*, 2016).

Il existe également des types particuliers de 5'UTR, tels que les riboswitches et les thermosenseurs (Oliva *et al.*, 2015).

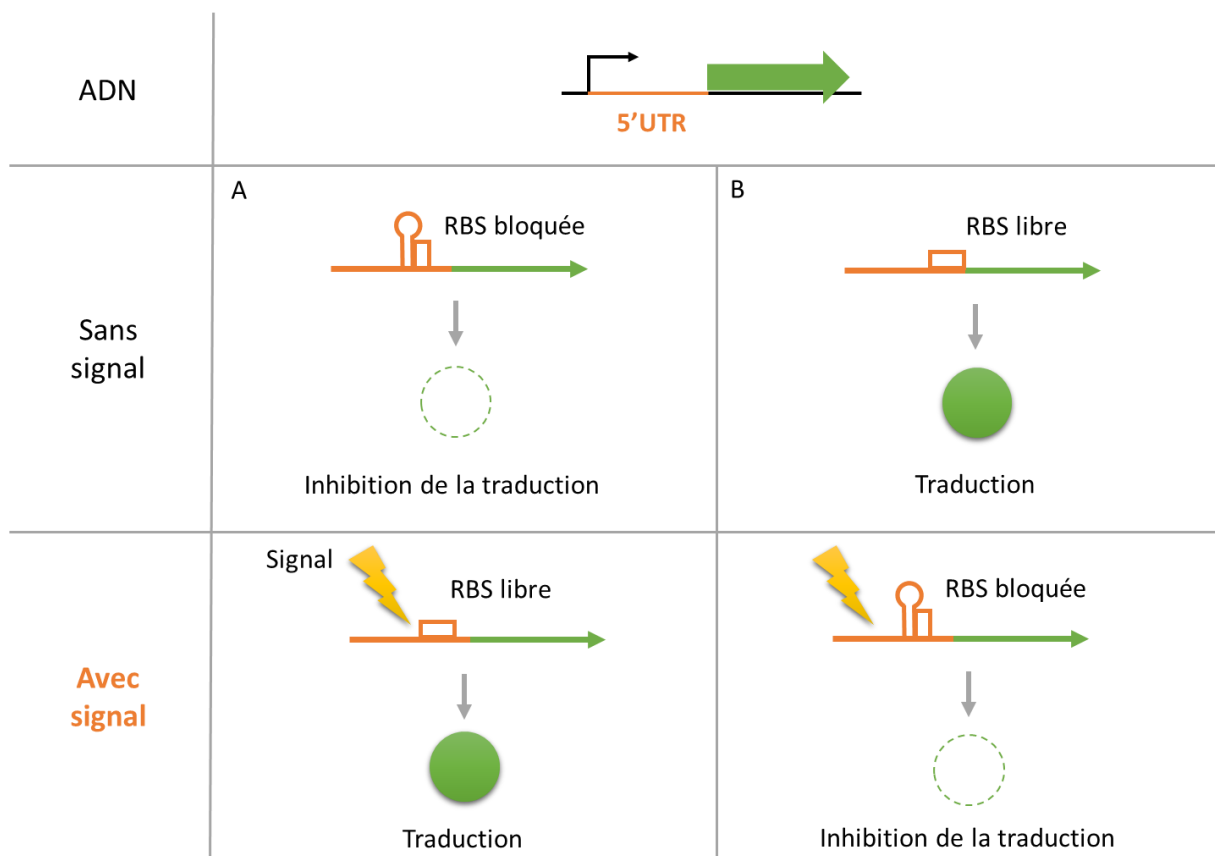


Figure 8 : Exemples d'action régulatrice par des 5'UTR. L'induction ou la répression du gène régulé est modulée par un signal inducteur (A) ou répresseur (B) qui aboutit au relâchement ou à la formation de structures secondaires. Ce signal peut être de nature protéique, ribonucléique, chimique ou physique.

Les riboswitches sont des 5'UTR capables de fixer un métabolite. Cela produit un changement de conformation de l'ARN, ce qui induit ou réprime la transcription ou la traduction du(es) gène(s) en aval, qui est(sont) le plus souvent impliqué(s) dans l'acquisition ou l'utilisation de ce métabolite (Waters et Storz, 2009; Oliva *et al.*, 2015). Certains riboswitches permettent la résistance aux antibiotiques. En effet, Jia *et al.* ont mis en évidence l'existence d'un riboswitch répondant aux aminoglycosides et présent dans la région 5'UTR de gènes impliqués dans la résistance à cette famille d'antibiotique (Jia *et al.*, 2013).

Les thermosenseurs, quant à eux, sont des 5'UTR capables de moduler l'expression des gènes par des changements de conformations dus à des variations de température (Oliva *et al.*, 2015). Classiquement, une augmentation de la température défait des structures secondaires bloquant la séquence RBS. Le thermosenseur fourU par exemple est situé sur la 5'UTR de plusieurs gènes de virulence chez des bactéries pathogènes comme *Yersinia pestis*, *Samonella Typhimurium* et *Listeria monocytogenes* (Kortmann et Narberhaus, 2012).

#### b) Régulation par terminaison prématurée de la transcription

D'autres 5'UTR agissent en provoquant une terminaison prématurée de la transcription de l'ARNm (Ruiz de los Mozos *et al.*, 2013) par la formation d'un terminateur sur le fragment d'ARN néo-transcrit. Ce terminateur peut être déroulé par exemple par l'action d'un anti-terminateur ou d'une autre molécule effectrice.

Ainsi chez *E. coli*, l'opéron *bgl* est constitué de deux gènes, *bglG* et *bglF*, le premier codant pour l'anti-terminateur, et le deuxième pour un système PTS spécifique des  $\beta$ -glucosides. Le rôle de BglG est d'auto-réguler l'opéron *bgl* en déroulant le terminateur situé sur sa 5'UTR (Amster-Choder, 2005) (Figure 9).

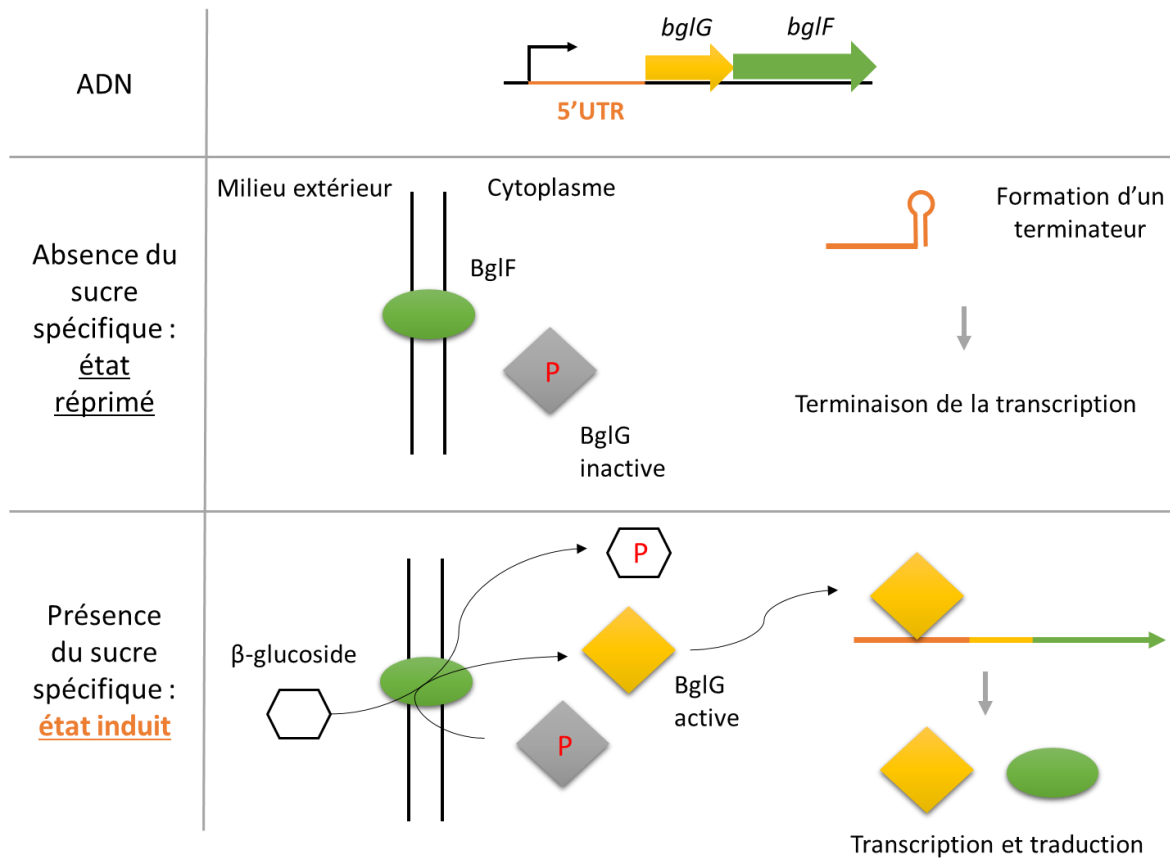


Figure 9 : Schéma explicatif de la régulation de type terminateur/anti-terminateur de l'opéron *bgl* chez *E. coli*. L'activité de la protéine BglG est modulée par BglF, qui la déphosphoryle en présence de son sucre spécifique. BglG est ainsi activée, et peut aller se fixer sur la 5'UTR en amont de l'opéron *bgl*, permettant ainsi la transcription et la traduction de l'opéron, ce qui à terme permet une augmentation de la quantité de protéines du système PTS et ainsi favorise l'importation du sucre (adapté de Amster-Choder, 2005).

## 6. Les 3'UTR régulatrices

Comme les 5'UTR, les 3'UTR ne sont pas traduites en protéines et ont elles aussi un rôle dans la régulation post-transcriptionnelle des gènes. Comme leur nom l'indique, elles sont situées en position 3' de certains ARNm et peuvent agir en *cis* sur ces ARNm.

Certaines 3'UTR ont été montrées comme étant régulatrices de l'expression de leur transcrit. Cette régulation peut être positive en protégeant l'ARNm contre l'action des RNAses, ou en favorisant l'initiation de la traduction en interagissant avec les 5'UTR. Les 3'UTR peuvent

aussi exercer une répression de leur transcrit en tant que cible de fixation de petits ARN ou de protéines (Ren *et al.*, 2017).

Chez *Salmonella enterica* serovar Typhimurium, l'ARNm *hilD* codant pour un régulateur de virulence, possède une 3'UTR agissant en *cis*. La liaison d'un sARN sur cette 3'UTR et l'intervention de protéines de liaison à l'ARN comme Hfq, Spot42 et ProQ comme co-facteurs influencent la traduction du transcrit (El Mouali et Balsalobre, 2018).

N'étant étudiées que depuis très récemment, peu d'informations sont disponibles concernant ces régions 3' non traduites. On estime cependant qu'un tiers des ARNm de *S. aureus* portent des 3'UTR d'une longueur supérieure à 100 nucléotides (Ruiz de los Mozos *et al.*, 2013). Beaucoup reste donc encore à découvrir sur ce mode de régulation post-transcriptionnelle.

#### IV. Les protéines impliquées dans le métabolisme des ARN

La régulation par les ARNs met également en jeu des protéines impliquées dans leur stabilisation ou leur dégradation.

##### 1. Les protéines chaperonnes à ARN

La liaison entre un sARN et un ARNm peut être modulée par une protéine chaperonne, comme la protéine Hfq découverte chez *E. coli*, et présente chez de nombreuses espèces bactériennes mais aussi chez certaines Archeae (Attaiech *et al.*, 2017; Santiago-Frangos et Woodson, 2018). Son rôle est de faciliter la liaison entre les sARN et leur ARNm cible, mais elle est aussi impliquée dans d'autres mécanismes tels que la traduction et le quorum-sensing (Lenz *et al.*, 2004; Santiago-Frangos et Woodson, 2018). Cependant, les Firmicutes des genres *Lactococcus*, *Streptococcus* et *Enterococcus* ne possèdent pas cette protéine (Sun *et al.*, 2002). D'autres

protéines chaperonnes de l'ARN ont été découvertes, comme ProQ chez *Salmonella enterica*, ou RocC chez *Legionella pneumophila*, mais il s'avère que ces protéines sont également absentes chez les Firmicutes (Attaiech *et al.*, 2017). Chez *E. faecalis*, une Cold-Shock Protéine, CspR, est soupçonnée de tenir ce rôle de chaperonne de l'ARN (Michaux *et al.*, 2012).

## 2. Le dégradosome

Le dégradosome est un complexe de protéines dont le rôle premier est la dégradation des ARNm. Il est notamment directement impliqué dans la régulation post-transcriptionnelle en dégradant les ARN bicaténaires formés par la liaison entre les sARN et leur ARNm cible. Le rôle du dégradosome est prépondérant pour le fitness de la cellule (Khemici et Carpousis, 2003; Lehnik-Habrink *et al.*, 2010).

Chez *E. coli*, les principales protéines du dégradosome sont les exonucléases RNase E et PNPase (polynucléotide phosphorylase), l'hélicase RhlB et l'enzyme glycolytique émolase. L'enzyme centrale de ce mécanisme est la RNase E, qui clive les ARNm par son activité endonucléase. La dégradation de l'ARNm se poursuit ensuite grâce à la PNPase, qui possède une activité 3'exonucléase.

Chez les bactéries à Gram positif, la RNase E est le plus souvent absente. Chez *B. subtilis* et *S. aureus*, les acteurs principaux du dégradosome ont été identifiés : la RNase J1, la RNase J2, la RNase Y, la PNPase, l'émolase, la phosphofructokinase PfkA et une hélicase à motif DEAD (CshA) (Figure 10) (Lehnik-Habrink *et al.*, 2010; Roux *et al.*, 2011).

Le rôle de l'émolase est peu connu, bien que son importance dans la dégradation des ARNm ait déjà été démontrée (Carpousis, 2007; Lehnik-Habrink *et al.*, 2010). Chez *B. subtilis*, elle interagit avec la phosphofructokinase PfkA, toutes deux étant des enzymes glycolytiques, et

avec la RNase Y, elle-même associée à la RNase J1, indiquant un rôle dans la coordination de la formation du dégradosome (Newman *et al.*, 2012; Lehnik-Habrink *et al.*, 2012).

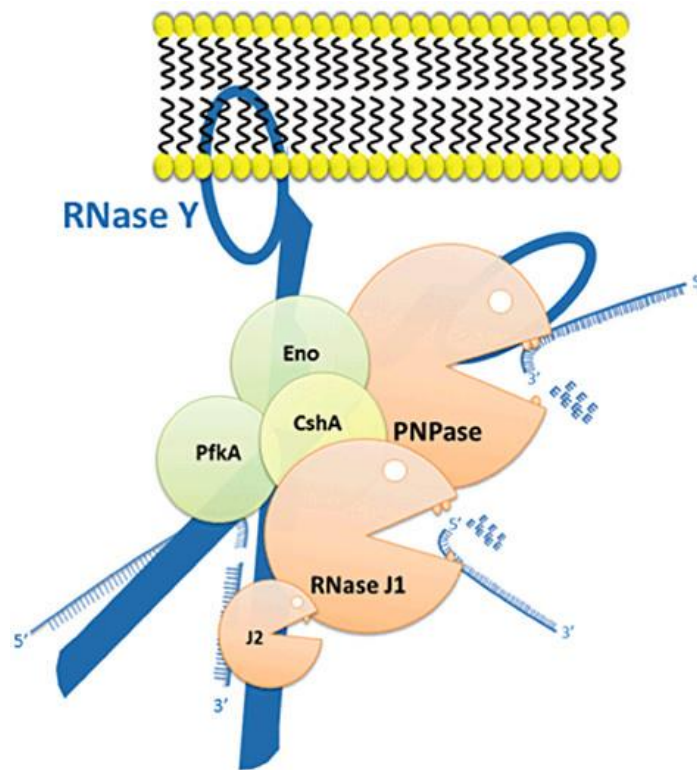


Figure 10 : Schéma des composants principaux du dégradosome de *Bacillus subtilis*. La RNase Y initie la formation du complexe et se lie à la PNPase, aux RNases J1 et J2, et avec les enzymes glycolytiques émolase (Eno) et phosphofruktokinase (PfkA). L'hélicase CshA permet de dérouler les structures secondaires de l'ARN pour le rendre accessible à l'action des RNases (Lehnik-Habrink *et al.*, 2012).

#### a) Les RNases

La RNase Y de *B. subtilis* est une endoribonucléase considérée comme l'équivalent fonctionnel de la RNase E d'*E. coli*. Liée à la membrane, elle initie la dégradation des ARN en interagissant avec le reste du dégradosome (Lehnik-Habrink *et al.*, 2011) et en effectuant un clivage interne des ARN à extrémité triphosphate, générant un ARN à extrémité monophosphate (Laalami et Putzer, 2011). Elle est également impliquée dans le turnover des ARN, dans le fonctionnement de certains riboswitches, et dans la dégradation d'ARNm très structurés (Lehnik-Habrink *et al.*, 2010; Roux *et al.*, 2011).



Les RNases J1 et J2 sont deux enzymes paralogues (70% de similarité d'acides aminés chez *B. subtilis*) possédant des activités endonucléolytique et exonucléolytique. Elles jouent un rôle prépondérant dans la maturation des ARN, mais aussi dans leur dégradation. Elles sont en effet des acteurs principaux du dégradosome chez *B. subtilis* et *S. aureus*. Elles exercent ensemble une activité de clivage interne équivalente à celle de la RNase Y et de la RNase E d'*E. coli* (Roux *et al.*, 2011; Jamalli *et al.*, 2014), mais aussi une activité exonucléasique des ARN à extrémités monophosphate issus de tels clivages (Laalami et Putzer, 2011).

La RNase III est une autre RNase présente chez *E. coli* et chez certaines bactéries à Gram positif comme *E. faecalis* (39% d'identité protéique, Johnson *et al.*, 2011), ou *S. aureus* (Lioliou *et al.*, 2012). Elle a un rôle dans la dégradation des ARN, notamment par sa capacité à cliver les ARN à extrémité triphosphate, fonction absente chez les RNases E, J1 et Y. Les ARN à extrémité monophosphate ainsi produits sont ensuite ciblés par les autres RNases du microorganisme. Elle est également capable de cliver les ARN double brins, ce qui indiquerait un rôle dans la dégradation des complexes sARN-ARNm. Elle est aussi connue pour son activité de répression sur l'expression de plusieurs gènes codant pour des facteurs de virulence chez *S. aureus* (Johnson *et al.*, 2011; Lioliou *et al.*, 2012).

La PNPase est une RNase avec une activité exonucléasique 3'-5'. Son rôle est de dégrader les ARN de petite taille générés par les clivages des RNases J1 et Y (Lehnik-Habrink *et al.*, 2012).

#### b) Les hélicases à motif DEAD

Les hélicases à ARN jouent un rôle crucial dans le métabolisme des ARN. Elles forment 6 superfamilles avec les hélicases à ADN et les translocases à ADN et ARN, selon leur séquence en acides aminés. Ces superfamilles regroupent des enzymes eucaryotes comme procaryotes, les hélicases bactériennes appartenant aux superfamilles 2 et 5 (Kaberdin et Bläsi, 2013). De manière générale, les hélicases à ARN sont capables de modifier les structures de l'ARN et les

interactions protéines-ARN. Elles sont ainsi impliquées dans de nombreux processus essentiels pour la cellule tels que le turnover des ARN, la synthèse des ribosomes, la traduction et le métabolisme des petits ARN (Owtrim, 2013; Khemici et Linder, 2016).

Les hélicases à motif DEAD sont une famille d'hélicases monomériques ou dimériques faisant partie de la superfamille 2. Elles possèdent au moins 12 motifs conservés en acides aminés (Figure 11), dont celui qui leur a donné leur nom, le motif DEAD (pour les 4 acides aminés qui le compose : Asp – Glu – Ala – Asp). Ces motifs sont contenus dans deux domaines répétés en tandem, appelés « RecA-like » en raison de leur ressemblance structurelle avec la protéine recombinase RecA. Ces deux domaines forment une crevasse qui correspond au site de liaison de l'ATP (Linder et Jankowsky, 2011). Leurs domaines terminaux (N-terminal, C-terminal ou les deux), déterminent leur spécificité avec leur substrat ARN ou un cofacteur protéique (Cordin *et al.*, 2006; Owtrim, 2013). Contrairement aux autres hélicases, elles ne sont capables de dérouler que des structures ARN ou ADN-ARN courtes (Khemici et Linder, 2016), mais elles sont impliquées dans presque l'ensemble des métabolismes de l'ARN, de leur synthèse à leur dégradation (Cordin *et al.*, 2006).

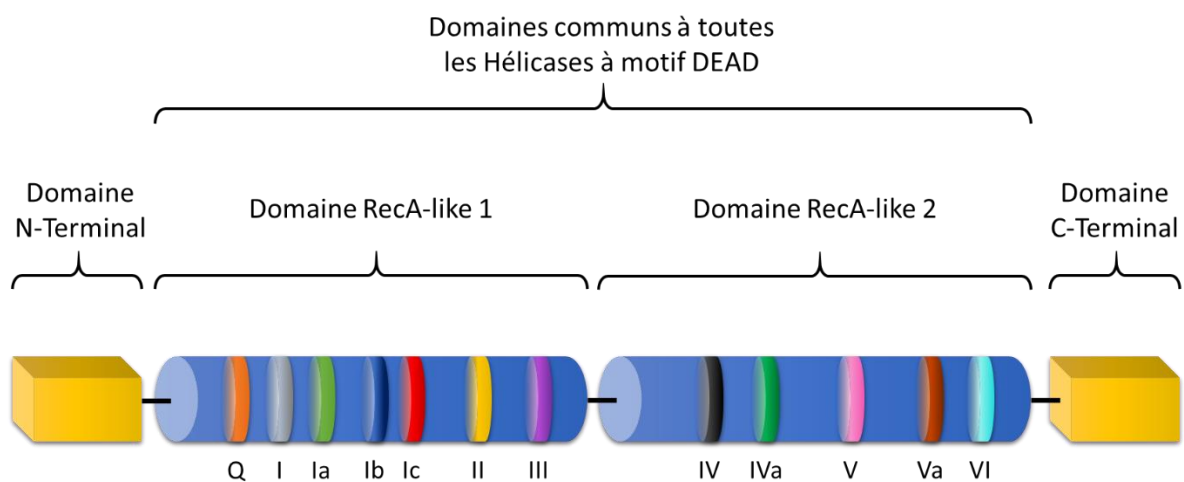


Figure 11 : Les différents domaines et séquences conservées des hélicases à motif DEAD. Ces domaines (au moins 12) sont à des positions conservées et sont caractéristiques de cette famille d'hélicase. Les lettres et chiffres présents sous la figure (Q, I, Ia, Ib, Ic, II, III, IV, IVa, V, Va, VI) correspondent à ces différents domaines. Le motif DEAD est localisé dans le domaine II. (adapté de Owtrim, 2013).

Chez *B. subtilis*, l'hélicase CshA possède les deux domaines RecA-like, ainsi qu'un domaine C-terminal. Ce domaine est essentiel pour la dimérisation de l'hélicase et pour son interaction avec les autres protéines du dégradosome, notamment la RNase Y (Lehnik-Habrink *et al.*, 2010).

De par leur rôle central dans le métabolisme des ARN, les hélicases à motif DEAD sont primordiales pour la croissance bactérienne. Mais il a été également montré qu'elles ont une fonction dans la virulence et la résistance aux stress, en particulier au choc froid. Ainsi la seule hélicase à motif DEAD du pathogène *Helicobacter pylori* est impliquée dans la colonisation chez la souris, et est essentielle pour la croissance à 33°C, mais pas à 37°C. Cela indique une fonction dans la croissance en choc froid (El Mortaji *et al.*, 2018). De même, les hélicases RhIE de *Caulobacter crescentus*, CrhB et CrhC de *Anabena* sp PCC7120 et CsdA de *E. coli* ont un rôle dans la résistance et l'adaptation au choc froid (Resch *et al.*, 2010; Kaberdin et Bläsi, 2013; Aguirre *et al.*, 2017).

## V. Les ARNs régulateurs chez *E. faecalis*

Chez *E. faecalis*, les ARNs régulateurs ont été relativement peu étudiés. Des études globales réalisées ces dernières années ont permis d'identifier près d'une centaine de sARN putatifs (Fouquier d'Hérouel *et al.*, 2011; Shioya *et al.*, 2011; Innocenti *et al.*, 2015), certains surexprimés en conditions de stress, notamment le stress acide ou le stress oxydatif (Fouquier d'Hérouel *et al.*, 2011; Shioya *et al.*, 2011).

Cependant, peu de ces sARN ont été confirmés et caractérisés (listés dans le Tableau 2).

Tableau 2 : Tableau récapitulatif des sARN caractérisés phénotypiquement chez *E. faecalis*

sARN	Localisation	Type	Fonction	Référence
<i>Qa</i>	pCF10	sARN antisens	Régulation de la conjugaison de pCF10	(Weaver, 2007; Michaux, 2013)
<i>mD</i>	pAD1	sARN antisens	Régulation de la conjugaison de pAD1	(Tomita et Clewell, 2000)
<i>ARNII</i>	pAD1	sARN antisens	Système toxine/antitoxine	(Weaver, 2007; Michaux, 2013; Weaver, 2015)
<i>eutX</i>	Chromosome	sARN liant les protéines	Métabolisme de l'éthanolamine	(DebRoy <i>et al.</i> , 2014a)
<i>ef0408-ef0409</i>	Chromosome	sARN antisens	Virulence, colonisation, résistance à la phagocytose, résistance aux stress	(Michaux <i>et al.</i> , 2014a)
<i>ef0605-ef0606</i>	Chromosome	sARN non caractérisé	Virulence, résistance à la phagocytose, résistance au stress oxydatif	(Michaux <i>et al.</i> , 2014a)
<i>ef1368-ef1369</i>	Chromosome	sARN antisens	Virulence, colonisation, résistance à la phagocytose, résistance au stress osmotique, stress acide	(Michaux <i>et al.</i> , 2014a)
<i>ef3314-ef3315</i>	Chromosome	sARN non caractérisé	Virulence, résistance aux sels biliaires	(Michaux <i>et al.</i> , 2014a)

Des gènes codant des sARN situés sur le chromosome ont été montrés comme étant impliqués dans la réponse au stress, notamment le stress acide, le stress oxydatif et la survie en macrophage, mais aussi dans la virulence d'*E. faecalis* (Tableau 2) (Michaux *et al.*, 2014a).

D'autres sont impliqués dans la régulation de voies métaboliques.

Le système le mieux caractérisé est le locus *eut*, composé de 19 gènes permettant de métaboliser l'éthanolamine et régulé par la présence d'éthanolamine et d'adénosyl cobalamine (Del Papa et Perego, 2008 ; DebRoy *et al.*, 2014b) (Figure 12). L'expression du gène codant le petit ARN *eutX* est contrôlée par un riboswitch liant l'adénosyl cobalamine. En présence de ce ligand, la transcription s'arrête et le sARN est inactif. En absence du ligand, la transcription continue et il se forme sur le sARN des structures tige-boucle (P3 et P4) capables de se lier et ainsi d'inhiber l'activité d'une protéine régulatrice, *EutV*. Cette protéine fait partie d'un système à deux composants avec le senseur *EutW*, sensible à l'éthanolamine. L'activité de l'opéron *eut* est

donc dépendante de la présence d'éthanolamine pour activer EutV et d'adénosyl cobalamine pour empêcher la production d'EutX, inhibiteur d'EutV.

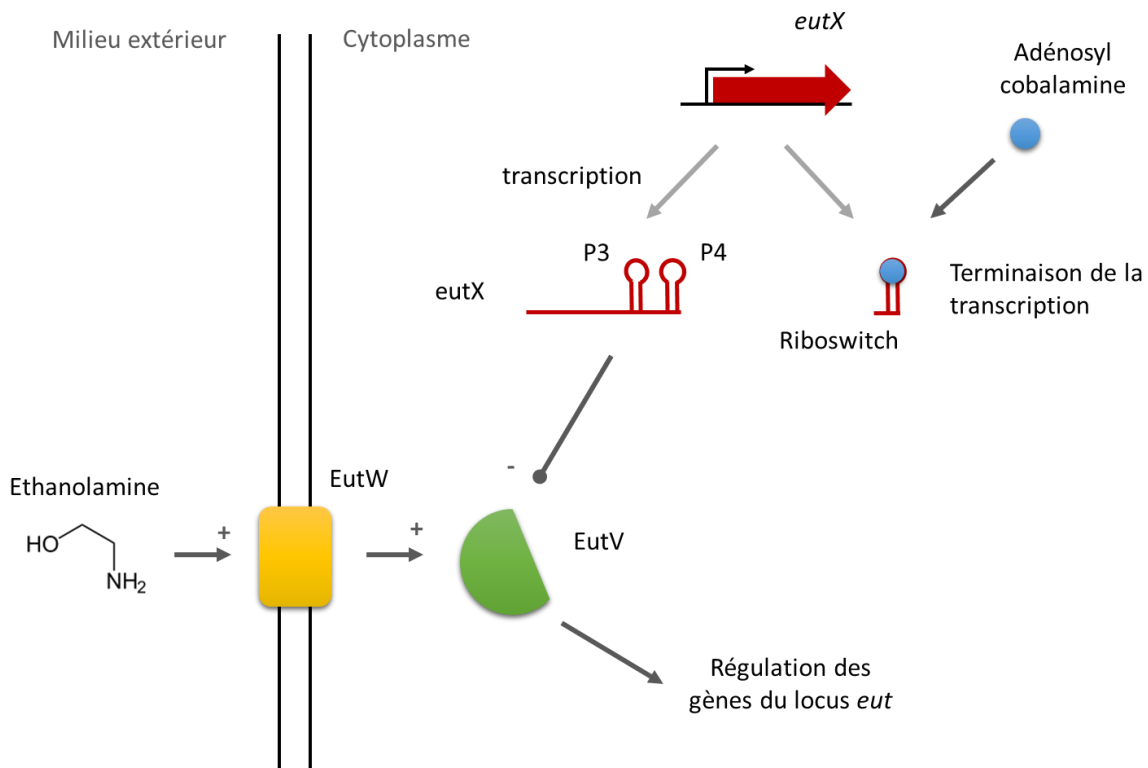
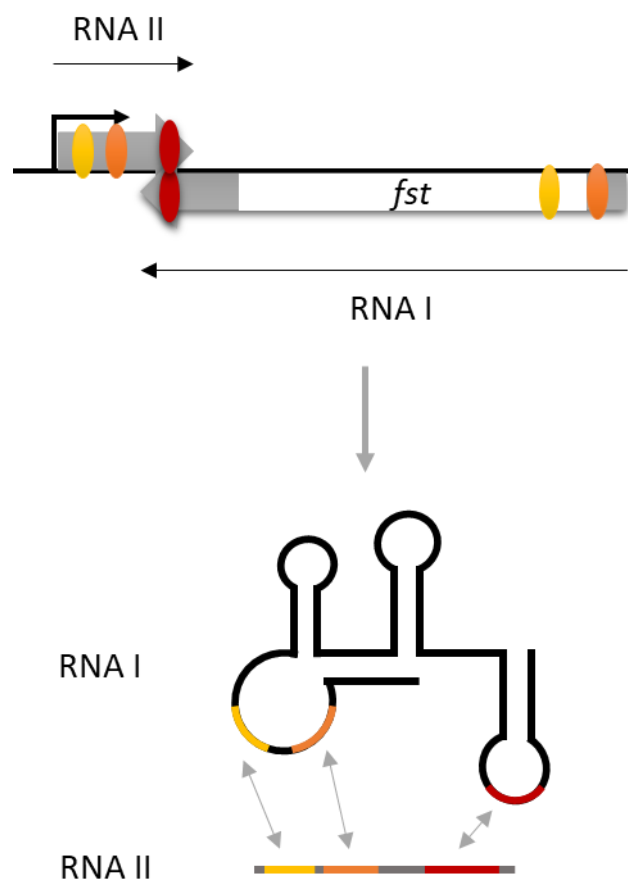


Figure 12 : Schéma explicatif de la régulation du système eut. La protéine régulatrice principale du système EutV (en vert) est elle-même régulée par deux mécanismes : le petit ARN répresseur eutX (en rouge), lui-même réprimé par l'adénosyl cobalamine, et le senseur histidine kinase EutW (en jaune), dépendant de la présence d'éthanolamine.

Des gènes codant pour des sARN impliqués dans la virulence ont aussi été identifiés sur des plasmides conjugatifs. Par exemple, le gène codant le sARN mD est localisé sur le plasmide pAD1 qui porte des gènes codants des hémolysines et des bactériocines. Ce sARN entre dans la régulation de la conjugaison du plasmide (Tomita et Clewell, 2000). Il est en effet induit en l'absence de la phéromone cAD1, et permet de contrôler la formation de terminateurs qui sont situés en amont de gènes codant des protéines régulatrices. Cela donne lieu à un contrôle fin et spécifique de la conjugaison de ce plasmide chez *E. faecalis*. Le sARN Qa dont le gène est porté par le plasmide pCF10 semble fonctionner selon le même principe (Weaver, 2007).

Un autre petit ARN dont le gène est également situé sur pAD1 a été identifié : RNA II, convergent et complémentaire de l'ARNm RNAI (Weaver, 2007; Weaver, 2015). RNA I porte un gène codant une toxine, nommée Fst. Leur extrémité 3' et deux motifs répétés en 5' portés à la fois par RNA I et II sont complémentaires (Figure 13). Une structure en U, présente sur RNAI, permet l'interaction des trois motifs. La liaison RNAII-RNAI au niveau de leurs motifs en 5' bloque la traduction de la toxine Fst. Ce système de répression est similaire à un système toxine-antitoxine.



*Figure 13 : Conformation et association entre les motifs complémentaires du locus RNAI/RNAII sur le plasmide pAD1. Chacun des motifs complémentaires est représenté de la même couleur (jaune, orange, rouge).*

## VI. Contexte et objectifs

L'Unité de Recherche Risques Microbiens (U2RM) est un laboratoire dont la thématique est centrée sur l'interaction des bactéries à Gram positif dont celles du genre *Enterococcus* avec leur environnement. Les études concernent le métabolisme, la résistance aux antibiotiques, la virulence et la réponse aux stress.

Ces dernières décennies, de nombreux petits ARN non codants ont été mis en lumière en tant que régulateurs de la virulence, de la réponse au stress, et d'autres processus impliqués dans la croissance et le fitness bactérien chez de nombreux pathogènes. Cependant, si des études globales ont permis la détection de nombreux sARN régulateurs chez *E. faecalis*, très peu d'entre eux ont été confirmés, et encore moins caractérisés de manière précise chez cette bactérie pathogène.

Lors de cette étude, nous avons commencé par la détection par RNA-seq de petits ARN putatifs potentiellement induits en conditions d'infections. Cette expérience a été initiée dans l'unité U2RM avant ma thèse, et le traitement et l'analyse des résultats ont constitué la première étape de ce travail. Cela a permis d'avoir une vue d'ensemble du transcriptome d'*E. faecalis* en condition d'infection et dans différentes conditions stressantes, et ainsi de sélectionner les meilleurs candidats ARN non codants pour la suite de cette thèse.

Les ARN sélectionnés n'avaient jamais été étudiés au début de mon travail. La deuxième étape a donc été de caractériser ces ARN non codants, au moyen d'études génomiques, protéomiques et physiologiques, et de déterminer leur mode d'action, leur fonction ainsi que celle du (des) gène(s) qu'ils régulent (métabolisme concerné, implication ou non dans la réponse au stress ou la virulence).





# *Matériel et Méthodes*

Cette section présente le matériel et méthodes des expériences réalisées hors publications.

## 1. Souches et conditions de culture

La souche *Enterococcus faecalis* V19 a été utilisée pour cette étude et a été cultivée à 37°C en milieu M17 supplémenté en glucose (0,5%) (GM17) ou en ribose (0,5%) (RM17) pour les tests devant être réalisés en absence de répression catabolique. Pour les conditions stressantes, le milieu a été supplémenté en sels biliaires (0,06%) ou son pH a été ajusté à 5,8. Le milieu a été supplémenté avec du chloramphénicol (10 µg/ml) pour les souches contenant le plasmide pAGent, pLT06 ou pNZ273, et avec de l'agmatine (40mM) pour toutes les souches contenant pAGent. Les souches *E. coli* DH5α ont été cultivées en milieu LB à 37°C. Pour les souches contenant le plasmide pLT06, les cultures ont été incubées à 30°C. *Lactococcus lactis* utilisé pour la réalisation des constructions dans le plasmide du pAGent a été cultivé en GM17 à 30°C.

## 2. Techniques de biologie moléculaire

Les PCR ont été réalisées avec les enzymes Phusion High-Fidelity DNA Polymerase (ThermoFisher, Waltham, Massachusetts, USA) ou GoTaq DNA Polymerase (Promega, Madison, Wisconsin, USA). Le kit 5'/3' RACE kit, 2<sup>nd</sup> generation (Roche, Bâle, Switzerland), a été utilisé suivant les recommandations du fournisseur pour les 5'RACE-PCR. Celles-ci ont été obtenues en ajoutant des extensions poly-A ou poly-C à l'ADN complémentaire et en utilisant les amorces SP1, SP2 et SP3 (Tableau 3). Ce même kit a été utilisé pour les 3'RACE-PCR après avoir ajouté une extension poly-A en 3' des ARN, avec le kit poly(A) polymerase tailing kit (Epicentre, Madison, Wisconsin, USA). Toutes les purifications de produits PCR ont été effectuées à l'aide du kit NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel, Düren, Germany), et toutes les extractions plasmidiques ont été réalisées avec le kit NucleoSpin Plasmid (Macherey-Nagel).

Tableau 3 : Liste des oligonucléotides utilisés dans les parties hors articles.

Nom	Séquence	Utilisation
SRC65_SP1	GGCTCCACCTACTGGCATA	RACE-PCR, Northern Blot
SRC65_SP2	TAGAGCCGTCAGAAATTCAG	RACE-PCR
SRC65_SP3	ACAAACACGTAATTATTGGTGT	RACE-PCR
SRC85_SP1	GAACAATCCCGTTAAACATGTT	RACE-PCR, Northern Blot
SRC85_SP2	ATTGTACACCACATTGTATTAC	RACE-PCR
SRC85_SP3	CACATTGTATTACGAAATTCCTC	RACE-PCR
ef1515R	AGTCCTTCTGGCTCCATCAC	RACE-PCR, RT-PCR
ef1515_SP2	TATTAAACCCGACGCCTTTAC	RACE-PCR
ef1515_SP3	TTCTCCTGCTTTCGTCAAG	RACE-PCR
SRC90_SP1	GGAAACATGCCAAGTATGGAA	RACE-PCR, Northern Blot
SRC90_SP2	ACTTGACACGCATTGACTAATT	RACE-PCR
SRC90_SP3	AACGTTGCCATAAGTTGATGT	RACE-PCR
ef1377-SP1	ATAGTTGGCTTAACTTTATCCA	RACE-PCR
ef1377-SP2	TTCGTTGATAAAAGTTGAAATT	RACE-PCR
ef1009-SP1	TTCAGCAATTTGCATAGCTAATT	RACE-PCR
ef1009-SP2	TTCCAGTTCCTGTCGGCGA	RACE-PCR
ef1009-SP3	AAACATTTTCTCCTTCGCGAA	RACE-PCR
src65_Forward	ACACCAATAATTACGTGTTTGT	Northern Blot
src85_Forward	GAAGAATTTTCGTAATACAATGTG	Northern Blot, RT-PCR
src90_Forward	ACATCAACTTATGGCAACGTT	Northern Blot
ef0160L	GGCAAACAGTCAATGGAGGT	qPCR
ef0160R	TGGCAAACATATGCCCTGTA	qPCR
ef2221L	TTATGAACCGTTACGCAGCA	qPCR
ef2221R	TATCCCAGCCAGCTGTTTT	qPCR
ef3251L	ACAGTACCGCCACAGGAAAC	qPCR
ef3251R	CCGTCTTCTCCACAAAGAA	qPCR
gusA-G	GTGTGGGTCAATAATCAGG	qPCR
gusA-D	GAAATCATGGAAGTAAGACTGC	qPCR
GyrAL	GATGGGGAAATCAGGGATTC	qPCR
GyrAR	TCTTTTCCATTTCGGCATTTC	qPCR
pAGEnt_For	GATAGTAAACTCTTCGCTAAA	Surexpression
pAGEnt_Rev	CGAATATCAGTAATTGCTTTATC	Surexpression
RPV_AGEnt+1	ACATTGACTTTACCAAGTTAG	Surexpression
FRV_AGEnt_Term	AATTACAGCACGTGTTGCTTT	Surexpression
FPI_SRC65_+2	CTAAACTGGTAAAGTCAATGTATCATAATAAAGGAGGTGAAATAT	Surexpression
RPI_SRC65_Term	AAAGCAACACGTGCTGTAATTA AAAACGTAACATCCTTGTAAAT	Surexpression
SRC_65_1	GGATACCGAGCTCAACTCACGAAGACTAATATTATT	Construction du mutant $\Delta$ SRC65
SRC_65_2	TCCACCTACTGGCATAATTTACCTCCTTTATTATGAT	Construction du mutant $\Delta$ SRC65
SRC_65_3	ATGCCAGTAGGTGGAGCC	Construction du mutant $\Delta$ SRC65
SRC_65_4	GGGTTTATCGACAATATTACTTAGCGAACCCAG	Construction du mutant $\Delta$ SRC65
SRC_65_5	TAATGTGTTAGCCTTCCCTG	Construction du mutant $\Delta$ SRC65
SRC_65_6	AAATAGTAGGGTATTTACTTTCT	Construction du mutant $\Delta$ SRC65
SRC_90_1	GGATACCGAGCTCTTATCCTGTAATCGTAAATCAC	Construction du mutant $\Delta$ SRC90
SRC_90_2	ATGTACCTCCTTTAATTGATAAT	Construction du mutant $\Delta$ SRC90
SRC_90_3	TTAAAGGAGGTACATATGGTCTTCCATACTTGCA	Construction du mutant $\Delta$ SRC90
SRC_90_4	GGGTTTATCGACTAACTATGGTAGAAACATCATTA	Construction du mutant $\Delta$ SRC90
SRC_90_5	ATTTGCGTTTCAAACGTGAAAA	Construction du mutant $\Delta$ SRC90
SRC_90_6	TTGCATACCGTTGAGTTAGG	Construction du mutant $\Delta$ SRC90
pLT06_1	GAATTCGAGCTCGGTATCC	Construction des mutants
pLT06_1_bis	GAATTCGAGCTCGGTACC	Construction des mutants
pLT06_2	CTGCAGGTCGATAAACCC	Construction des mutants
ef1518-1b	ATCACCCGGGGCTGATGCTAATGTAACTTCT	Construction du mutant $\Delta$ ef1518
ef1518_2	TAGACTCGAGTTACTCTACGCATCGAATCAT	Construction du mutant $\Delta$ ef1518
ef1518_3	ACTTCTCGAGAAATGGTGGCAATATGTTCTAT	Construction du mutant $\Delta$ ef1518
ef1518-4b	TCATGGATCCAACCTAATGACACCTGTTGTC	Construction du mutant $\Delta$ ef1518
ef1518_5	AATGATGTTGGTGGTTCAAGG	Construction du mutant $\Delta$ ef1518
ef1518_6	TTTACCTTCTGTAATCGTTCC	Construction du mutant $\Delta$ ef1518
1518_Seq_for	AACGTTCAATCTTGCGAAGTT	Construction du mutant $\Delta$ ef1518
1518_Seq_rev	CTTGGCAACTTCTTCTGTACT	Construction du mutant $\Delta$ ef1518

### 3. Construction d'un plasmide permettant la surexpression de SRC65

Le plasmide pAGent (Linares *et al.*, 2014) a été utilisé pour surexprimer SRC65 grâce au promoteur inductible par l'agmatine. Le plasmide a été amplifié par les oligonucléotides RPV\_Agent+1 et FRV\_Agent\_Term, sans inclure le His-Tag et la séquence Shine-Dalgarno. Le fragment à insérer a été amplifié avec les oligonucléotides FPI\_SRC65\_+2 et RPI\_SRC65\_Term qui possèdent des queues flottantes complémentaires au plasmide (Tableau 3). L'ensemble a été assemblé par recombinaison *in vivo*, ou avec le kit NEBuilder HiFi DNA Assembly Cloning Kit (New England BioLabs, Ipswich, Massachusetts, USA) en utilisant *L. lactis* MG1363 comme hôte. Le plasmide a ensuite été extrait et introduit dans *E. faecalis* V19.

### 4. Construction du mutant $\Delta$ SRC65

La souche mutante  $\Delta$ SRC65 a été construite chez *E. faecalis* V19 et en utilisant la souche *E. coli* DH-5 $\alpha$  pour toutes les étapes de clonage. Le plasmide pLT06 a été amplifié avec les amorces pLT06-1 et pLT06-2, et les régions flanquantes de SRC65 ont été amplifiées avec les couples d'oligonucléotides SRC\_65\_1/SRC\_65\_2 et SRC\_65\_3/SRC\_65\_4 (Tableau 3). Le clonage a été réalisé à l'aide du kit NEBuilder HiFi DNA Assembly Cloning Kit (New England BioLabs) et le plasmide obtenu a été utilisé pour transformer *E. coli*. Un fois répliqué, le plasmide a été introduit chez *E. faecalis* par électroporation selon le protocole décrit par Bae *et al.*, 2002 (avec 10 g/ml de lysozyme et mutanolysine 10 U pour le traitement des cellules compétentes) et intégré dans le chromosome par double crossing over, comme décrit auparavant (Thurlow *et al.*, 2009). Le mutant a été vérifié par PCR et par séquençage avec les amorces SRC\_65\_5 et SRC\_65\_6. La procédure utilisée pour la tentative de construction du mutant  $\Delta$ SRC90 a été la même, avec les oligonucléotides SRC\_90\_1, SRC\_90\_2, SRC\_90\_3, SRC\_90\_4, SRC\_90\_5 et SRC\_90\_6.

## 5. Northern blot

Les expériences ont été réalisées avec 10 µg d'ARN total comme décrit antérieurement (Shioya *et al.*, 2011). La sonde utilisée a été synthétisée par PCR avec les couples d'oligonucléotides SRC65\_Forward/SRC65\_SP1 et src90\_Forward/SRC90\_SP1 pour SRC65 et SRC90, respectivement (Tableau 3). Des dATP-[ $\alpha^{32}$ P] ont été incorporés aux sondes lors de l'étape d'élongation de la PCR.

## 6. Extraction d'ARN total

Pour les conditions non stressantes, 10 à 100 ml de cultures en bouillon GM17 ont été incubés à 37°C jusqu'à DO<sub>600</sub> 0,5. Pour les conditions stressantes, les bactéries ont été cultivées jusqu'à DO<sub>600</sub> 0,3, centrifugées et resuspendues dans du GM17 liquide supplémenté en sels biliaries (0,06%) ou avec un pH ajusté à 5,8 et incubées à 37°C pendant 30 minutes. Les cellules ont été centrifugées puis lysées au FastPrep (MP Biomedicals, Illkirch Graffenstaden, France). L'extraction des ARN a été réalisée au moyen d'une séparation avec du TRIzol Reagent (ThermoFisher) et du chloroforme isoamylique. Puis les ARNs ont été purifiés à l'aide du kit Direct-Zol RNA Miniprep kit (Zymo-Research, Irvine, Californie, USA), et quantifiés au Nanodrop™ 2000 (ThermoFisher). Leur qualité a été vérifiée par migration sur un gel d'agarose contenant de la javel, comme précédemment décrit (Aranda *et al.*, 2012).

## 7. RT-PCR et RT-PCR quantitative (RT-qPCR)

Les étapes de rétrotranscription ont été réalisées à l'aide du kit QuantiTect Reverse Transcription kit (Qiagen, Venlo, Pays-Bas). Les PCR pour les expériences de RT-PCR ont été réalisées comme une PCR classique, avec une amorce sens complémentaire du gène amont, et une amorce antisens complémentaire du gène aval (Tableau 3). Le kit GoTaq qPCR Master Mix (Promega, Madison, Wisconsin, USA) a été utilisé pour les expériences de qPCR, avec

l'appareil C1000™ Thermal Cycler (Bio-Rad, Hercules, Californie, USA). Les oligonucléotides utilisés (L/R ou G/D) sont listés dans le tableau 3. Les conditions suivantes ont été utilisées pour toutes les expériences : 95°C pendant 3 minutes, suivies de 40 cycles à 95°C pendant 15 secondes et 60°C pendant 1 minute. La spécificité de l'amplification a été vérifiée grâce à une courbe de fusion faisant suite aux cycles d'amplification. Le gène de ménage *gyrA* (oligonucléotides GyrAL et GyrAR) a servi de gène de référence pour normaliser les niveaux de transcription de chaque gène. L'ADN génomique de la souche V19 a été utilisé pour l'établissement des courbes standards et le calcul de l'efficacité de la qPCR.

## 8. Extraction des protéines totales

Pour chacune des extractions, 50 ml de culture ont été incubées jusqu'à DO<sub>600</sub> 0,5. Les cellules ont été lavées avec du Tampon 2DE (Tris 50 mM pH8.0, 1 mM PMSF) et lysées au FastPrep (MP Biomedicals). Les cellules lysées ont été centrifugées et le surnageant a été récupéré. Les acides nucléiques ont été éliminés en ajoutant 10 µl de Nuclease Mix (GE Healthcare, Chicago, Illinois, USA) par ml de surnageant. L'ensemble a été incubé à température ambiante pendant 45 min. Les protéines ont été précipitées en ajoutant 4 volumes de tampon de précipitation (Acétone 80%, 20% Tris 50 mM pH8.0, DTT 20 mM, TCEP 5mM) et en laissant incuber une nuit à -20°C. Les protéines ont été quantifiées par la méthode de Bradford (Bradford, 1976).

## 9. Expériences d'iTraQ

Ces expériences ont été réalisées à la plateforme PROTEOGEN de l'Université de Caen par Benoit Bernay et Julien Pontin.

- a) Préparation des échantillons, réduction, digestion, marquage iTraq et fractionnement.

Les échantillons ont été réduits et alkylés avant d'être digérés par la trypsine puis le marquage iTraq avec les étiquettes 114 et 117 a été réalisé selon les recommandations du fabricant (AB Sciex, USA) avant que les deux échantillons ne soient combinés. Après marquage, les peptides tryptiques marqués ont été fractionnés, sur la base du point isoélectrique, sur un appareil 3100 OFFGel fractionator selon les recommandations du fabricant (Agilent Technologies, Japan) en utilisant une barette de gel Immobiline DryStrip de 13 cm avec un pH 4-7 (GE Healthcare). Après électrofocalisation, 12 échantillons ont été collectés puis concentrés.

- b) Analyse en spectrométrie de masse

Les extraits peptidiques ont été dans un premier temps concentrés et dessalés sur des microcolonnes  $\mu$ C18 Omix (Varian) avant fractionnement en nanoLC (Prominence, Shimadzu). Les peptides sont concentrés dans un premier temps sur une précolonne Zorbax 5x0.3mm C18 precolumn (Agilent) avant d'être séparés sur une colonne Zorbax 150x75 $\mu$ m C18 (Agilent). Les phases mobiles sont constituées de 0,1% acide trifluoroacétique (A) et 0,1% acide trifluoroacétique ; 99,9% acétonitrile (B). Le gradient appliqué est le suivant : 2% B constant pendant 5 minutes, de 2 à 5% B en 1 min, de 5 à 32% B en 144 min, de 32 à 70% B en 10 min, de 70 à 90% B en 5 min et retour à 2% B en 10min. Le débit est de 300nl/min, auxquels sont ajoutés 1,2 $\mu$ L/min d'une solution contenant 5mg/mL de matrice CHCA diluée dans une solution de 50% ACN ; 0,1% TFA. Les fractions sont déposées toutes les 26 secondes sur une cible Opti-TOF™ 384 par un "spotter" AccuSpot (Shimadzu).

Les analyses en spectrométrie de masse sont réalisées sur un instrument ABSciex 5800 équipé d'une source MALDI, d'un analyseur TOF TOF et d'un laser OptiBeam™. L'appareil est calibré,

avec une précision supérieure à 50 ppm, avant chaque analyse avec un mélange de 5 peptides (Angiotensin I & II, Neurotensin, ACTH (1-17), ACTH (18-39)).

Les analyses sont réalisées en mode automatique avec les paramètres suivants :

- Les spectres MS sont acquis en mode réflecteur positif avec une intensité de laser de l'ordre de 3400 en sommant 1000 spectres indépendants (5 x 200) dans la gamme de masse 700 à 4000 m/z.

- Les spectres de fragmentation sont réalisés avec une intensité de laser de 4300, en mode MSMS réflecteur positif, en sommant 2500 spectres (10 x 250). L'air est utilisé en tant que gaz de collision.

c) Traitement des résultats, recherches en banque

La recherche en banque, pilotée par le logiciel Mascot 2.6.1 (Matrix Science), est réalisée sur la base des fragmentations obtenues en MSMS en utilisant, selon l'étude, deux banques, la première contenant des séquences de contaminants connus (kératines et trypsine, 21 séquences), les séquences protéiques d'*Enterococcus faecalis* V583 (3264 séquences). Les modifications « non fixes » sélectionnées pour l'identification sont les suivantes : oxydation et dioxydation des méthionines, actéylation des lysines et methyl methanethiosulfonate des cystéines. Le paramètre trypsine est sélectionné et 2 erreurs de coupures sont autorisées. Le paramètre quantification est sélectionné sur la base « iTraQ4plex ». L'erreur de masse est fixée, pour la MS et la MSMS, à 300 ppm et 0.6 Da respectivement. Les données Mascot ainsi obtenues sont ensuite transférées dans le logiciel de validation Proline (<http://www.profiroteomics.fr/proline/>) afin d'être filtrées selon un seuil <0,05 et d'éliminer les redondances protéiques. Les données validées sont ensuite analysées par le logiciel Protein Pilot, en utilisant l'algorithme Paragon (Sciex) en sélectionnant les paramètres suivants : digestion tryptique, modification methyl methanethiosulfonate des cystéines, marquage iTRAQ 4-plex, paramètres de correction



automatique, quantification, correction des biais, modification biologique et identification stringente. Le seuil de détection protéique a été fixé à 95% et une analyse de faux positifs a été réalisée.

## 10. Expérience de retard sur gel

Les protéines CcpA, Hpr et HprKV267F ont été extraites comme décrit précédemment (Galinier *et al.*, 1997). La phosphorylation de Hpr sur la sérine 46 (HprP) a été obtenue par l'action de HprKV267F comme précédemment décrit (Monedero *et al.*, 2001). CcpA et HprP ont été diluées en séries, mises en contact avec les ADN test et témoins dans du tampon de liaison (Tris pH 7,5 40 mM, DTT 4 mM, EDTA 4 mM, KCl 200 mM, MgCl<sub>2</sub> 80 mM, Fructose-1,6 biphosphate 10 mM, Glycerol 20 %, ADN de saumon 2 mM, BSA 0.04%) et incubées 15 minutes à température ambiante. Les mélanges ont ensuite été mis à migrer en gel de polyacrylamide en condition native dans du tampon TBE 0,5 X. Un fragment d'ADN du gène *gusA* d'*E. coli* a été utilisé comme témoin négatif, et la séquence promotrice du gène *ef0292*, contenant une boîte *cre* comme témoin positif.

## 11. Détermination de la CMI et de la CMB

Du milieu RM17 a été distribué dans une microplaque 96 puits (Bio-Rad) à la moitié du volume final, puis chaque antibiotique de concentration 4X a été ajouté au premier puits puis dilué en série d'un facteur 2. Les cultures de nuit ont été diluées pour obtenir une DO<sub>600</sub> de 0,02 dans du milieu RM17 et incubées à 37°C jusqu'à DO<sub>600</sub> environ 0,2, puis ont été diluées au 1/100<sup>ème</sup> et distribuées (moitié du volume final) pour obtenir environ 10<sup>6</sup> UFC/ml par puits. Les microplaques ont ensuite été incubées 24 h à 37°C. La quantité d'inoculum a été vérifiée par dénombrement sur boîtes de M17 glucosé (incubées 24h à 37°C). La concentration minimale inhibitrice (CMI) correspond à la concentration en antibiotique la plus basse ayant inhibé la croissance bactérienne. Pour déterminer la concentration minimale bactéricide (CMB), 100 µl

de chaque puits où la croissance bactérienne est inhibée ont été étalés sur boîte de GM17 gélosé. Les boîtes ont été incubées à 37°C pendant 24h et le nombre d'UFC/ml a été déterminé. La CMB correspond à la concentration la plus faible pour laquelle on observe une baisse du nombre d'UFC de  $10^3$  par rapport à l'inoculum de départ.

## 12. Microscopie électronique à balayage

Les expériences ont été réalisées par le Dr. Didier Goux à la plateforme CMAbio<sup>3</sup> d'ICORE à l'Université de Caen Normandie.

## 13. Analyses bio-informatiques

Les recherches d'homologies de séquences nucléotidiques et protéiques ont été faites à l'aide de l'outil BLAST (Madden, 2003). Les recherches de terminateurs de transcription ont été effectuées avec l'outil FindTerm, SoftBerry (Li, 2011) avec un seuil d'énergie de -11. Les alignements de séquences ont été réalisés à l'aide du logiciel Multalin (Corpet, 1988). Les outils Phobius (Käll *et al.*, 2004) et InterPro (Mitchell *et al.*, 2018) ont été utilisés pour la recherche de domaines caractéristiques dans une séquence en acides aminés. Pour prédire les cibles potentielles d'un sARN, le logiciel TargetRNA2 (Kery *et al.*, 2014) a été utilisé, avec une recherche localisée à 80 nucléotides en amont et 20 nucléotides en aval du +1 de transcription des gènes cibles. La base de données DOOR (Mao *et al.*, 2009) qui référence les structures opéroniques chez les bactéries a servi à détecter d'éventuelles co-transcriptions entre les gènes étudiés.

# *Résultats*

## **Chapitre 1 : Identification de petits ARN putatifs induits in vivo et en conditions de stress**

Cette première partie a été écrite sous la forme d'un article actuellement soumis pour publication dans le journal FEMS Microbiology Letters. La détection d'ARN régulateurs non codants dans le génome d'*Enterococcus faecalis*, ainsi que la mise en lumière de leur expression *in vivo* et en conditions stressantes grâce à une approche transcriptomique constituent le point de départ de cette thèse. Les études qui suivront concernent des ARN régulateurs détectés lors de ce travail.

**Identification of the general stress stimulon related to colonization in  
*Enterococcus faecalis***

Marine Salze<sup>a</sup>, Jean-Christophe Giard<sup>a</sup>, Eliette Riboulet-Bisson<sup>a</sup>, Torsten Hain<sup>b</sup>, Alain Rincé<sup>a</sup>  
and Cécile Muller<sup>a,\*</sup>

<sup>a</sup>Normandie Univ, UNICAEN, U2RM, Caen, France

<sup>b</sup>Institute of Medicine Microbiology, Gießen University, 35392 Gießen, Germany

\*Corresponding author: [cecile.muller@unicaen.fr](mailto:cecile.muller@unicaen.fr)

**Key words:** *Enterococcus faecalis*, transcriptome, stress response, colonization, small RNA

## **Abstract**

*Enterococcus faecalis* is an opportunistic pathogen responsible for nosocomial infections which have to cope with major stress factors during colonization. To better understand the effects of stress encountered during infection and colonization by *E. faecalis*, the present study assessed the transcriptomic response of the bacteria facing exposure to serum, urine, bile salts, acid pH or oxidative stress. Compared to non-stressed culture in exponential growth phase, 30% of the *E. faecalis* genes were differentially expressed, that may represent the infection stimulon. The transcriptome analysis reveals common but also specific responses, depending on the stress encountered. Urine exposure has the most important impact on mRNA abundance, and the highest number of genes with modified expression encode proteins involved in transport and metabolic pathways. The results also pinpoint many stress-related sRNA not yet characterized or intergenic regions that require further investigations to precisely characterize their cellular function. This study allowed the identification of the general stress stimulon related to infection. Indeed, *E. faecalis* confronted to stress potentially encountered during host infection induces a stress response more than expression of virulence factors. Thus when the bacterium shifts from commensal to pathogen, it increases its ability to survive to rough conditions allowing it to colonize its host, and becomes this redoubtable opportunistic pathogen that thrives in hospitals.

## Introduction

*Enterococcus faecalis* is a Gram positive bacterium, facultative anaerobic and common resident of mammals' gastrointestinal tract. It is part of the microorganisms used as indicator of faecal contamination in environment, water and food (Ostrolenk *et al.*, 1947). This bacterium has a very wide genome plasticity: studies showed that *E. faecalis* strain V583 genome contains more than a quarter of mobile and foreign DNA, like a pathogenicity island (PAI) (Paulsen *et al.*, 2003) or prophages (Matos *et al.*, 2013). Its ability to grow in hostile environment and its high resistance to several antibiotics (Arias and Murray, 2012; O'Driscoll and Crank, 2015; van Harten *et al.*, 2017) make it an important cause of hospital-acquired infections (urinary tract infection, bloodstream infection, endocarditis...) especially in immunocompromised patients (Arias and Murray, 2012; Babady, 2016; Murray, 1990).

Thus, *E. faecalis* is a model for opportunistic pathogens, through several factors that contribute to enterococcal virulence (Garsin *et al.*, 2014), like aggregation substance, hemolysins, gelatinase, cytolysin, general stress protein 24 (Gls24) (Madsen *et al.*, 2017) or molecules exposed on the surface of the cell like EfaA and SalB antigen (Xu *et al.*, 1997). However, opportunism mainly relies on the ability to cope with stresses or to grow in a wide range of environmental conditions, thanks to stress response proteins, metabolism adaptation or nutrient acquisition, that allow bacteria to colonize and invade their host. These mechanisms can be investigated by studying their expression with global transcriptomic studies to obtain an overall vision of all these characteristics. *E. faecalis* can replicate in presence of bile salts (Rincé *et al.*, 2003), and grow in urine (Vebø *et al.*, 2010) that bacteria encounter in the intestinal tract or during urinary tract infection. This bacterium can also translocate across the intestinal tract through intestinal epithelial cells and causes systemic infection (Wells *et al.*, 1990). The *E. faecalis* abilities to survive in serum and blood (Nallapareddy and Murray, 2008; Vebø *et al.*, 2009) allow it to invade the entire host body and to generate persistent bloodstream infection.

It was also shown that serum resistance is a trait associated with strains that cause bacteremia (Johnson, 1991). Despite the fact that enterococci are considered mainly as extracellular pathogen, they have the ability to persist and survive within macrophages, fully armed with destructive antimicrobial features, like acid pH or production of reactive oxygen species (Flannagan *et al.*, 2009).

The global transcriptomic response study includes the expression of non-coding RNA. These molecules are generally post-transcriptional regulators that allow bacteria to adapt rapidly and subtly to hostile environmental conditions (Oliva *et al.*, 2015). They include riboswitches that are untranslated region in 5' of the coding sequence (5'UTR) capable of binding ligand to fold or unfold secondary structures, small RNA (sRNA) binding proteins able to modulate their activity, CRISPRs (clustered regulatory interspaced short palindromic repeats), and mRNA binding sRNA (Michaux *et al.*, 2014a; Waters and Storz, 2009). This last type of sRNA includes the *cis*- and the *trans*-encoded sRNA that act through a base-pairing mechanism between the sRNA and its target, leading to a change in the efficiency of the translation or the degradation of the double-stranded RNA (Oliva *et al.*, 2015; Waters and Storz, 2009). Experimental studies showed that enterococci sRNA are involved in numerous cellular mechanisms like metabolism (carbon, amino-acids, iron...), quorum sensing, biofilm formation, stress response, pathogenesis (Michaux *et al.*, 2014b; Shioya *et al.*, 2011) and antibiotic resistance (Sinel *et al.*, 2017).

The inherent abilities of *E. faecalis* to adapt rapidly (notably thanks to sRNA) and to survive to stress make it a redoubtable opportunism pathogen. The aim of this study was to establish the transcriptomic response of *E. faecalis* in order to determine how this bacterium adapts to its environment during its commensal lifestyle or during colonization, and we focused on stressors relevant to the human host, like body fluids (urine, serum) or stressors mimicking *in vivo* scenario (bile salts (BS), acid pH or oxidative stress).



## **Materials and Methods**

### ***Bacterial strain and growth media***

*E. faecalis* V19 (plasmid-cured V583 strain) was cultured in BHI medium until exponential growth phase ( $OD_{600nm}=0.6$ ). This culture was divided in order to have the same conditions right before stress exposure. After centrifugation, the cell pellets were re-suspended in fresh BHI medium (control), BHI containing 0.06% Bile Salts (Sigma, Saint-Louis, MO, USA) (BS), BHI set at pH5.8 (acid), BHI with 1 mM  $H_2O_2$  (Sigma) (oxidative), horse serum (serum), or filtered human urine from 3 different volunteers (urine). For the urine samples, written consent from all volunteers was obtained and they were informed of the goal of this study. No health information was collected and no analysis of the samples was performed. Cultures were incubated at 37°C during 15 min and then cells were harvested by centrifugation at 6000g during 10 min at 4°C supernatant were eliminated and pellets were stored at -80°C prior to RNA isolation.

### ***RNA-extraction and sequencing***

Pellets were resuspended in TE (10 mM Tris pH8.0, 1 mM EDTA). Suspensions were added to 2 volumes of acid phenol (Ambion Austin, TX, USA) and 1 volume of glass beads, and then agitated during 30 min at 30 Hz using a MM200 system (Retsch GmbH, Haan, Germany). After centrifugation, aqueous phase was added to 1 volume of TRIZol (Invitrogen, Carlsbad, CA, USA), mixed and incubated 5 min at room temperature. After centrifugation, one volume of ethanol was added to the supernatant and RNA was purified using Direct-Zol RNA Miniprep (Zymo Research, Irvine, CA, USA) according to the manufacturer's instructions.

RNA integrity was checked on a denaturing agarose gel. RNA quality and quantity were determined using a NanoDrop spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). RNA-Seq assays were performed as described by Muller *et al.* (Muller *et al.*, 2015).

***Reads processing, transcriptome de novo assembly, annotation***

Raw sequencing reads were trimmed by removing low quality bases and adapter sequences, as well as sequences shorter than 50 base pairs. To determine gene expression levels, the CLC Genomics Workbench v.10.0 software (<http://www.clcbio.com/genomics/>) was used to align reads on *E. faecalis* V583 reference transcriptome (AE016830). RNA-Seq analysis was carried out for sequence reads obtained from control and stressed conditions using default parameters as previously described (Aedo *et al.*, 2015). Gene expressions were represented by the logCPM (Counts Per Million) values based on the TMM (Trimmed Mean of M values) normalization.

RNA-Seq data from this study are available in NCBI SRA repository, reference number PRJNA507568.

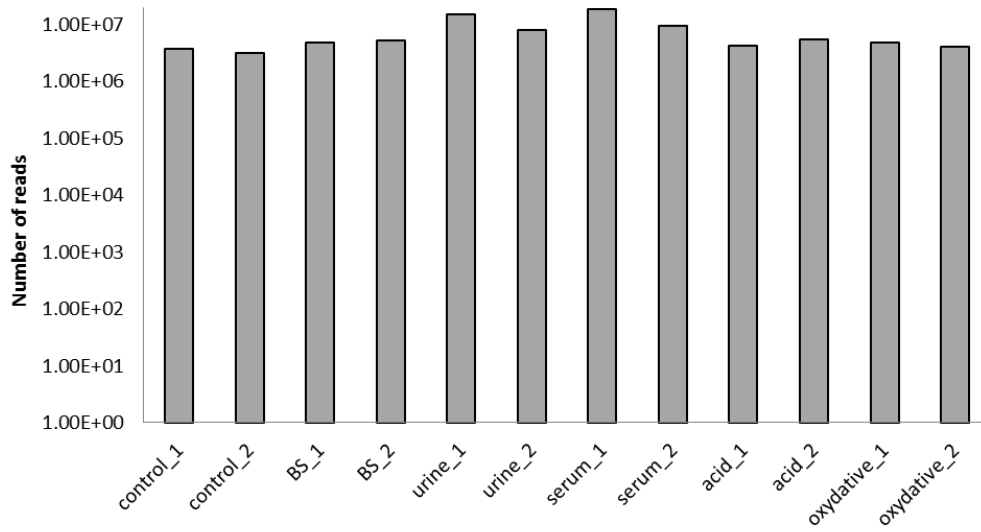
## Results and discussion

### *Global transcriptional analysis*

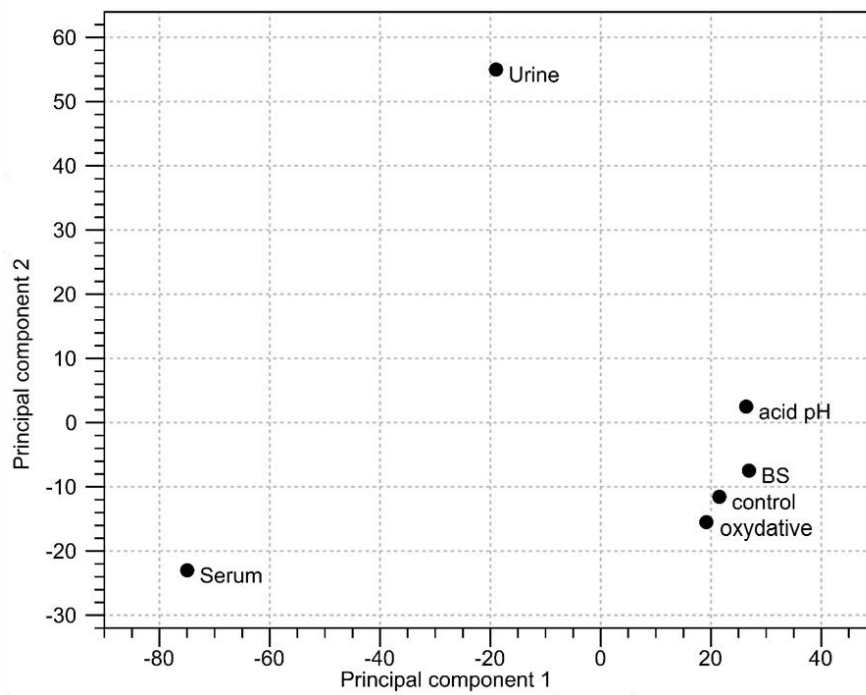
Individual stress conditions were chosen to be representative of stresses encountered by *E. faecalis* during colonization. We studied the transcriptomic response of *E. faecalis* V19 after 15 min incubation in media mimicking *in vivo* conditions encountered during its commensal life or during infection, *i.e.* in urine or serum, or in presence of bile salts (0.06%). We also tested stresses encountered in phagosome like acid pH or oxidative stress using BHI medium adjusted to pH 5.8, or containing 1 mM H<sub>2</sub>O<sub>2</sub>. All these stress conditions were previously studied and no growth alteration of *E. faecalis* was observed (Le Breton *et al.*, 2003).

Gene expression under these conditions was compared to that of a control corresponding to V19 strain in exponential growth phase (OD = 0.6) and harvested after 15 min of incubation at 37°C in BHI medium. RNA-sequencing gave a total of  $8.7 \times 10^7$  reads and *de novo* assembly mapped  $8.4 \times 10^7$  reads on the genome of *E. faecalis* V583 (96% of the reads). Figure 1A gives the size of the libraries that varies between  $3.2 \times 10^6$  and  $1.8 \times 10^7$  reads per condition. The transcriptomes of the bacteria exposed to urine or serum were highly different from that obtained for the conditions tested in BHI medium, like exponential growth phase (control), acid pH, oxidative stress or bile salts (BS) (Figure 1B). Table 1 lists the counts of affected genes according to their functional category (fold-change FC > 5, p < 0.05), which represents 997 genes encoding proteins or 30% of the genome. Figure 2 shows the distribution of differentially expressed genes for each stress condition. The number of genes that showed increased or decreased expression ranged from 327 for the BS stress that had the least impact to 526 for bacterial cells exposed to urine (Table 1 and Table S1).

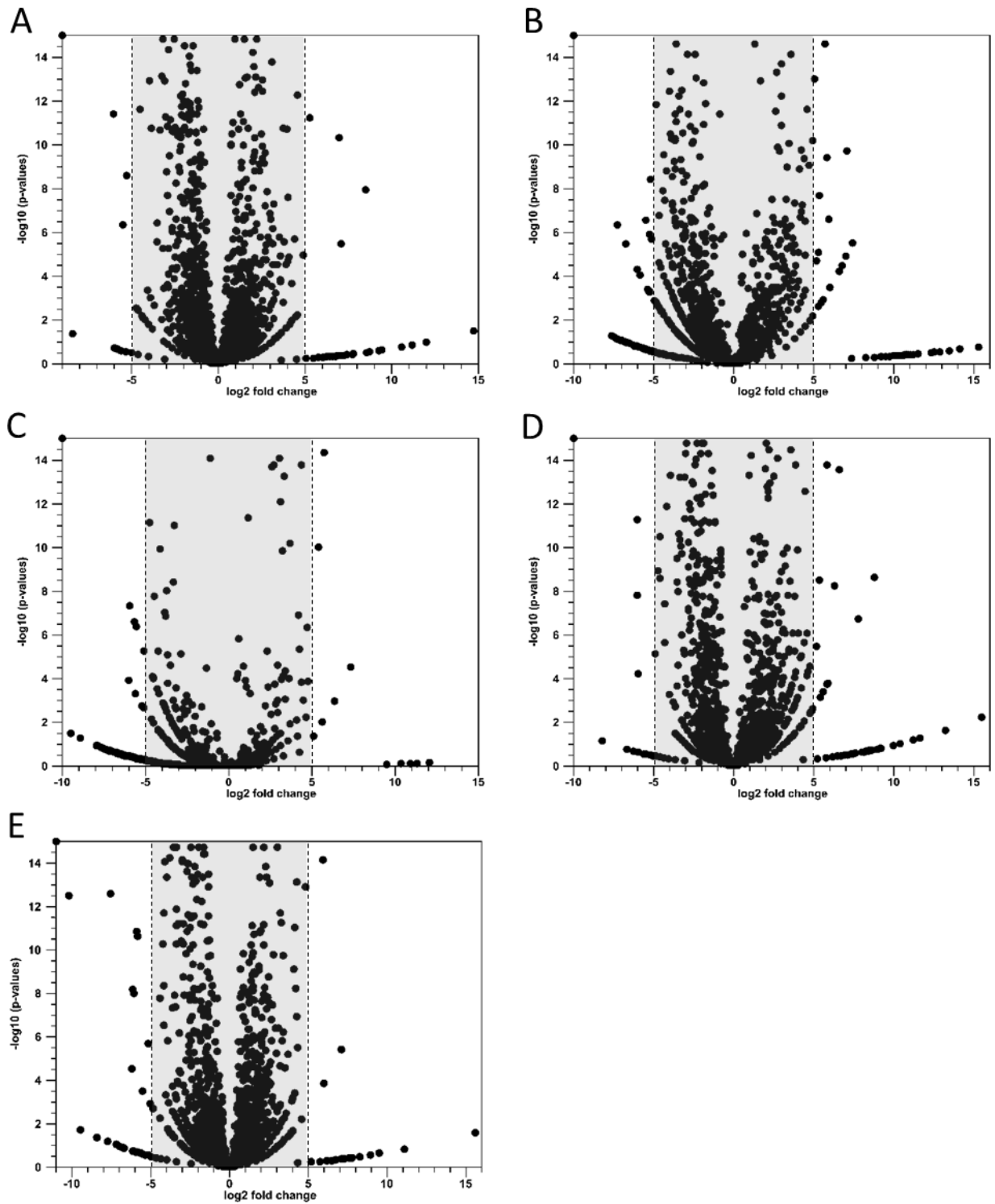
A



B



**Figure 1:** Global analysis of the RNA-Seq data (CLC Genomics Workbench). (A) Total number of RNA-Seq reads per sample. (B) Principal Component Analysis (PCA) plot of the culture conditions used in this study.



**Figure 2:** Volcano plots of the stress conditions displaying significance vs. expression change. Outside the gray areas: genes with FC >5 and  $p < 0.05$ . (A) BS, (B) urine, (C) serum, (D) acid pH, (E) oxidative stress.

**Table 1:** Numbers of differentially expressed genes after 15 min exposure to the stress encountered during infection by *E. faecalis* (FC>5, p<0.05).

	BS		Urine		Serum		Acid		Oxidative	
	up	down	up	down	up	down	up	down	up	down
Amino-acid biosynthesis	3	10	3	7	10	1	10	9	11	9
Cell envelope	12	12	14	25	8	18	10	19	7	12
Cellular process	8	5	9	17	6	9	17	10	7	12
Central and intermediary metabolism	27	15	39	23	32	12	34	15	22	20
DNA metabolism	3	3	4	2	1	3	9	3	6	1
Fatty acid/phospholipid metabolism	0	5	0	21	0	11	9	8	8	8
Mobile/extrachromosomal	10	11	30	13	17	19	27	5	4	19
Other metabolism	1	1	2	5	2	2	3	3	1	1
Proteins fate	7	6	9	5	9	2	7	3	6	3
Proteins synthesis	2	14	7	46	1	38	9	33	24	5
Purines, pyrimidines, nucleosides, nucleotides	1	13	14	26	18	15	3	16	3	1
Regulatory functions	9	11	19	15	6	7	7	9	4	16
Signal transduction	2	12	9	9	4	12	10	9	9	14
Transport	41	43	58	50	50	30	44	39	40	27
Unknown/Hypothetical proteins	22	18	27	18	11	15	41	11	20	29
<b>Total</b>	<b>148</b>	<b>179</b>	<b>244</b>	<b>282</b>	<b>175</b>	<b>194</b>	<b>240</b>	<b>192</b>	<b>172</b>	<b>177</b>

A common response was observed when we cross-checked transcriptomes (Table 2). Thus, almost 40% of the genes with higher or lower expression during urine exposure were common with those of serum or acid stress exposure (a total of 218 and 191 genes, respectively). In the same way, response to BS, acid and oxidative stresses are closer with almost half of the genes common between acid or oxidative conditions and BS exposure (154 and 150 genes, respectively). In contrast, serum and oxidative stress conditions transcriptomes shared less similarities (81 genes).

**Table 2:** Numbers of genes common between the different stresses studied. The total number of genes are indicated and in parenthesis the number of genes with higher (+) or lower (-) expression compared to the control condition.

	<b>BS</b>	<b>Urine</b>	<b>Serum</b>	<b>Acid</b>	<b>Oxidative</b>
<b>BS</b>	327 (148+,179-)				
<b>Urine</b>	137 (47+,90-)	526 (244+,282-)			
<b>Serum</b>	109 (44+,65-)	218 (87+,131-)	369 (175+,194-)		
<b>Acid</b>	154 (70+,84-)	191 (72+,119-)	140 (52+,88-)	432 (240+,192-)	
<b>Oxidative</b>	150 (59+,91-)	102 (28+,74-)	81 (30+,51-)	105 (50+,55-)	349 (172+,177-)

This analysis enforces the previous transcriptomic study on *E. faecalis* during mice peritoneum infection (Muller *et al.*, 2015). If all stresses tested were considered, 51% and 64% of the differentially expressed genes were common to this previous study.

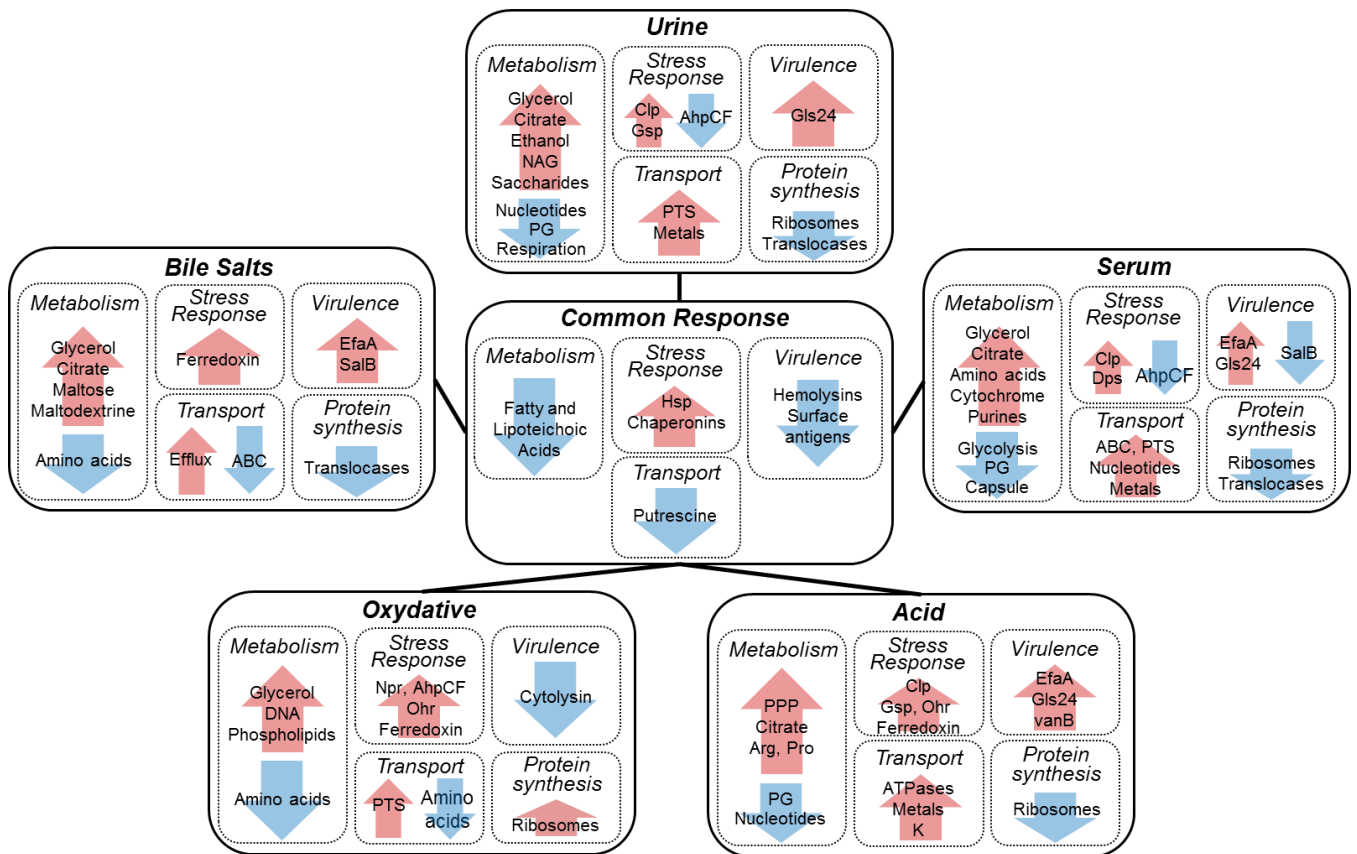
### ***Identification of the general stress stimulon related to infection***

The RNA-Seq data revealed changed expression in most functional gene categories for all 5 conditions compared to exponential growth phase in BHI medium (Table S1).

Genes encoding hypothetical or unknown proteins constituted a large part of genes with differential expression including 126 core-genome ORFs.

It is interesting to note that categories that undergo the most important changes during stress exposures were transport and metabolism, especially for urine and serum that may represent the most drastic conditions for nutrients recovery (Table 1). Some categories were more affected by a specific stress, summarized in Figure 3. For example, many genes of the protein synthesis group (translation or ribosomal structure) were less expressed in all conditions except

for the oxidative stress group that showed only 5 genes with lower expression and 24 genes with higher expression compared to the control.



**Figure 3:** Simplified summary figure of the transcriptomic study. The principal pathways that undergo changes in their expression are indicated. Regulation, mobile elements and hypothetical proteins were excluded for more clarity. Hsp: Heat shock proteins, Gsp: General and universal stress protein, PPP: Pentose Phosphate Pathway, PG: Peptidoglycan.

*Fermentation and citrate metabolism are the preferred metabolic pathways during stress exposure.*

A number of genes showed pattern expression similarities between the different stresses, *e.g.* genes encoding proteins involved in fermentation. Indeed, for almost all stress conditions (BS, serum, urine, oxidative stress), glycerol assimilation genes were highly induced (Figure 3 and Table S1). This issue is in agreement with results previously observed in urine (Ipe *et al.*, 2016;



Vebø *et al.*, 2010), or BS in *Lactobacillus* (Lv *et al.*, 2017). *E. faecalis* uses two pathways for glycerol metabolism: the dihydroxyacetone pathway (DHA) and the GlpK pathway (Bizzini *et al.*, 2010). The DHA pathway allows oxidation of glycerol to DHA by glycerol dehydrogenase GldA that reduces one molecule of NAD<sup>+</sup>, and then DHA kinase DhaKL phosphorylates DHA in DHA-P. For the second pathway, glycerol kinase GlpK phosphorylates glycerol by consuming ATP, and then glycerol-3-P oxidase GlpO oxidizes it to DHA-P by generating one molecule of H<sub>2</sub>O<sub>2</sub>. Thus, the no energy consuming DhaK pathway is used preferentially during stress exposure. DHA-P phosphate is then metabolized by methylglyoxal synthase (EF0939) whose corresponding gene was overexpressed in the same conditions, giving fermented products as 1,2 propanediol (Saxena *et al.*, 2010). Inversely, genes of the GlpK pathway and glycerol transport were less expressed in BS and oxidative stress conditions.

Citrate metabolism allows the synthesis of pyruvate using a pathway independent from glycolysis, through citrate lyase CitDEF that synthesizes oxaloacetate and acetyl CoA. The *citCDEFXM* cluster expression increased during BS stress (5.1 fold) and more highly after serum, urine and acid pH exposure (19.3, 30.8, and 10.5 fold, respectively). It was the same for the citrate transporter CitH (13.5 and 136.2 fold for acid and urine, respectively), indicating that this metabolism is probably promoted during colonization (Figure 3 and Table S1). Moreover, glucose uptake was less expressed in almost all conditions (except urine). There were important changes in the expression of genes involved in glycolysis, pentose phosphate pathway, or metabolic pathways of carbohydrates other than glucose (maltose, galactose, maltodextrine, N-acetylglucosamine), specifically depending on stress.

*Cell surface molecules undergo high expression alteration.*

Due to their spatial location, bacterial surface proteins are naturally involved in complex interactions of bacterial cells with host tissues. Some of these interactions, and proteins or other molecules involved in these interactions, might be essential during the colonization stage, invasion, or spread to other host tissues. Changes in gene expression encoding 36 proteins thought to be bound to the cell membrane or exposed on the cell surface, like membrane proteins and lipoproteins were observed in all conditions. Among them, the *dltXABCD* operon (EF2750-EF2746), which is responsible for lipoteichoic acid D-alanylation, was shown to be less expressed in all conditions (Figure 3 and Table S1). This repression (or the absence of induction) could contribute to the autolysis process and cell wall remodeling in the different stress conditions (Steen *et al.*, 2005). Different external stressors also trigger modification in the fatty acid composition in the cell membrane. Actors identified as involved in lipid, isoprenoid or phospholipid metabolism had lower expression in most of our stress conditions, especially *fabZ1-fabF1*, *ispE* or *pgsA*, suggesting that *E. faecalis* adjusts its fatty acid composition and membrane fluidity (Figure 3 and Table S1).

It has also to be noted the reduction in all conditions of the expression of the ABC transporter EF2652-EF2649, annotated as a spermidine/putrescine transporter (Figure 3 and Table S1). The putrescine is involved in a variety of cellular processes such as cell growth, stress response or pathogenesis in *Salmonella enterica* serovar *Typhimurium* (Gevrekci, 2017), and its intracellular concentration has to be strictly regulated. The low expression of the transporter could suggest that bacterial cell synthesizes polyamines during stress exposure.

*E. faecalis* prioritizes stress response than virulence factors expression when it is confronted to host stress.

Genes that encode virulence factors were also highly differentially expressed, but surprisingly mostly downregulated such as *salB* and *salA* antigens involved in biofilm formation (Mohamed *et al.*, 2006), or *hlyIII*, *tlyA* and EF0700 encoding hemolysins (Figure 3 and Table S1). Stresses encountered by *E. faecalis* during infection also have an important impact on stress-related genes, like those encoding heat-shock proteins *hrcA-grpE-dnaK-dnaJ* or *groEL-groES* that have an induced expression in at least 4 conditions, and mostly in urine. Consequently, it appears that *E. faecalis* may adapt its transcriptome to colonize its host, rather than to increase its virulence.

*Transcription regulators expression is highly impacted by stress related to colonization.*

Modulation of gene expression is a fundamental adaptation response to varying extracellular environment, and is mandatory for bacteria to replicate and survive. Transcriptional regulators or signal transduction systems coordinate genes expression to allow cell to adapt rapidly. A total of 57 transcriptional regulators were found to have a modified expression in all conditions such as HypR (EF2958, +7.3 fold) or SoxS (EF2063, -9.4 fold) involved in H<sub>2</sub>O<sub>2</sub> adaptation.

Genes encoding 10 of the 18 TCS present in *E. faecalis* V583 (Hancock and Perego, 2002) were also found to be significantly regulated, like *croRS* that is upregulated during serum exposure (EF3289-EF3290, +7.0), and was previously shown to be involved in virulence and antibiotic resistance (Comenge *et al.*, 2003; Kellogg *et al.*, 2017; Muller *et al.*, 2018, 2006), or *err10-ehk10* that was less expressed in all tested conditions except acid stress (Muller *et al.*, 2008).

*Stress encountered during host colonization influence mobile elements of E. faecalis genome.*

Most of the genes differentially expressed during host encountered stresses were found into the core genome. The V19 strain derives from V583 strain and harbors an incomplete pathogenicity island (PAI) (Shankar *et al.*, 2002). Genes within the PAI were expressed in all stress conditions tested here but did not exhibit transcriptional variation. However, each stress exposure appeared to significantly regulate genes of the 7 prophages found in the V583 genome (Matos *et al.*, 2013) (Table S1). Some prophages had a specific expression profile depending on stress. For example, more than 30% of the pp5 prophage genes showed more transcripts as a consequence of urine exposure. In the same way, transposases had stress-related altered expression, like during culture in serum where we detected less mRNA for 16 transposases (with fold change (FC) between -5.7 to -90.6 compared to our reference).

*Stress encountered during infection induced a specific response*

*Bile salts stress response involved specific metabolisms.*

Bacteria encountered very arduous conditions during infection, like in the gastrointestinal tract, where they are confronted to bile salts that facilitate digestion through emulsification of fatty food. This biological detergent induces membrane and nucleic acids damages, protein misfolding or denaturation, and osmotic stress. Bile salts can be considered as a signal of entering the small intestine (duodenum), and are found at a concentration between 0.05 to 2%. Some observed changes in gene expression correlated with previous studies (Lv *et al.*, 2017; Solheim *et al.*, 2007). For example, among the highly differentially expressed genes (FC>10), we found that genes related to metabolism and transport of maltose but also of maltodextrine were highly expressed suggesting that these carbohydrates were used preferentially by the cell during BS stress (Figure 3 and Table S1). We also found genes specifically induced by BS, like

that encoding the efflux pump EF0420 (+15.6 fold), or several cell membrane-associated proteins (EF0673, EF2794, EF2795, EF2929, EF3245). The secreted antigen SalB expression was induced (EF0394, +16.6 fold) as it was previously shown (Muller *et al.*, 2006), as well as the adhesion lipoprotein EF0577 (+5.5 fold) and the ferredoxin EF1543 (+8.7).

*Urine exposure had the most impact on E. faecalis transcriptome.*

In its quality of opportunistic pathogen, *E. faecalis* is able to colonize the urinary tract, which in turn is a common source of origin for bloodstream infections. During urinary infection, *E. faecalis* is confronted to several antimicrobial characteristics of urine: low pH (around 6.0), hypertonicity, high urea concentration, iron limitation, and is starvation of several nutrients like amino acids, fatty acids, citrate or sucrose. The urine stress condition displayed the larger transcriptomic changes, especially for gene categories involved in membrane or cell wall metabolism/constitution (cell envelope, lipids, transport, and stress response). As expected, the *E. faecalis* transcriptomes in urine previously published (Vebø *et al.*, 2010) for MMH, OG1RF and Symbioflor 1 strains showed a lot of similarities with our RNA-Seq study, especially for genes linked to metabolism (citrate, ethanol, N-acetylglucosamine, xylose or galactose utilization), and stress response (Gls24 or Clp protease) which were more strongly expressed, or for genes of peptidoglycan synthesis which were less expressed (EF0992-EF0994) (Figure 3 and Table S1). Nevertheless, we showed in our V19 strain that exposure to urine causes additional high transcriptional responses corresponding to strong modifications in several gene categories such as transport with 58 genes highly expressed, notably metal transport like iron (EF1641-EF1639, +10.8), manganese (EF1057, +28.1), copper (EF0298-EF0299, +34.2), or cadmium (EF1400, +12.6). *E. faecalis* adapts the anabolism of proteins, fatty acids or nucleic acid bases and nucleotides to stress conditions, as evidence the lower expression in each

category (46, 21 and 26 genes downexpressed during urine exposure, respectively). Moreover, genes encoding enzyme potentially involved in respiratory metabolism such as the fumarate reductase (EF2556, -35.9 fold) or enzyme of demethylmenaquinone metabolism (EF0445-EF0450, -11.9 fold) have a reduced expression compared to our control condition.

#### *E. faecalis exposed to serum shows transcriptomic response closed to that of urine.*

In the case of microorganisms that enter the bloodstream and cause infection, bacterial cells are exposed to the host innate immune response, through the proteins of the complement system or antimicrobial peptides/proteins (phospholipases, lysozyme, defensin or lactoferrin) found in the serum. Vebø *et al.* (Vebø *et al.*, 2009) studied the *E. faecalis* transcriptome during growth in blood, which partially overlaps with the one in serum condition (Figure 3 and Table S1), as example genes predicted to be involved in virulence and stress response: the *gls24* operon (EF0076-EF0080, +7.0) (Nannini *et al.*, 2005; Teng *et al.*, 2005), *dps* (EF0606, +5.8) (Xia *et al.*, 2017; Yan *et al.*, 2014), *salB* (EF0394, -84.9) (Mohamed *et al.*, 2006), *ahpCF* (EF2739-EF2738, -24.9) (La Carbona *et al.*, 2007), a gene encoding a cell wall surface anchor family protein (EF2713, +13.0) (Paulsen *et al.*, 2003), *efaA* (EF2076, +5.8) (Garsin *et al.*, 2014), or *cpsG* that intervenes in capsule synthesis (EF2489, -6.5) (Hancock *et al.*, 2003). About metabolic pathways, we also found enhanced expression of genes involved in citrate (EF3322-EF3317, +19.3) and arginine (EF0104-EF0106, +6.7) metabolisms or branched chain amino-acids catabolism (EF1663-EF1659, +19.3). A total number of 50 transport-related genes were highly transcribed during serum exposure including 7 PTS and 10 ABC transport systems, favoring the entry of gluconate, mannitol, galactose or maltose, but also iron and nucleotides. It also has to be noted that this transcriptome due to serum exposure shared 60% with the urine gene expression response (Table 2), suggesting that cells have close behavior in these two liquids.

*Transport and stress response genes are highly expressed in order to respond to acid stress.*

Many enterococcal infections are from endogenous origins (gastrointestinal tract), and macrophages may serve as a vehicle facilitating invasion. Once the infection started, phagocytic cells are recruited in order to eliminate invading pathogens by phagocytosis and incorporation into a phagosome. The phagosome matures and leads to phagolysosomes (Flannagan *et al.*, 2009) that is achieved in 12-15 min after internalization of pathogen (Russell *et al.*, 2009). During maturation, phagosomal membrane is altered and lumen becomes an extremely acidic (pH 4.5 to 5.5), degradative (proteases, lytic enzymes, antimicrobial peptides) and oxidative (reactive oxygen or nitrogen species) environment. After phagosome maturation the immune functions intervene, with inflammation and antigen presentation following lysosomal proteolysis. Thus, bacterial abilities to survive to acid pH are essential for the colonization process and have to be considered. As expected, expression of several ATPases was increased during acid pH to maintain intracellular pH and metal ion homeostasis (Figure 3 and Table S1). This was the case of *kdpABC* (EF0567-EF0570, +9.6 fold), EF0875 (+9.6 fold), EF0298-EF0299 (+11.6 fold), EF0758 (+11.6), *atpFHAGDC* (EF2607-EF2612, +12.2 fold), and also the arginine deiminase system (EF0104-EF0106, +8.3 fold) that releases alkaline molecules such as ammonia. The RNA-Seq results revealed that the expression of several genes associated with pentose phosphate pathway was enhanced by acid pH (EF0556-EF0557, EF2961, and EF3235). Acid pH exposure also increased the level of expression of general stress protection genes including *gsp64* (EF2797, +5.6 fold), EF1744 (+5.5 fold), EF3035 (+5.8 fold), *ohr* (EF0453, +17.8 fold), *gls24* (EF0080, +19.1 fold), *fer* (EF1543, +34.8 fold), *groEL-groES* (EF2634-EF2633, +6.6 fold), genes involved in virulence like *efaA* (EF2076, +14.3), and the *vanB* operon (EF2297-EF2293, +17.3 fold). Cell envelope and membrane fatty acid synthesis was much less expressed, as example the *mraY-murDG* (EF0992-EF0994, -5.4 fold) or the

*dltXABCD* (EF2750-EF2746, -9.8 fold) operons, suggesting that cells may inhibit their growth to use their resources to survive.

*H<sub>2</sub>O<sub>2</sub> causes specific oxidative response genes and proteins synthesis.*

As previously mentioned, during the process of infection, bacteria are also exposed to massive oxidative stress (Thomas *et al.*, 1988) that damages lipids, proteins and DNA biomolecules. NADPH oxidase of the phagosome generates superoxide during the first 20 min following the phagosome formation. This superoxide can be converted into hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>) by superoxide dismutase and can combine with nitric oxide, producing peroxynitrite, or iron, generating hydroxyl radical and hypervalent iron (Russell *et al.*, 2009). Oxidative stress has been well studied in *E. faecalis*, and a link was reported with survival within macrophages (La Carbona *et al.*, 2007; Riboulet *et al.*, 2007). As expected, in presence of H<sub>2</sub>O<sub>2</sub>, we found that genes encoding the NADH peroxidase Npr (EF1211), the alkyl peroxide reductase AhpC (EF2739) and Ohr (EF0453) had an enhanced expression (+10.2, +8.2, +6.3 fold, respectively) (Figure 3 and Table S1). Also was awaited a higher expression of DNA repair functions. Sub lethal dose of hydrogen peroxide induced several metabolic functions, such as protein (24 ribosomal proteins) and phospholipid synthesis (EF2882-EF2875, +5.4 fold), or general stress response chaperones (*dnaK*, EF1308, +5.6 fold). In addition, significantly altered expression was observed for gene of amino-acid synthesis and transport.

Thus this study allowed to complete other transcriptomic studies carried out in *E. faecalis* with cells grown in presence of bovine bile, urine or blood (Solheim *et al.*, 2007; Vebø *et al.*, 2010, 2009), but also gave additional information about gene expressions during serum exposure, acid



or oxidative stress. Moreover, we dissect the non-coding transcriptome involved in stress response.

***Expression of non coding-RNA was strongly affected when E. faecalis copes with host stresses.***

*RNA-Seq allowed identification of new expressed intergenic regions and sRNA.*

In this RNA-seq assay, the expression of many stress-related sRNA was modified in response to the stress exposure: we chose to select sRNA with a cutoff of 2 ( $p < 0.05$ ) to allow the identification of a higher number of relevant intergenic regions. We have detected 65 putative sRNA and other untranslated intergenic regions (that we designed by the genes flanking the sRNA/untranslated region), with 25 of them that have never been highlighted so far (Livny *et al.*, 2008; Innocenti *et al.*, 2015; Shioya *et al.*, 2011; Fouquier d'Hérouel *et al.*, 2011) (Table 3). Although, not all the putative sRNA listed by Fouquier d'Hérouel *et al.* or Shioya *et al.* have been experimentally confirmed as being sRNA. Among these, 10 sRNA have been confirmed by Northern Blot or other methods (Fouquier d'Hérouel *et al.*, 2011; Shioya *et al.*, 2011) (Table 3). EF0129\_EF0130, EF0147\_EF0149, EF1146\_EF1147, EF1350\_EF1351, EF1412\_EF1413 and EF2085\_EF2086 have been found to probably act by pairing to mRNA, and then affect the translation or modulate the activity of proteins (Innocenti *et al.*, 2015). The other intergenic transcripts can potentially be sRNA or 5'UTR, and more investigations are needed to determine which categories they belong to. For instance, EF0869\_EF0871 is strongly suspected to be a riboswitch or a 5'UTR of the EF0869 gene, encoding a putative transcriptional regulator (Fouquier d'Hérouel *et al.*, 2011).

**Table 3:** Differential expression of sRNA genes and intergenic regions during host relevant stress. sRNA and intergenic regions differentially expressed in at least one stress condition, with a cut-off fold  $\geq 2$ ,  $p \leq 0.05$  and  $\log\text{CPM} > 1$  are indicated.

Small RNA	BS	Urine	Serum	Acid	Oxidative
EF0096_EF0097				2.2	
EF0129_EF0130				4.3	3.8
EF0147_EF0149	2.4				
EF0156_EF0157				-2.6	
EF0401_EF0402	-2.0				2.3
EF0408_EF0409	4.3	4.1	3.3	4.5	4.0
EF0513_EF0514		2.4			
EF0595_EF0598	2.2			2.8	2.3
EF0605_EF0606	2.3	2.0	2.4	2.8	2.3
EF0628_EF0629	2.8				3.1
EF0631_EF0633		2.3		3.3	
EF0723_EF0724		-2.0	-2.8		
EF0747_EF0748		2.0			
EF0806_EF0807				2.5	
EF0827_EF0828			-3.2		
EF0845_EF0846	3.0			3.2	
EF0865_EF0867				2.0	2.2
EF0868_EF0869				4.7	
EF0869_EF0871		4.1	4.8		
EF0913_EF0914			-2.4		
EF1097_EF1098	2.0	2.4	3.4	2.7	2.4
EF1146_EF1147	2.1				
EF1151_EF1152				-2.1	
EF1204_EF1205	4.4	3.1	2.4	4.3	2.4
EF1244_EF1246			-2.1		
EF1303_EF1304				2.0	
EF1350_EF1351					-2.2
EF1368_EF1369					-2.7
EF1378_EF1379	2.7			2.6	
EF1412_EF1413	-2.4				
EF1452_EF1453		2.7	3.1	2.6	2.1
EF1513_EF1515	2.7	2.7	3.2	2.9	2.2
EF1540_EF1541		-3.5	-2.9		
EF1599_EF1601	3.0	2.4		3.3	
EF1634_EF1635	2.7			3.1	
EF1635_EF1637		2.1			-2.2
EF1935_EF1936		2.7			
EF1979_EF1980	-6.8	-3.8	-2.9	-4.4	
EF1989_EF1990				3.3	2.2
EF2021_EF2022		4.6	3.9		3.9
EF2085_EF2086	3.1	2.8	3.8	4.3	3.3
EF2205_EF2206		2.4	2.9	2.2	
EF2255_EF2257		-3.0		-3.4	
EF2365_EF2366				2.2	
EF2463_EF2464		-2.3			
EF2484_EF2485		2.0	2.2	2.6	
EF2542_EF2543	2.4	2.0			3.2
EF2665_EF2666					-2.2
EF2696_EF2697	-2.2				
EF2740_EF2741		4.6			3.1
EF2770_EF2771		-2.5			-3.0
EF2771_EF2772	2.1			2.9	

EF2778_EF2780						-4.5
EF2858_EF2859		-2.1				
EF2863_EF2864	2.3			2.8		
EF2944_EF2945	2.1		2.4	2.2		
EF3081_EF3082				2.9		
EF3086_EF3087	2.3	-2.4				-2.3
EF3087_EF3088	2.6					
EF3158_EF3160				-2.3		
EF3160_EF3161		8.7	8.8	6.1		
EF3247_EF3248	2.9	3.1	3.3	3.4		2.2
EF3260_EF3261		5.7	3.0			
EF3262_EF3263		4.1	4.5			
EF3314_EF3315	2.5	2.2	2.6			3.8

*Untranslated regions had common or specific expression during host stress exposure.*

Most of the sRNA or untranslated intergenic regions had higher expression following stress exposure, especially for acid pH and urine: thus, among 34 and 32 sRNA differentially expressed, 29 and 23 were more abundant, respectively (Table 3). Seven of them, like EF1097\_EF1098 (SsrA) and EF0605\_EF0606, had a higher expression in all the tested stresses, suggesting a major role in the general stress response. SsrA is the tmRNA involved in the *trans*-translation process, which is a RNA quality control mechanism rescuing blocked ribosome on aberrant mRNA, and releasing the incomplete protein. Studies suggest that it could also be involved in virulence, sporulation, cell division, antibiotic resistance and stress response in other bacterial species (Hayes and Keiler, 2010). The fact that SsrA expression was elevated in all the tested stresses agrees with its crucial function during stress response, also documented by Shioya *et al.* (Shioya *et al.*, 2011).

EF0605\_EF0606, located in the V583 pathogenicity island, was shown to be more transcribed when *E. faecalis* is confronted to an oxidative stress (Fouquier d'Hérouel *et al.*, 2011), but also during bile salts, acid stress, urine or serum exposure (Table 3).

EF3314\_EF3315 showed a higher expression in almost all the tested stresses, except acid stress compared to our control condition. This could mean also an implication in global stress response. Indeed, this sRNA have been already detected and studied by Shioya *et al.* and

Michaux *et al.*, who found that it could be involved in the turnover of some abundant proteins (Shioya *et al.*, 2011) and stress response (Michaux *et al.*, 2014b).

Some sRNA or intergenic regions were differentially expressed in one specific condition, like EF1412\_EF1413 and EF2696\_EF2697 that had a reduced expression specifically in bile salts condition, or EF0806\_EF0807, EF1303\_EF1304, EF0096\_EF0097, EF0868\_EF0869, EF2365\_EF2366 and EF3081\_EF3082 that were more expressed only in acid stress condition.

*Some sRNA could impact expression of genes located in their direct genomic environment.*

As small RNAs often regulate mRNA stability of genes located immediately upstream or downstream of the sRNA locus, we were interested in expression of genes located in the direct genomic environment of the sRNA (Table S2). Interestingly, EF0806\_EF0807 upstream gene, EF0806, encoding an amino acid ABC transporter, had also an increased expression in acid condition. In contrast, EF0097 (encoding a hypothetical regulatory protein), downstream of EF0096\_EF0097, was less expressed in acid conditions. There were other sRNA and their upstream or downstream (sometimes both) genes that were expressed in the same way, like EF1989\_EF1990 and EF1990 in acid stress, EF2778\_EF2780 and EF2780 in oxidative stress, and EF3087\_EF3088 and EF3087 in bile salt (Table S2). The three corresponding genes are all encoding hypothetical proteins.

EF1989\_EF1990 had also a higher expression in oxidative stress, which could be interesting considering its upstream gene, EF1989, encoding the ferrochelatase HemH, an enzyme able to produce heme by inserting a Fe<sup>2+</sup> into porphyrin ring (Almirón *et al.*, 2001). Although, as *E. faecalis* is known to be unable to synthesize heme (Frankenberg *et al.*, 2002), the function of this enzyme in this bacterium is not clear. Nevertheless, it has been shown that this enzyme is involved in intracellular survival in *Neisseria gonorrhoeae* and *Brucella abortus* (Almirón *et*

*al.*, 2001), where it could be linked to an oxidative stress. This gene showed no variation of expression in any tested stress, but may be regulated at the translational level by its associated sRNA.

The putative sRNA EF0628\_EF0629 was more expressed in bile salts and oxidative stress, as well as its downstream gene (Table S2). The latter gene EF0629 was also more present in serum, and encodes an oxidoreductase, which is coherent with its expression during an oxidative stress. This suggests that EF0628\_EF0629 could be involved in response to this stress.

On the contrary, transcription of some sRNA varied in an opposite manner to that of the gene located upstream or downstream. For example, EF1368\_EF1369 was less expressed in oxidative stress while its downstream gene (EF1368) was more expressed in this same stress (Table S2). It is also the case in urine for EF2858\_EF2859 and its downstream gene, encoding a 4-oxalocrotonate tautomerase, an enzyme involved in degradation of aromatic hydrocarbons (Whitman, 2002). Also, EF1204\_EF1205 and EF1204 (hypothetical protein) have opposite expression in almost all the stresses (except oxidative stress). It may be speculated that the interaction between the sRNA and mRNA in these cases led to an increase of mRNA degradation.

## **Conclusion**

With the present study we provide novel information about the gene regulation relevant to an adaptation process during commensal life or colonization. The overview of transcriptomic response in five stress conditions encountered in host, *i.e.* bile salts, urine, serum, acid pH or oxidative stress, is usefully informative. Indeed, this work allowed the detection of 65 intergenic regions and sRNA, including 25 unknown until yet, involved in stress response encountered during host colonization. Moreover, these transcriptome studies highlight the fact that *E.*

*faecalis* represents a redoubtable opportunistic bacterium because of its very efficient stress responses rather than because of virulence factors.

### **Supplementary data**

*Note : Afin de privilégier la clarté du manuscrit, les données supplémentaires précédemment citées sont disponibles en Annexe 1.*

### **Funding**

This work was partly supported by grants from the Agence Nationale de la Recherche in the frame of a transnational ERA-NET PathoGenoMics program (ANR-08-PATH-008-01).

### **Acknowledgments**

The expert technical assistance of Isabelle Rincé was greatly appreciated. We would like to thank all members of our labs for helpful discussions.

## References

- Aedo, J.E., Maldonado, J., Aballai, V., Estrada, J.M., Bastias-Molina, M., Meneses, C., *et al.* (2015) mRNA-seq reveals skeletal muscle atrophy in response to handling stress in a marine teleost, the red cusk-eel (*Genypterus chilensis*). *BMC Genomics* **16**: 1024.
- Almirón, M., Martínez, M., Sanjuan, N., and Ugalde, R.A. (2001) Ferrochelatase is present in *Brucella abortus* and is critical for its intracellular survival and virulence. *Infect Immun* **69**: 6225–6230.
- Arias, C.A., and Murray, B.E. (2012) The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat Rev Microbiol* **10**: 266–278.
- Babady, N.E. (2016) Hospital-associated infections. *Microbiol Spectr* **4** <http://www.asmscience.org/content/journal/microbiolspec/10.1128/microbiolspec.DMIH2-0003-2015>. Accessed June 27, 2017.
- Bizzini, A., Zhao, C., Budin-Verneuil, A., Sauvageot, N., Giard, J.-C., Auffray, Y., and Hartke, A. (2010) Glycerol is metabolized in a complex and strain-dependent manner in *Enterococcus faecalis*. *J Bacteriol* **192**: 779–785.
- Comenge, Y., Quintiliani, R., Jr, Li, L., Dubost, L., Brouard, J.-P., Hugonnet, J.-E., and Arthur, M. (2003) The CroRS two-component regulatory system is required for intrinsic beta-lactam resistance in *Enterococcus faecalis*. *J Bacteriol* **185**: 7184–7192.
- Flannagan, R.S., Cosío, G., and Grinstein, S. (2009) Antimicrobial mechanisms of phagocytes and bacterial evasion strategies. *Nat Rev Microbiol* **7**: 355–366.
- Fouquier d’Hérouel, 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 Res* **39**: e46.
- Frankenberg, L., Brugna, M., and Hederstedt, L. (2002) *Enterococcus faecalis* heme-dependent catalase. *J Bacteriol* **184**: 6351–6356.
- Garsin, D.A., Frank, K.L., Silanpää, J., Ausubel, F.M., Hartke, A., Shankar, N., and Murray, B.E. (2014) Pathogenesis and models of enterococcal infection, in Gilmore, M.S., Clewell, D.B., Ike, Y., Shankar, N. (Eds.), *Enterococci: From commensals to leading causes of drug resistant infection*. Massachusetts Eye and Ear Infirmary, Boston. <http://www.ncbi.nlm.nih.gov/books/NBK190426/>. Accessed June 25, 2014.
- Gevrekci, A.Ö. (2017) The roles of polyamines in microorganisms. *World J Microbiol Biotechnol* **33**: 204.
- Hancock, L., and Perego, M. (2002) Two-component signal transduction in *Enterococcus faecalis*. *J Bacteriol* **184**: 5819–5825.
- Hancock, L.E., Shepard, B.D., and Gilmore, M.S. (2003) Molecular analysis of the *Enterococcus faecalis* serotype 2 polysaccharide determinant. *J Bacteriol* **185**: 4393–4401.

- Harten, R.M. van, Willems, R.J.L., Martin, N.I., and Hendrickx, A.P.A. (2017) Multidrug-resistant enterococcal infections: New compounds, novel antimicrobial therapies? *Trends Microbiol* **25**: 467–479.
- Hayes, C.S., and Keiler, K.C. (2010) Beyond ribosome rescue: tmRNA and co-translational processes. *FEBS Lett* **584**: 413–419.
- Innocenti, N., Golumbeanu, M., Hérouël, A.F. d', 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.
- Ipe, D.S., Horton, E., and Ulett, G.C. (2016) The basics of bacteriuria: strategies of microbes for persistence in urine. *Front Cell Infect Microbiol* **6** <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4744864/>. Accessed February 13, 2018.
- Johnson, J.R. (1991) Virulence factors in *Escherichia coli* urinary tract infection. *Clin Microbiol Rev* **4**: 80–128.
- Kellogg, S.L., Little, J.L., Hoff, J.S., and Kristich, C.J. (2017) Requirement of the CroRS two-component system for resistance to cell wall-targeting antimicrobials in *Enterococcus faecium*. *Antimicrob Agents Chemother* **61**.
- La Carbona, S., 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*. *Mol Microbiol* **66**: 1148–1163.
- Le Breton, Y., Boël, G., Benachour, A., Prévost, H., Auffray, Y., and Rincé, A. (2003) Molecular characterization of *Enterococcus faecalis* two-component signal transduction pathways related to environmental stresses. *Environ Microbiol* **5**: 329–337.
- Livny, J., Teonadi, H., Livny, M., and Waldor, M.K. (2008) High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. *PloS One* **3**: e3197.
- Lv, L.-X., Yan, R., Shi, H.-Y., Shi, D., Fang, D.-Q., Jiang, H.-Y., *et al.* (2017) Integrated transcriptomic and proteomic analysis of the bile stress response in probiotic *Lactobacillus salivarius* LI01. *J Proteomics* **150**: 216–229.
- Madsen, K.T., Skov, M.N., Gill, S., and Kemp, M. (2017) Virulence factors associated with *Enterococcus faecalis* infective endocarditis: a mini review. *Open Microbiol J* **11**: 1–11.
- Matos, R.C., Lapaque, N., Rigottier-Gois, L., Debarbieux, L., Meylheuc, T., Gonzalez-Zorn, B., *et al.* (2013) *Enterococcus faecalis* prophage dynamics and contributions to pathogenic traits. *PLoS Genet* **9**: e1003539.
- 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.
- Michaux, C., Verneuil, N., Hartke, A., and Giard, J.-C. (2014b) Physiological roles of small RNA molecules. *Microbiol Read Engl* **160**: 1007–1019.



- Mohamed, J.A., Teng, F., Nallapareddy, S.R., and Murray, B.E. (2006) Pleiotrophic effects of 2 *Enterococcus faecalis* sagA-like genes, salA and salB, which encode proteins that are antigenic during human infection, on biofilm formation and binding to collagen type i and fibronectin. *J Infect Dis* **193**: 231–240.
- Muller, C., Cacaci, M., Sauvageot, N., Sanguinetti, M., Rattei, T., Eder, T., *et al.* (2015) The intraperitoneal transcriptome of the opportunistic pathogen *Enterococcus faecalis* in mice. *PLoS One* **10**: e0126143.
- Muller, C., Le Breton, Y., Morin, T., Benachour, A., Auffray, Y., and Rincé, A. (2006) The response regulator CroR modulates expression of the secreted stress-induced SalB protein in *Enterococcus faecalis*. *J Bacteriol* **188**: 2636–2645.
- Muller, C., Massier, S., Le Breton, Y., and Rincé, A. (2018) The role of the CroR response regulator in resistance of *Enterococcus faecalis* to D-cycloserine is defined using an inducible receiver domain. *Mol Microbiol* **107**: 416–427.
- Muller, C., Sanguinetti, M., Riboulet, E., Hébert, L., Posteraro, B., Fadda, G., *et al.* (2008) Characterization of two signal transduction systems involved in intracellular macrophage survival and environmental stress response in *Enterococcus faecalis*. *J Mol Microbiol Biotechnol* **14**: 59–66.
- Murray, B.E. (1990) The life and times of the *Enterococcus*. *Clin Microbiol Rev* **3**: 46–65.
- Nallapareddy, S.R., and Murray, B.E. (2008) Role played by serum, a biological cue, in the adherence of *Enterococcus faecalis* to extracellular matrix proteins, collagen, fibrinogen, and fibronectin. *J Infect Dis* **197**: 1728–1736.
- Nannini, E.C., Teng, F., Singh, K.V., and Murray, B.E. (2005) Decreased virulence of a gls24 mutant of *Enterococcus faecalis* OG1RF in an experimental endocarditis model. *Infect Immun* **73**: 7772–7774.
- O’Driscoll, T., and Crank, C.W. (2015) Vancomycin-resistant enterococcal infections: epidemiology, clinical manifestations, and optimal management. *Infect Drug Resist* **8**: 217–230.
- Oliva, G., Sahr, T., and Buchrieser, C. (2015) Small RNAs, 5’ UTR elements and RNA-binding proteins in intracellular bacteria: impact on metabolism and virulence. *FEMS Microbiol Rev* **39**: 331–349.
- Ostrolenk, M., Kramer, N., and Cleverdon, R.C. (1947) Comparative studies of enterococci and *Escherichia coli* as indices of pollution. *J Bacteriol* **53**: 197–203.
- Paulsen, I.T., Banerjee, L., Myers, G.S.A., Nelson, K.E., Seshadri, R., Read, T.D., *et al.* (2003) Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* **299**: 2071–2074.
- Riboulet, E., Verneuil, N., La Carbona, S., Sauvageot, N., Auffray, Y., Hartke, A., and Giard, J.-C. (2007) Relationships between oxidative stress response and virulence in *Enterococcus faecalis*. *J Mol Microbiol Biotechnol* **13**: 140–146.

- Rincé, A., Le Breton, Y., Verneuil, N., Giard, J.-C., Hartke, A., and Auffray, Y. (2003) Physiological and molecular aspects of bile salt response in *Enterococcus faecalis*. *Int J Food Microbiol* **88**: 207–213.
- Russell, D.G., VanderVen, B.C., Glennie, S., Mwandumba, H., and Heyderman, R.S. (2009) The macrophage marches on its phagosome: dynamic assays of phagosome function. *Nat Rev Immunol* **9**: 594–600.
- Saxena, R.K., Anand, P., Saran, S., Isar, J., and Agarwal, L. (2010) Microbial production and applications of 1,2-propanediol. *Indian J Microbiol* **50**: 2–11.
- Shankar, N., Baghdayan, A.S., and Gilmore, M.S. (2002) Modulation of virulence within a pathogenicity island in vancomycin-resistant *Enterococcus faecalis*. *Nature* **417**: 746–750.
- 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**: e23948.
- 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. *Sci Rep* **7**: 11067.
- Solheim, M., Aakra, Å., Vebø, H., Snipen, L., and Nes, I.F. (2007) Transcriptional responses of *Enterococcus faecalis* V583 to bovine bile and sodium dodecyl sulfate. *Appl Environ Microbiol* **73**: 5767–5774.
- Steen, A., Palumbo, E., Deghorain, M., Cocconcelli, P.S., Delcour, J., Kuipers, O.P., *et al.* (2005) Autolysis of *Lactococcus lactis* is increased upon D-alanine depletion of peptidoglycan and lipoteichoic acids. *J Bacteriol* **187**: 114–124.
- Teng, F., Nannini, E.C., and Murray, B.E. (2005) Importance of gls24 in virulence and stress response of *Enterococcus faecalis* and use of the Gls24 protein as a possible immunotherapy target. *J Infect Dis* **191**: 472–480.
- Thomas, E.L., Lehrer, R.I., and Rest, R.F. (1988) Human neutrophil antimicrobial activity. *Rev Infect Dis* **10 Suppl 2**: S450-456.
- Vebø, H.C., Snipen, L., Nes, I.F., and Brede, D.A. (2009) The transcriptome of the nosocomial pathogen *Enterococcus faecalis* V583 reveals adaptive responses to growth in blood. *PloS One* **4**: e7660.
- Vebø, H.C., Solheim, M., Snipen, L., Nes, I.F., and Brede, D.A. (2010) Comparative genomic analysis of pathogenic and probiotic *Enterococcus faecalis* isolates, and their transcriptional responses to growth in human urine. *PloS One* **5**: e12489.
- Waters, L.S., and Storz, G. (2009) Regulatory RNAs in bacteria. *Cell* **136**: 615–628.
- Wells, C.L., Jechorek, R.P., and Erlandsen, S.L. (1990) Evidence for the translocation of *Enterococcus faecalis* across the mouse intestinal tract. *J Infect Dis* **162**: 82–90.
- Whitman, C.P. (2002) The 4-oxalocrotonate tautomerase family of enzymes: how nature makes new enzymes using a beta-alpha-beta structural motif. *Arch Biochem Biophys* **402**: 1–13.

Xia, X., Larios-Valencia, J., Liu, Z., Xiang, F., Kan, B., Wang, H., and Zhu, J. (2017) OxyR-activated expression of Dps is important for *Vibrio cholerae* oxidative stress resistance and pathogenesis. *PLoS One* **12**: e0171201.

Xu, Y., Jiang, L., Murray, B.E., and Weinstock, G.M. (1997) *Enterococcus faecalis* antigens in human infections. *Infect Immun* **65**: 4207–4215.

Yan, X., Budin-Verneuil, A., Auffray, Y., and Pichereau, V. (2014) Proteome phenotyping of  $\Delta$ relA mutants in *Enterococcus faecalis* V583. *Can J Microbiol* **60**: 525–531.

## Chapitre 2 : Etude du rôle et de l'action d'un petit ARN régulateur agissant en trans

### 1. SRC65 est un petit ARN régulateur

SRC65 est un petit ARN régulateur putatif mis en évidence lors de précédentes études (Shioya *et al.*, 2011; Innocenti *et al.*, 2015) et détecté comme étant fortement exprimé *in vivo* lors d'infections intra-péritonéales chez la souris (Muller *et al.*, 2015).

SRC65 est situé sur le brin négatif du chromosome bactérien. Il mesure approximativement 166 bases selon les données du RNA-Seq, et son locus est situé entre les gènes *ef1204* et *ef1205* de la souche V19 (Figure 14). Le gène *ef1204* code une petite protéine de fonction inconnue (protéine hypothétique) tandis qu'*ef1205* code un régulateur transcriptionnel non caractérisé. Ces deux gènes sont situés sur le brin positif de l'ADN ce qui écarte la possibilité que SRC65 soit une 5'UTR.

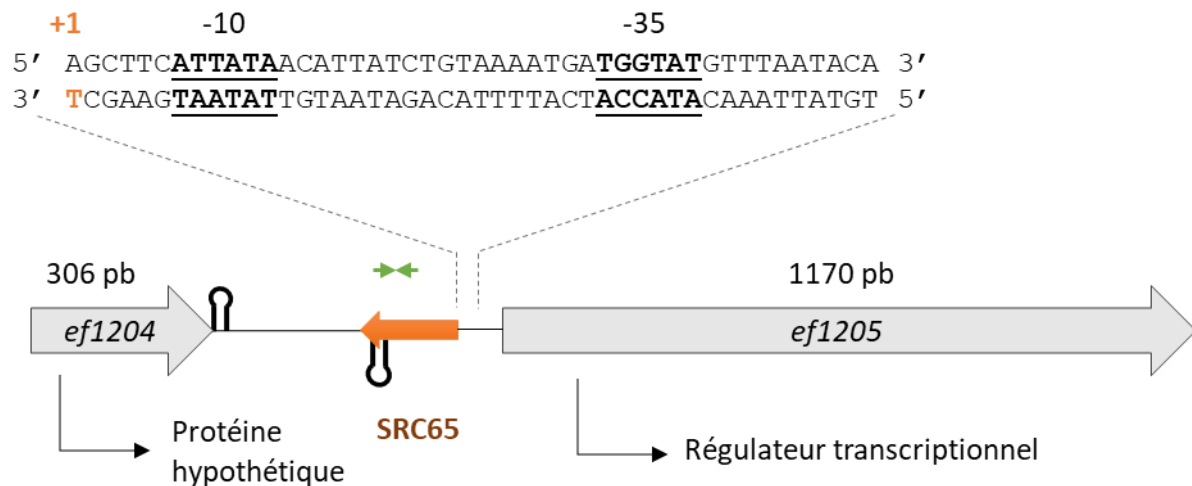
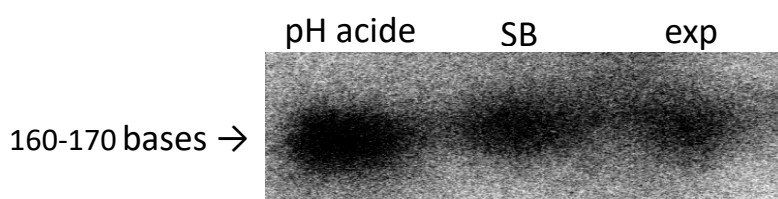


Figure 14 : Environnement génomique du petit ARN SRC65 (en orange). Les terminateurs détectés par l'outil FindTerm et la localisation de la sonde pour le northern blot (*SRC65\_SP1* et *src65\_Forward*, en vert) sont indiqués. Les tailles des parties codantes sont précisées. La séquence promotrice et le +1 de transcription (en orange) de SRC65 sont indiqués (en double brin).

Des terminateurs de transcription ont été détectés grâce à l'outil FindTerm en aval d'*efl204* et en fin de séquence de SRC65. Nous avons aussi décelé la présence de séquences caractéristiques d'un promoteur (boîtes -10 et -35), et observé l'absence d'une séquence Shine-Dalgarno, ce qui sous-entend que l'ARN SRC65 n'est pas traduit.

La réalisation d'une 5'RACE-PCR nous a permis d'identifier le +1 de transcription de SRC65, situé au même endroit que celui suggéré par le RNA-seq (Muller *et al.*, 2015). Une 3'RACE-PCR nous a également montré que le transcrit avait une taille comprise entre 149 et 153 bases, une séquence riche en T n'ayant pas permis de situer précisément l'extrémité 3' par rapport à la queue poly-A ajoutée *in vitro* pour la réaction de 3'RACE-PCR.

Pour prouver définitivement l'existence de SRC65, une expérience de northern blot a été réalisée avec des ARN extraits de la souche V19 en phase exponentielle de croissance, en condition de stress acide, ou en présence de sels biliaires (Figure 15), avec une sonde localisée comme indiqué sur la figure 14. Les données de RNA-Seq ont montré que l'expression de SRC65 était induite dans ces conditions de stress (voir Chapitre I et Chapitre II partie 3), qui sont fréquemment rencontrées par *E. faecalis* lors de la colonisation.



*Figure 15 : Northern blot mettant en évidence le sARN SRC65. L'expérience a été réalisée avec des ARN extraits en phase exponentielle de croissance (exp) et en condition de stress : sels biliaires 0,06% (SB) et pH 5,8 (pH acide) et hybridés avec une sonde générée par PCR à l'aide des amorces SRC65\_SP1 et src65\_Forward (voir Tableau 3 et Figure 14).*

Nous avons obtenu un signal correspondant à un transcrit d'environ 160-170 bases, ce qui concorde approximativement avec la taille prédite par le RNA-Seq et avec celle déduite des

expériences de RACE-PCR. Ce transcrit est exprimé dans les trois conditions testées, avec un signal observé sur le northern blot plus fort pour la condition pH acide par rapport aux conditions phase exponentielle de croissance et sels biliaires.

SRC65 réunit donc toutes les caractéristiques d'un petit ARN régulateur : pas de co-transcription avec les gènes flanquants, absence de séquence Shine-Dalgarno, et une taille qui concorde avec celle d'un sARN.

## 2. SRC65 est homologue à un autre sARN putatif

Afin d'identifier les cibles de SRC65, une recherche d'homologie en utilisant l'outil BLAST a été réalisée. Une séquence homologue à SRC65 (66% d'identité dont 79% de bases identiques sur une région de 72 nucléotides) a ainsi été mise en évidence, correspondant à un autre sARN putatif détecté par le RNA-Seq et nommé SRC90 (voir Chapitre I et Shioya *et al.*, 2011). L'homologie entre les deux sARN se situe principalement sur leurs 100 premières bases (Figure 16).

```

          1                               50
SRC65  UAUCAUAAUA AAGGAGGUGA AAUAUAUGUU UGAUUUACUU GCUUUUUAA
SRC90  .AUCA.AUUA AAGGAGGUAC AUGCUAUGUU UGAUUUACUU GCAUUUUUGA

          51                               100
SRC65  CUGGAAUUUC AGUUCGUGGC AAUGCACCAC ACCAAUAAUU ACGUGUUUGU
SRC90  CUGGAACAUC ACUUAUGGC AACGUUCCCC AUGUUUAAUU A.GUCAAUGC

          101                              150
SRC65  UUAACUGAAU UUCUGACGGC UCUACCUAUG CCAGUAGGUG GAGCCAUUUU
SRC90  GUGUC..AAG UAUGGUCUUC CAUACUUG.G CAUGUUUC... ..

          151                              169
SRC65  UUUUUUUAAU AUUAUAAA
SRC90  .....

```

Figure 16 : Alignement des sARN potentiels SRC65 et SRC90. L'alignement a été réalisé en utilisant le programme MultAlin. Les bases conservées sont montrées en rouge. Les points noirs représentent les bases absentes sur la séquence de SRC90.

En utilisant la même démarche que pour SRC65 (northern blot présenté sur la Figure 17, et la localisation de la sonde sur la figure 18), nous avons pu montrer que SRC90 possédait son propre promoteur, qu'il n'est pas co-transcrit avec les gènes localisés immédiatement en amont ou en aval, et qu'il mesure environ 150 bases (Figure 18) : il s'agit donc également d'un sARN. De plus, nous avons relevé que l'expression de ce sARN était fortement induite en condition de pH acide, et plus légèrement en présence de sels biliaires par rapport à la phase exponentielle (Figure 17).

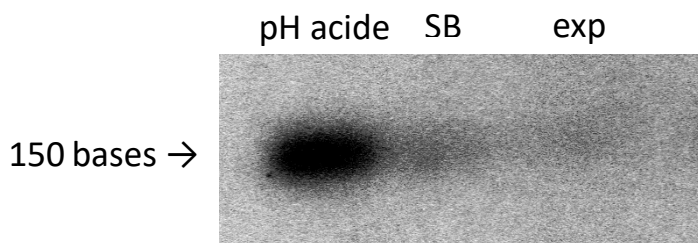


Figure 17 : Northern blot mettant en évidence le sARN SRC90. L'expérience a été réalisée avec des ARN extraits en phase exponentielle de croissance (exp) et en condition de stress : sels biliaires 0,06% (SB) et pH 5,8 (pH acide). L'hybridation a été réalisée avec une sonde générée par PCR à l'aide des amorces SRC90\_SP1 et src90\_Forward (voir Tableau 3 et Figure 18).

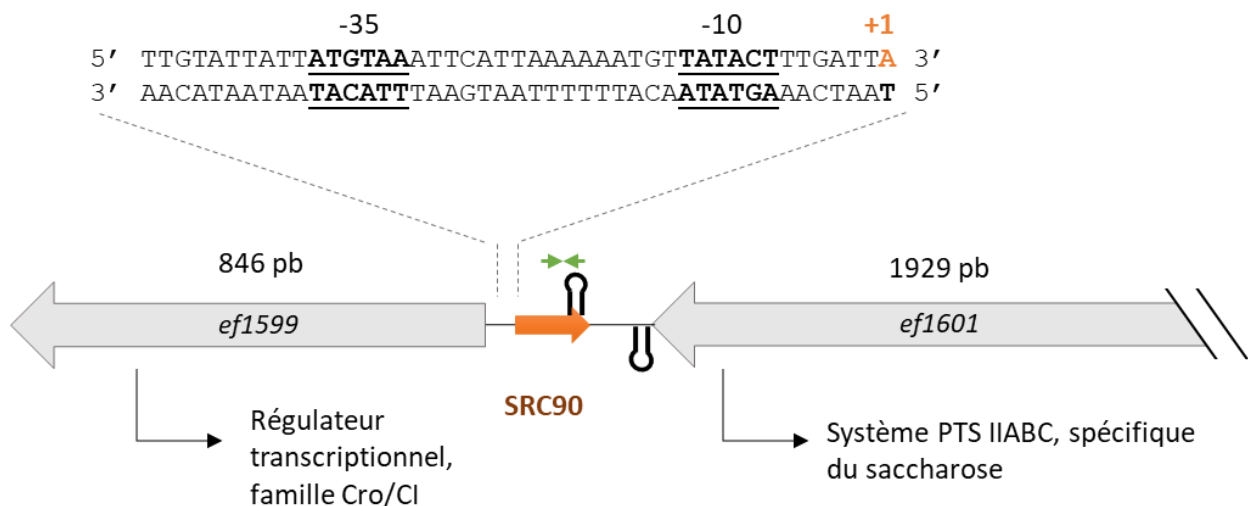


Figure 18 : Environnement génomique du petit ARN SRC90 (en orange). Les terminateurs détectés par l'outil FindTerm et la localisation de la sonde pour le northern blot (SRC90\_SP1 et src90\_Forward, en vert) sont indiqués. Les tailles des parties codantes sont précisées. La séquence promotrice et le +1 de transcription (en orange) de SRC90 sont indiquées.

### 3. Implication dans la réponse générale au stress

Nous nous sommes appuyés sur les résultats du RNA-Seq effectué précédemment (voir Chapitre I) pour observer dans quelles conditions les sARN étaient exprimés.

Nous avons ainsi relevé que SRC65 était induit en présence de sels biliaries (0,6%), d'urine, de sérum, en condition de pH acide (pH 5,8) et de stress oxydatif (H<sub>2</sub>O<sub>2</sub> 1mM) en comparaison avec une culture en phase exponentielle de croissance (milieu non stressant) (Tableau 4). SRC90 est quant à lui surexprimé en présence de sels biliaries, d'urine, et en condition de stress acide. Pour les deux sARN, les surexpressions en pH acide et en sels biliaries ont été confirmées par northern blot (voir résultats précédents décrits en Figures 15 et 17).

*Tableau 4 : Facteurs de variation de l'expression de SRC65 et SCR90 chez la souche V19 d'E. faecalis soumise à des conditions de stress, et comparés aux cellules en phase exponentielle de croissance en milieu BHI non stressant (p-value <0,05). Cases vides : pas de changement d'expression statistiquement significatif.*

<b>Conditions (15 minutes d'exposition)</b>	<b>Sels Biliaires (0,06%)</b>	<b>Urine</b>	<b>Sérum</b>	<b>pH acide (5,8)</b>	<b>Stress oxydatif (H<sub>2</sub>O<sub>2</sub> 1 mM)</b>
<i>SRC65</i>	4,4	3,1	2,4	4,3	2,4
<i>SRC90</i>	3,0	2,4		3,3	

On peut donc émettre l'hypothèse que SRC65 et SRC90 sont potentiellement impliqués dans la réponse générale au stress.

### 4. Implication dans le déclenchement de la phase exponentielle

Afin de déterminer dans quel(s) métabolisme(s) pouvait être impliqué SRC65, un mutant portant une délétion du gène correspondant, ainsi qu'une souche le surexprimant ont été construites, nommées ΔSRC65 et SRC65+, respectivement. La surexpression du sARN a été



obtenue par clonage de la séquence correspondante dans le plasmide pAGent sous le contrôle du promoteur *aguB* inductible par l'agmatine, et a été vérifiée par qPCR (données non montrées). Une souche d'*E. faecalis* hébergeant le plasmide pAGent vide a été utilisée comme contrôle pour SRC65+ (nommée « souche pAGent »).

Nous avons ensuite analysé la croissance des souches V19,  $\Delta$ SRC65, pAGent et SRC65+ en milieu RM17 supplémenté en agmatine (Figure 19).

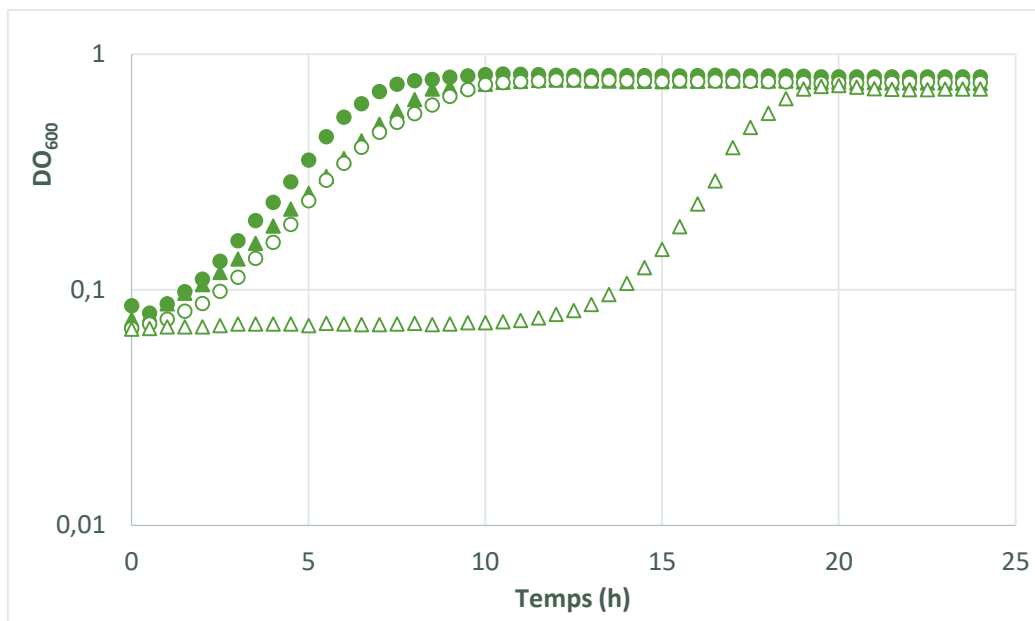


Figure 19: Courbes de croissance des souches V19 (cercles pleins),  $\Delta$ SRC65 (cercles vides), pAGent (triangles pleins) et SRC65+ (triangles vides) à 37°C en RM17 supplémenté en agmatine 40 mM. Expérience répétée trois fois.

Dans ces conditions, la croissance du mutant  $\Delta$ SRC65 est similaire à celle de la souche V19. De même, la souche pAGent ne présente pas de différence de croissance avec la souche V19, ce qui montre que le plasmide vide n'a pas d'influence sur la croissance de la bactérie. En revanche, nous avons constaté chez la souche de surexpression SRC65+ un allongement de la phase de latence de 12 heures.

L'agmatine pouvant être métabolisée par la bactérie, nous avons ensuite cherché à déterminer si la reprise de croissance au bout de 12h n'était pas due à un épuisement de ce substrat. Lorsque

des bactéries SRC65+ collectées au cours de la phase exponentielle de croissance (c'est-à-dire après 16 heures d'incubation) sont transférées dans du milieu RM17 supplémenté en agmatine fraîchement préparé et préincubé à 37°C, la croissance exponentielle reprend sans repasser par une phase de latence (données non montrées). Ceci signifie donc que le déclenchement de la phase exponentielle n'était pas dû à un épuisement de l'inducteur agmatine. Nous avons également répété cette expérience en présence de chloramphénicol (antibiotique dont le gène de résistance est porté par le vecteur pAGEnt) avec les mêmes observations, ce qui montre que ces résultats ne sont pas non plus la conséquence d'une perte du plasmide de surexpression (données non montrées).

Nous pouvons donc en conclure que SRC65 a une influence négative sur le déclenchement de la phase exponentielle de croissance. Il exercerait donc un rôle dans une fonction primordiale de la bactérie pour sa croissance.

## 5. Mise en évidence de cibles potentielles

Pour détecter *in silico* les cibles potentielles de SRC65, l'outil TargetRNA2 qui permet d'identifier des cibles d'un sARN en se basant sur la séquence des ARNm cibles a été utilisé. Il prend notamment en compte la conservation du sARN, son accessibilité ainsi que celle de l'ARNm, et l'énergie nécessaire à l'hybridation (plus elle est faible, plus l'interaction est probable).

L'analyse nous a révélé 56 cibles potentielles, listées dans le tableau 5 dans l'ordre décroissant de probabilité de liaison calculée par le logiciel.

Tableau 5 : Cibles potentielles de SRC65 identifiées à l'aide du logiciel TargetRNA2 ( $p$ -value < 0,05). Les gènes sont listés dans l'ordre croissant d'énergie nécessaire à la liaison.

Rang	Gène	Locus	Produit	Rang	Gène	Locus	Produit
1	-	ef3251	hypothetical protein	29	pepQ-2	ef1743	proline dipeptidase
2	-	ef2221	ABC transporter substrate-binding protein	30	-	ef0886	hypothetical protein
3	-	ef2207	Fis family DNA-binding protein	31	-	ef0888	hypothetical protein
4	foIE	ef3267	GTP cyclohydrolase	32	-	ef3212	PTS system transporter subunit IIC
5	-	ef2820	major capsid protein	33	-	ef1272	hypothetical protein
6	-	ef1925	hypothetical protein	34	feoB	ef0476	ferrous iron transport protein B
7	-	ef0571	DNA-binding response regulator	35	-	ef3197	sensor histidine kinase
8	-	ef0160	hypothetical protein	36	-	ef2165	NAD-dependent epimerase/dehydratase
9	-	ef0740	deoxynucleoside kinase	37	lplA-2	ef2741	lipoate-protein ligase A
10	-	ef1804	PTS system transporter subunit IIB	38	-	ef1240	Lacl family sugar-binding transcriptional regulator
11	rpmG-2	ef2731	50S ribosomal protein L33	39	ftsQ	ef0995	cell division protein FtsQ
12	-	ef1326	TetR family transcriptional regulator	40	-	ef0941	ABC transporter ATP-binding protein/permease
13	-	ef1720	uracil permease	41	rexB	ef1112	exonuclease RexB
14	recN	ef0984	DNA repair protein RecN	42	-	ef2777	hydroxyethylthiazole kinase
15	-	ef2832	hypothetical protein	43	-	ef2910	potassium uptake protein
16	-	ef1018	PTS system transporter subunit IIA	44	-	ef3138	PTS system transporter subunit IID
17	proS	ef2379	prolyl-tRNA synthetase	45	-	ef0660	MATE efflux family protein
18	-	ef3168	hypothetical protein	46	-	ef2621	hypothetical protein
19	-	ef1023	hypothetical protein	47	-	ef0707	hypothetical protein
20	-	ef1871	thioredoxin family protein	48	-	ef2229	hypothetical protein
21	-	ef0873	Cro/Cl family transcriptional regulator	49	drrC	ef1032	daunorubicin resistance protein
22	-	ef1017	PTS system transporter subunit IIB	50	-	ef1938	cation transporter E1-E2 family ATPase
23	thrC	ef2421	threonine synthase	51	-	ef3216	transcriptional regulator
24	-	ef0336	hypothetical protein	52	-	ef0738	hypothetical protein
25	-	ef0022	PTS system mannose-specific transporter subunit IID	53	nagA-1	ef1317	N-acetylglucosamine-6-phosphate deacetylase
26	pyrD-2	ef1714	dihydroorotate dehydrogenase	54	-	ef0318	hypothetical protein
27	mtID	ef0413	mannitol-1-phosphate 5-dehydrogenase	55	-	ef1123	hypothetical protein
28	aceF	ef1355	dihydrolipoamide acetyltransferase	56	-	ef1887	hypothetical protein

Nous avons ensuite cherché à évaluer si la transcription des gènes et/ou la stabilité des ARNm correspondant aux cibles potentielles prédites *in silico* était contrôlée par SRC65. L'expression des 15 premiers gènes de la liste ainsi que celle des gènes encadrant SRC65 et SRC90 a été déterminée par une approche de RT-qPCR, à partir d'ARN extraits des souches V19,  $\Delta$ SRC65, pAGEnt et SCR65+. L'expérience a été répétée pour les gènes *ef0160*, *ef2221* et *ef3251* (Tableau 3) qui ont présenté une potentielle variation d'expression entre les souches sur la première analyse.

Ces analyses ont révélé que la quantité d'ARNm de la grande majorité des gènes testés ne variait pas selon la souche, à l'exception d'*ef0160* dont l'ARNm était présent en plus grande quantité chez la souche SRC65+, et *ef3251* dont l'ARNm était moins abondant chez le mutant  $\Delta$ SRC65 (Figure 20). Ces deux gènes codent des protéines hypothétiques.

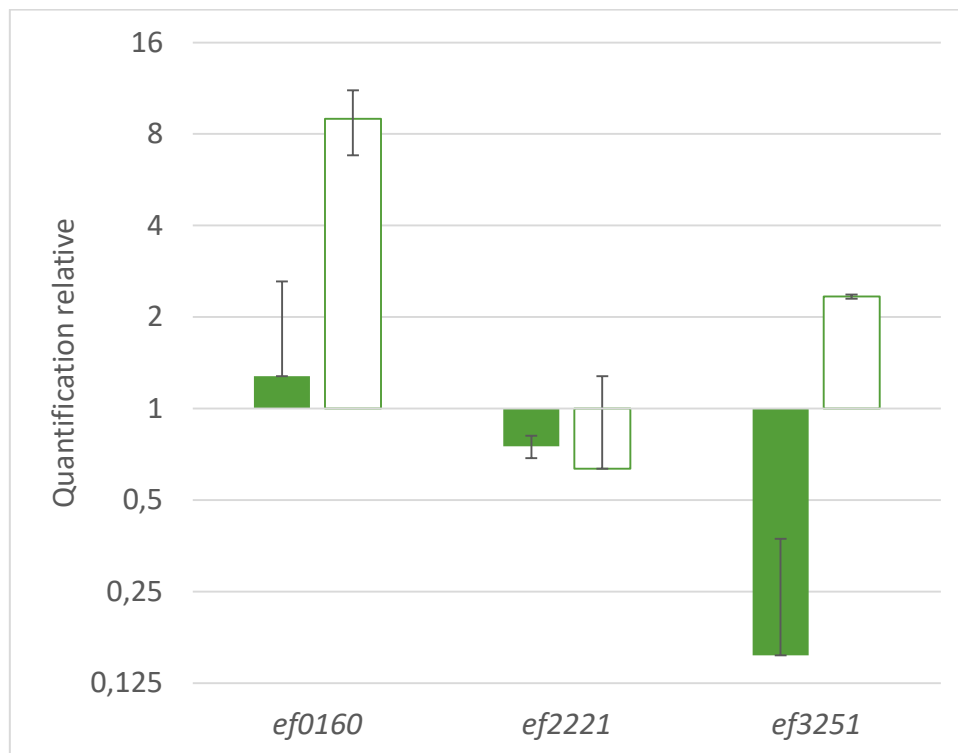


Figure 20 : Résultats de l'analyse de l'abondance des ARNm codés par les gènes *ef0160*, *ef2221* et *ef3251* par RT-qPCR, réalisée avec les ARN des souches  $\Delta$ SRC65 (barres pleines) et SCR65+ (barres vides). Quantification relative : rapport avec l'expression du gène de l'ARN 5S, puis avec les témoins respectifs des deux souches (V19 pour  $\Delta$ SRC65 et pAGEnt pour SCR65+).

Cependant, les sARN agissant en *trans* exercent généralement leur action au niveau post-transcriptionnel. Il n'est donc pas aberrant d'observer peu de variation au niveau transcriptionnel, bien que la liaison sARN-ARNm puisse affecter la stabilité de l'ARNm et ainsi diminuer sa quantité dans la cellule (Storz *et al.*, 2004; Waters et Storz, 2009). Il convient donc d'analyser l'expression de ces gènes au niveau post-transcriptionnel.

Pour cela, une expérience d'iTRAQ a été réalisée en partenariat avec le plateau PROTEOGEN, à partir de protéines totales (à l'exception des protéines sécrétées) extraites des souches V19 et  $\Delta$ SRC65 (Tableau 6).

*Tableau 6 : Liste des protéines détectées lors des expériences d'iTRAQ avec un facteur d'induction ou de répression ( $\Delta$ SRC65/V19) d'au moins 2, et pour lesquelles le test statistique (t-test) montre une p-value inférieure à 0,05.*

Nom	Produit	$\Delta$ SRC65/V19	P-value $\Delta$ SRC65/V19
EF0106	carbamate kinase	-2.11	0,026
EF1963	phosphoglycerate kinase	-2.17	0,011
EF2406	glycyl-tRNA synthetase subunit beta	-2.23	0,031
EF0210	30S ribosomal protein S19	-2.25	0,039
EF1613	formate acetyltransferase	-2.29	0,005
EF2718	50S ribosomal protein L1	-2.58	0,03
EF2398	30S ribosomal protein S2	-2.68	0,002
EF0284	(3R)-hydroxymyristoyl-ACP hydratase	-2.91	0,001
EF0916	50S ribosomal protein L20	-3.16	0,021
EF0022	PTS system mannose-specific transporter subunit IID	-3.44	0,026
EF0419	hypothetical protein EF0419	-3.56	0,017
EF3256	pheromone cAD1 lipoprotein	-4.29	0,006
EF0105	ornithine carbamoyltransferase	-5.81	0
EF0104	arginine deiminase	-6.10	0,008
EF1167	fructose-bisphosphate aldolase	-7.46	0,002

Au terme de cette expérience, 15 protéines dont l'expression est réprimée d'au moins un facteur 2 chez la souche  $\Delta$ SRC65 ont été identifiées ( $p < 0,05$ ). Aucune protéine avec une expression induite d'un facteur supérieur à 2, et associée à une p-value acceptable, n'a été mise en évidence.

La protéine EF0022, réprimée d'un facteur 3,44 était la seule protéine correspondant au produit d'un gène identifié *in silico* comme cible potentielle. Cette protéine est l'un des composants d'un système phosphotransférase (PTS) spécifique du mannose. On observe aussi la répression de trois protéines d'un même opéron (EF0104, EF0105 et EF0106) impliqué dans le métabolisme de l'arginine. Cependant il a été observé au laboratoire que ces gènes avaient une expression modifiée chez de très nombreuses souches mutantes (Giard *et al.*, 2001; Yan *et al.*, 2015; Ladjouzi *et al.*, 2015). Des protéines impliquées dans la traduction (protéines ribosomales, aminoacyl ARNt synthétase), la biosynthèse des acides gras (EF0284), ainsi qu'un précurseur de phéromone (EF3256) avaient aussi leur expression altérée. Les protéines EF1963 (phosphoglycerate kinase) et EF1167 (fructose-bisphosphate aldolase) impliquées dans la glycolyse présentaient également une différence d'expression significative.

Afin d'affiner la recherche de cibles potentielles de SRC65, nous avons réalisé une recherche *in silico* des cibles potentielles de son homologue SRC90. Nous avons ainsi pu révéler 42 cibles potentielles pour SRC90, qui feront l'objet d'une prochaine étude (Tableau 7).

Il est intéressant de noter que le gène en amont de SRC90, *ef1599*, a été détecté comme cible potentielle, bien que ce gène ne figure pas dans les cibles potentielles de SRC65.

Nous avons relevé trois gènes communs parmi les cibles potentielles de SRC65 et de SRC90 : *ef3251*, *ef2731* et *ef2165*, codant respectivement une protéine hypothétique, la protéine ribosomale L33 et une épimérase/déhydratase NAD-dépendante.

Tableau 7 : Cibles potentielles de SRC90 identifiées à l'aide du logiciel TargetRNA2 (p-value < 0,05).

Rang	Gène	Locus	Produit	Rang	Gène	Locus	Produit
1	-	ef1599	Cro/Cl family transcriptional regulator	22	<i>cyIM</i>	ef0527	CyIM protein
2	-	ef2132	RecT protein	23	-	ef1732	ABC transporter ATP-binding protein/peptidase
3	-	ef0964	hypothetical protein	24	-	ef0033	hypothetical protein
4	-	ef0266	chaperonin, 33 kDa	25	-	ef3170	HD/KH domain-containing protein
5	-	ef1456	hypothetical protein	26	<i>murAA</i>	ef2605	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1
6	-	ef3023	polysaccharide lyase 8	27	-	ef2165	NAD-dependent epimerase/dehydratase
7	-	ef2124	methyltransferase	28	-	ef2138	Cro/Cl family transcriptional regulator
8	-	ef3251	hypothetical protein	29	-	ef0871	cation transporter E1-E2 family ATPase
9	-	ef1745	hypothetical protein	30	-	ef3135	mannonate dehydratase
10	<i>rpmG-2</i>	ef2731	50S ribosomal protein L33	31	-	ef2603	PTS system transporter subunit IIA
11	-	ef1496	V-type ATPase subunit C	32	<i>dnal</i>	ef0883	primosomal protein Dnal
12	-	ef1879	hypothetical protein	33	-	ef3081	pheromone binding protein
13	-	ef2205	hypothetical protein	34	-	ef2357	hypothetical protein
14	-	ef2389	hypothetical protein	35	-	ef2672	hypothetical protein
15	<i>msrA</i>	ef1681	peptide methionine sulfoxide reductase	36	-	ef0887	glycosyl transferase group 2 family protein
16	-	ef1695	acetyltransferase	37	-	ef0161	hypothetical protein
17	-	ef0850	PemK family transcriptional regulator	38	<i>purK-1</i>	ef1786	phosphoribosylaminoimidazole carboxylase ATPase
18	<i>radC</i>	ef2926	DNA repair protein RadC	39	<i>thyA</i>	ef1576	thymidylate synthase
19	-	ef2851	hypothetical protein	40	-	ef1259	HAD superfamily hydrolase
20	-	ef2786	hypothetical protein	41	-	ef1010	sigma-54 factor interaction domain-containing protein
21	-	ef1537	integrase/recombinase XerD	42	-	ef2484	hypothetical protein

EF3251 a été analysée par les logiciels Interpro et Phobius qui permettent de détecter des domaines caractéristiques des protéines. Cette protéine ne présente pas de domaine spécifique, mais elle possède cependant un peptide signal et serait donc extracytoplasmique. Nous avons également observé qu'elle présente de fortes homologies avec des protéines non caractérisées chez le genre *Enterococcus* mais aussi chez *Listeria monocytogenes* et *Mycobacteroides abscessus* (99% d'identité sur 95% de couverture), *Streptococcus agalactiae* et *Streptococcus*

*pneumoniae* (29% d'identité sur 67% de couverture) et *Chlamydia trachomatis* (23% d'identité sur 98% de couverture). D'après la base de données DOOR (Dam *et al.*, 2007; Mao *et al.*, 2009), *ef3251* est transcrit en opéron avec *ef3250*, codant une autre protéine hypothétique contenant un domaine WxL, qui est un domaine de liaison à la paroi (Brinster *et al.*, 2007).

## Discussion

A l'aide d'analyses *in silico* et de techniques de biologie moléculaire, nous avons pu acquérir la certitude que SRC65 est un petit ARN agissant probablement en *trans* étant donné l'absence de variation des gènes flanquants chez les souches  $\Delta$ SRC65 et SRC65+.

Nous avons également mis en évidence une homologie importante entre SRC65 et un autre sARN, SRC90. Chez *Brucella abortus*, deux sARN homologues sont impliqués dans la virulence et la survie en macrophages au travers de fonctions régulatrices redondantes (Caswell *et al.*, 2012). Dans l'étude réalisée chez *B. abortus*, le double mutant de ces deux sARN homologues est affecté dans ses facultés de colonisation mais pas les simples mutants, ce qui pourrait expliquer l'absence de phénotype de croissance chez notre mutant SRC65. La réalisation d'un double mutant SRC65/SRC90 pourrait être intéressante. Cependant, malgré différentes tentatives, le mutant de SRC90 n'a pu être obtenu.

Nous avons utilisé plusieurs méthodes pour tenter de détecter la ou les cibles de SRC65 et SRC90. Nous avons commencé par des études *in silico* qui nous ont permis de repérer trois potentiels gènes cibles communs : *ef3251*, *ef2731* et *ef2165*. *ef3251* nous semble être le plus prometteur, puisque les données de RT-qPCR ont montré que l'ARN messager correspondant était moins abondant chez le mutant  $\Delta$ SRC65. La protéine correspondante n'a pas de fonction ni de domaine fonctionnel identifié à l'exception d'un peptide signal, ce qui signifie qu'elle est



probablement sécrétée. Cela pourrait expliquer le fait qu'elle n'ait pas été détectée lors des expériences d'iTRAq, puisque la technique d'extraction des protéines utilisée ne permettait pas de récupérer les protéines sécrétées.

Un gène semblable à *ef3251* est retrouvé chez toutes les espèces du genre *Enterococcus* mais aussi chez *L. monocytogenes*, *M. abscessus*, *S. agalactiae*, *S. pneumoniae* et *C. trachomatis*, qui sont des pathogènes connus de l'homme.

Il est intéressant de noter qu'*ef3251*, d'après la base de données DOOR, est transcrit en opéron avec un gène codant pour une autre protéine hypothétique, EF3250, porteuse d'un domaine WxL. Il s'agit d'un domaine de liaison à la paroi bactérienne présent chez de nombreuses bactéries à Gram positif (Brinster *et al.*, 2007). Il a récemment été démontré que des protéines d'*E. faecium* porteuses d'un tel domaine étaient capables de se fixer à la matrice extracellulaire, en particulier au collagène de type I et à la fibronectine, mais également entre elles (Galloway-Peña *et al.*, 2015). L'implication de ces protéines dans la résistance aux sels biliaires a également été mise en évidence (Galloway-Peña *et al.*, 2015).

Par ailleurs, il semble y avoir un lien entre SRC65 et la réponse au stress, puisqu'il a été observé que ce sARN est surexprimé en présence de sels biliaires, urine, sérum, et en condition de stress acide et stress oxydatif.

Nous pouvons également penser que SRC65 pourrait être impliqué dans le métabolisme des sucres. En effet, la protéine EF0022 détectée grâce aux expériences d'iTRAq et dont le gène était prédit *in silico* comme cible potentielle de SRC65, code pour un système PTS spécifique du mannose.

D'autres protéines en rapport avec le métabolisme des sucres ont également été relevées par les expériences d'iTRAq. EF1963 et EF1167 sont des protéines impliquées dans la glycolyse, mais aussi vraisemblablement dans la virulence : EF1963 est une phosphoglycerate kinase, exprimée

en surface et jouant un rôle dans l'activation des neutrophiles chez *S. pyogenes* (Uhlmann *et al.*, 2016) ; et EF1167 est une fructose-biphosphate aldolase également impliquée dans l'interaction avec le système immunitaire chez *L. monocytogenes* (Mendonça *et al.*, 2016), mais est également soupçonnée de lier le plasminogène chez *Neisseria meningitidis* (Shams *et al.*, 2016). Cependant, aucune de ces protéines n'est apparue comme cible potentielle de SRC65 *in silico*.

Nous avons également pu observer que la surexpression de SRC65 prolongeait la phase de latence lors de la croissance bactérienne, ce qui signifie que SRC65 serait probablement impliqué dans le déclenchement de la phase exponentielle de croissance. Il est possible qu'un tel phénotype soit lié à un défaut métabolique (répression d'une enzyme métabolique par exemple), ou encore à la réponse au stress. En effet, des liens entre modification de la phase de latence et réponse au stress ont été mis en évidence auparavant chez *Salmonella enterica*. Chez cette bactérie, le facteur sigma alternatif RpoS, connu pour son implication dans la réponse générale au stress, influence la durée de la phase de latence en condition de choc osmotique (Shiroda *et al.*, 2014).

De plus, il a été montré que des souches d'*E. coli* ayant acquis des mutations provoquant un allongement de la phase de latence développaient une tolérance accrue à l'ampicilline (Levin-Reisman *et al.*, 2017). Des manipulations préliminaires ont été entreprises, mais nous n'avons pas observé ce phénotype chez la souche surexprimant SRC65.

## **Chapitre 3 : Etude des mécanismes de régulation d'une 5'UTR possédant un terminateur de transcription**

Dans cette partie, nous nous sommes intéressés à un sARN putatif qui s'est révélé être une 5'UTR. Nous avons commencé par révéler l'existence de cette 5'UTR. La suite de l'étude est présentée sous forme d'un article qui sera prochainement soumis. Des résultats complémentaires que nous n'avons pas souhaité inclure dans l'article viendront clore ce chapitre.

### 1. Contexte : découverte de la 5'UTR

Dans ce chapitre, nous avons étudié la régulation de l'opéron *ef1515-ef1516* via sa 5'UTR, 5'1515, ainsi que le rôle des protéines EF1515 et EF1516 dans le métabolisme des sucres.

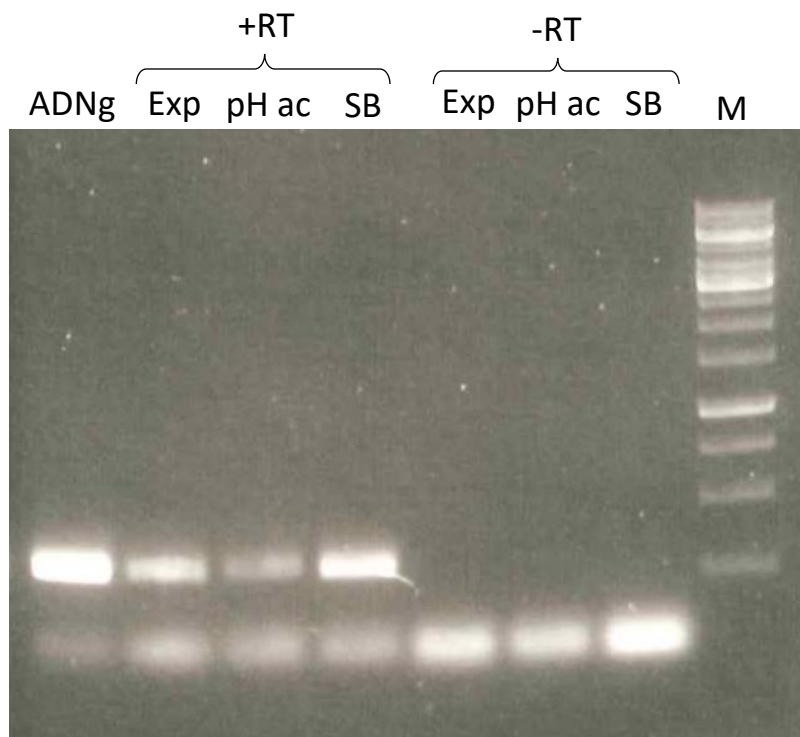
5'1515 a d'abord été détectée en tant que sARN putatif par des expériences de RNA-Seq (Muller *et al.*, 2015). D'après ces données, 5'1515 alors nommée SRC85, était surexprimée en condition d'infection intra-péritonéale chez la souris (Muller *et al.*, 2015). Nous avons pu observer que SRC85 était aussi surexprimé en condition de stress acide et en présence de sels biliaires (voir Chapitre I). Ce sARN putatif est localisé entre les gènes *ef1513* (annoté comme codant une protéine de liaison aux phéromones) et l'opéron *ef1515-ef1516* (annotés comme correspondants à un antitermineur de la famille Bgl et une protéine PTS spécifique du N-acétylglucosamine, respectivement).

Une expérience de 5'RACE-PCR ciblant SRC85 a confirmé la présence du site d'initiation de la transcription de SRC85 dans la région prédite par le RNA-Seq. Par 3'RACE-PCR, un petit transcrit a clairement été mis en évidence avec une extrémité 3' localisée 120 bases en aval de l'extrémité 5'. Cependant, cette même réaction de 3'RACE-PCR a montré l'existence d'un

autre ARN dont l'extrémité 3' se situe bien au-delà de celle du premier petit transcrit (données non montrées).

Afin de rechercher la présence de terminateur, une analyse *in silico* a été effectuée entre *ef1513* et *ef1515* au moyen du logiciel FindTerm. Cet outil a localisé un terminateur entre *ef1513* et SRC85, mais aucun n'a été détecté entre SRC85 et *ef1515* (voir Figure 1A de la publication « Characterization of a Bgl antiterminator in *Enterococcus faecalis* » ci-après, page 124).

Afin d'explorer la possibilité de co-transcription de SRC85 et *ef1515*, une RT-PCR a été réalisée au moyen d'amorces ciblant SRC85 pour l'une (*src85\_Forward*) et *ef1515* pour l'autre (*ef1515R*) (Tableau 3 et Tableau S1 de la publication « Characterization of a Bgl antiterminator in *Enterococcus faecalis* », Annexe 2, page 225) et en utilisant des ARN extraits de cellules d'*E. faecalis* V19 en phase exponentielle de croissance, en condition de pH acide ou en présence de sels biliaries. L'expérience a également été réalisée avec de l'ADN génomique comme matrice pour le témoin positif (Figure 21).



*Figure 21 : RT-PCR effectuées avec des oligonucléotides localisés sur SRC85 (src85\_Forward) et sur ef1515 (ef1515R). Chaque test a été réalisé avec (+RT) ou sans (-RT servant de témoin négatif) rétrotranscriptase, à partir d'ARN extraits en phase exponentielle de croissance (Exp), en condition de pH acide (pH ac) et en présence de sels biliaries (SB). De l'ADN chromosomique (ADNg) de la souche V19 a été utilisé comme contrôle positif. M : 1 kb Plus DNA Ladder (ThermoFisher).*

L'apparition d'une bande de taille égale au témoin positif (333 pb) a été observée pour les tests réalisés en présence de rétrotranscriptase (Figure 21). L'absence de bande pour les témoins négatifs montre que ces bandes sont bien dues à l'ADNc provenant de la rétrotranscription de l'ARN, et pas à une amplification d'ADN génomique résiduel.

De plus, une expérience de 5'RACE-PCR ciblant *ef1515* avec les oligonucléotides ef1515R, ef1515\_SP2 et ef1515\_SP3 a montré que le site d'initiation de la transcription d'*ef1515* était le même que celui de SRC85.

Ces données mettent en évidence une co-transcription de SRC85 avec *ef1515* dont le produit correspondrait au transcrit de plus grande taille observé lors de l'expérience de 3' RACE-PCR précédente. Ainsi, la région SRC85 serait donc très probablement la 5'UTR d'*ef1515*.

## 2. Publication 2 : Characterization of a Bgl antiterminator in *Enterococcus faecalis*

La publication suivante porte sur l'étude du rôle régulateur de la 5'UTR précédemment décrite, ainsi que sur l'implication des protéines EF1515 et EF1516 dans le métabolisme des sucres et la virulence. Cette publication est en phase finale de rédaction, et les manipulations pour répliquer les expériences et approfondir certains points sont en cours.

## Characterization of a Bgl antiterminator in *Enterococcus faecalis*

Marine Salze <sup>a</sup>, Cécile Muller <sup>a</sup>, Benoit Bernay <sup>b</sup>, Nicolas Sauvageot <sup>a</sup>, Thomas Clamens <sup>c</sup>, Olivier Lesouhaitier <sup>c</sup>, and Alain Rincé <sup>a</sup>

<sup>a</sup> Normandie Univ, UNICAEN, Unité de Recherche Risques Microbiens U2RM, Caen, France

<sup>b</sup> Normandie Univ, UNICAEN, SFR ICORE 4206, Proteogen platform, Caen, France

<sup>c</sup> Normandie Univ, University of Rouen, LMSM EA 4312 Laboratoire de Microbiologie Signaux et Microenvironnement, Evreux, France

### Abstract

Carbohydrate is essential for growth of the chemo-organotrophic bacterium *Enterococcus faecalis*. The metabolism in microorganism responsible for hospital-acquired infections is crucial for its colonization abilities. This includes carbohydrate transport by Phosphotransferase systems (PTS). In this study, we were interested in an operon encoding a PTS EIIABC subunit (EF1516) and a BglG/SacY family antiterminator (EF1515). The involvement of this operon into sugar metabolism was studied and revealed it was implicated in N-acetyl-glucosamine and D-gluconate metabolism. The characterization of the genomic environment has demonstrated the presence of a 5'UTR with regulatory functions on the operon through binding of the antiterminator EF1515. This BglG-like protein also regulates the expression of the gene *ef3023* which encodes a polysaccharide lyase and is shown to be involved in virulence.

## Introduction

*Enterococcus faecalis* is a ubiquitous, aero-anaerobic gram-positive bacterium that is most commonly found in mammal intestinal tract. While it is mostly harmless for immunocompetent subjects, this bacterium is a major hospital-acquired infection agent that can cause bacteremia, urinary tract infection (Murray, 1990; Rincé *et al.*, 2003) and that is responsible of almost 20% of bacterial endocarditis (Madsen *et al.*, 2017).

*E. faecalis* is present in the human gut microbiota at a small proportion but a nosocomial infection by this bacterium is often linked to an intensification of colonization in the gastrointestinal tract (Arias and Murray, 2012). As a chemo-organotrophic and homofermentative bacterium, *E. faecalis* relies on fermentation of carbohydrates as energy source (Brock *et al.*, 2007). Previous work has shown that *E. faecalis* metabolism and carbohydrate transport, including phosphotransferase systems (PTS), are crucial for its colonization abilities (Muller *et al.*, 2015). These protein complexes transport sugars into the cell while phosphorylating them, making them immediately available for metabolism. This occurs from a phosphorylation cascade, where the phosphoryl group is successively transferred from phosphoenolpyruvate (PEP) to Hpr, EI, and EII enzymes. This last enzyme assumes the role of transport and phosphorylation of the carbohydrate (Deutscher *et al.*, 2006).

PTS encoding genes can be regulated by the Bgl antiterminator that is a good example of sugar-dependant adaptation. In *Escherichia coli*, the *bgl* (for  $\beta$ -glucoside) operon is composed of three genes, *bglG*, *bglF* and *bglB* (Houman *et al.*, 1990; Amster-Choder, 2005). BglB is a phosphor- $\beta$ -glucosidase that hydrolyzes phosphorylated  $\beta$ -glucosides. BglG is a RNA specific antiterminator that unwinds a transcription terminator located in the 5' Untranslated Region (5'UTR) of the operon and allows transcription downstream of the terminator. BglF is a EIIABC subunit of a PTS transporter, able to import specific sugars ( $\beta$ -glucosides in this case), phosphorylates them during the transport, and consequently makes them available for the



processing enzymes. BglF also plays a sensor and regulator role, as it phosphorylates BglG in the absence of the specific sugar, making the antiterminator inactive (Amster-Choder, 2005; Harwani *et al.*, 2012).

Although the *bgl* system was firstly described in *E. coli*, systems that work on the same pattern are also present in gram-positive bacteria, like the *sacX-sacY* operon in *Bacillus subtilis* (Crutz *et al.*, 1990; Aymerich and Steinmetz, 1992). In this system, the antiterminator SacY is regulated by phosphorylation by SacX, a sucrose-specific PTS, and induces *sacB* gene expression with its antitermination activity (Idelson and Amster-Choder, 1998). In this system, SacY is an antiterminator regulating the expression of the gene *sacB* which encodes a levansucrase involved in the metabolism of sucrose and SacX is a saccharose-specific PTS protein which controls the activity of the SacY antiterminator by phosphorylation (Idelson and Amster-Choder, 1998).

To our knowledge, no *bgl* system was studied until now in *E. faecalis*, but genes encoding proteins of this family have been identified by homology in the genome of V583 strain. A RNA-seq experiment revealed a transcript corresponding to a non-coding region located upstream of *ef1515* gene and overexpressed during mouse peritonitis (Muller *et al.*, 2015). The *ef1515* and *ef1516* genes are annotated as encoding a BglG/SacY antiterminator and a PTS transporter, respectively.

In the present study, we showed that this region is in fact a 5'UTR, renamed 5'1515, and we highlighted its controls over the *ef1515-ef1516* operon. The role of each gene was investigated and the regulation mechanism of EF1515 identified. In addition, we found that this antiterminator controls the expression of EF3023, a potential virulence factor.

## **Material and Methods**

### ***Bacterial strains and growth conditions***

The reference strain used in this study is *E. faecalis* V19 (Zhao *et al.*, 2010), which corresponds to a plasmid-cured strain derived from the V583 strain of clinical origin (Paulsen *et al.*, 2003). Overnight cultures were achieved in M17 medium supplemented with 0.5% glucose (GM17). Cultures for physiological tests were carried out using a Model 680 Microplate reader (Bio-Rad, Hercules, Californie, USA) in carbon-cured M17 MOPS (M17cc MOPS) (Bizzini *et al.*, 2010), supplemented with 0,5% glucose, ribose, N-acetylglucosamine or D-gluconate and incubated at 37°C, or 30°C to maintain pLT06-derived plasmids. Media were supplemented with chloramphenicol (10 µg/ml), kanamycin (25 µg/ml) or ampicillin (100 µg/ml) when needed, and agmatin 40 mM was added for pAGEnt containing strains. *E. coli* strains TOP10, NEB-5α (New England BioLabs, Ipswich, Massachusetts, USA) and M15 pRep4 were used for RNA *in vitro* production, mutant constructions and recombinant protein synthesis, respectively. They were grown in LB medium at 37°C or 30°C for strains containing pLT06-derivatives.

### ***Molecular biology techniques***

5'RACE experiments were performed with the 5'/3' RACE kit, 2<sup>nd</sup> generation (Roche, Bale, Switzerland), using SP1 or EF1515R (retrotranscription), SP2 (PCR with anchor primers) and SP3 (sequencing) primers (Table S1) and poly-A or poly-C tailing. Phusion High-Fidelity DNA Polymerase (ThermoFisher, Waltham, Massachusetts, USA) and GoTaq DNA Polymerase (Promega, Madison, Wisconsin, USA) were used for the PCR. Purifications of PCR products were performed with NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel, Düren, Germany), and plasmid extractions were achieved using NucleoSpin Plasmid kit (Macherey-Nagel), following the manufacturer's recommendations.

### ***Construction of E. faecalis mutant strains***

All mutant strains ( $\Delta 5'1515$ ,  $\Delta ef1515$ ,  $\Delta ef1516$ ,  $\Delta ef3023$ ) were constructed in *E. faecalis* V19 using *E. coli* DH-5 $\alpha$  as an intermediate cloning host. For all constructions except  $\Delta ef1516$ , pLT06 vector (Thurlow *et al.*, 2009) was amplified with pLT06\_1\_bis and pLT06\_2 primers, and flanking regions of the region to be deleted were amplified by PCR using oligonucleotides 1 and 2 (SRC85 primers for  $\Delta 5'1515$ ; 1\_IVR and 2\_IVR\_bis for *in vivo* recombination method) for the upstream fragment, and oligonucleotides 3 and 4 (3\_IVR\_bis and 4\_IVR) for downstream fragment (Table S1). Primers 1 and 4 have overlapping tails compatible with pLT06\_1\_bis and pLT06\_2 primers, and primers 2 and 3 have overlapping tails compatible with each other. Cloning was performed using the *in vivo* recombination method (Huang *et al.*, 2017) or the NEBuilder HiFi DNA Assembly Cloning Kit (New England BioLabs). Deletion was obtained by double crossing over, as previously described (Thurlow *et al.*, 2009), and was verified by PCR with primers 5 (SRC85\_5 primer for *ef1515*) and 6 (Table S1) and sequencing. The  $\Delta ef1516$  mutant was constructed by insertion of the plasmid composed of the pUCB300 suicide vector containing an internal fragment of the *ef1516* gene beforehand amplified with 1516F\_BamHI and 1516R\_PstI primers (Table S1), and cloned into pUCB300 using the corresponding restriction enzymes. Vector insertion into V19 chromosome was checked with 1516\_verifF and 1516\_verifR primers (Table S1).

### ***Construction of a strain overexpressing 5'1515***

The pAGent vector (Linares *et al.*, 2014) used for the overexpression of 5'1515 was amplified using RPV\_AGent+1 and FRV\_AGent\_Term primers, without including the His-Tag and Shine-Dalgarno sequence (Table S1). Insert was amplified with FPI\_SRC85\_+1 and

RPI\_SRC85\_Term primers that possess plasmid-compatible overlapping tails (Table S1). Assembly was performed with NEBuilder HiFi DNA Assembly Cloning Kit (New England BioLabs) and the product was used to transform *E. coli* NEB-5 $\alpha$ . The resulting plasmid was introduced into *E. faecalis* V19, and the induction was performed in ribose-supplemented medium in presence of agmatin 40 mM.

### ***Total RNA extraction***

Cultures of 10 ml were performed at 37°C in M17 supplemented with the appropriate sugar to OD<sub>600</sub> 0.5 for non-stressing conditions. For stressing condition, strains were grown at 37°C to OD<sub>600</sub> 0.3, resuspended in GM17 supplemented with bile salts (0.06%) or in GM17 with pH adjusted to 5.8, and grown at 37°C for 30 minutes. Cells were pelleted and lysed using a FastPrep device (MP Biomedicals, Illkirch Graffenstaden, France). RNA extraction and purification were achieved with TRIzol Reagent (ThermoFisher) and chloroform/isoamyl alcohol separation before using Direct-Zol RNA Miniprep kit following the manufacturer's recommendations (Zymo-Research, Irvine, Californie, USA). RNAs were quantified using Nanodrop™ 2000 (ThermoFisher) and their quality was checked by electrophoresis in bleach agarose gel as previously described (Aranda *et al.*, 2012).

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

Reverse Transcription for RT-PCR and RT-qPCR assays were performed using QuantiTect Reverse Transcription kit (Qiagen, Venlo, Netherlands) with L/R oligonucleotides (Table S1). The GoTaq qPCR Master Mix (Promega, Madison, Wisconsin, USA) was used for quantitative PCR (qPCR), as well as the C1000™ Thermal Cycler (Bio-Rad) apparatus, using the following conditions: 95°C for 3 min, and 40 cycles at 95°C for 15sec, 60°C for 1 min. We used the *gyrA* (primers GyrAL and GyrAR) reference gene in order to normalize transcription levels for each

gene. Standard curves of each gene and PCR efficiency were obtained using genomic DNA of the V19 strain.

### ***Northern blotting***

Northern blot assays were performed using 10 µg of total RNA as previously described (Shioya *et al.*, 2011). The probe targeting 5'1515 was synthesized by PCR using src85\_Forward and SRC85\_SP1 primers (Table S1) and dATP- $\alpha^{32}$ P. Storage phosphor screens (Packard Instrument Company, Canberra, Australia) were used for exposition of hybridized membranes, and Cyclone (Packard Instrument Company) for the reading.

### ***In vitro production of RNA***

RNA synthesis was achieved with the pTOPO plasmid and *in vitro* synthesis. DNA regions of interest were amplified by PCR using primers containing tails overlapping with pTOPO: topo85\_FP1 and topo85\_RP1 for 5'1515, topo5'3023\_FP1/topo5'3023\_RP1 for 5'3023, and topo65\_FP1 and topo65\_RP1 for SRC65 (Table S1). Plasmid was amplified with topo\_FP2 and topo\_RP2 primers (Table S1). PCR products were inserted into the plasmid with NEBuilder HiFi DNA Assembly Cloning Kit (New England BioLabs) and products used to transform *E. coli* TOP10. RNA production was performed on the resulting plasmid, linearized with the *SpeI* restriction endonuclease, using MAXIScript™ T7 *in vitro* Transcription Kit (Invitrogen, Carlsbad, California, USA). Unincorporated nucleotides were eliminated by ammonium acetate/ethanol precipitation, as recommended on the kit protocol, and RNA were quantified with Nanodrop™ 2000 (ThermoFisher).

### ***Synthesis and purification of a recombinant EF1515 protein***

*ef1515* gene was amplified with primers *ef1515\_pQE70\_SphI* and *ef1515\_pQE70\_BglII* carrying restriction sites (Table S1). pQE70 plasmid and PCR products were digested with *SphI* and *BglII* enzymes, ligated together and used to transform *E. coli* M15 pRep4. The resulting construction was verified by sequencing. The bacteria obtained were grown at 37°C with agitation in Terrific Broth medium supplemented with kanamycin (25 µg/ml) and ampicillin (100 µg/ml) until OD<sub>600</sub> 0.5. Transcription induction was triggered with 0.5 mM IPTG during 4 h at 37°C under agitation. EF1515 protein (carrying His-Tag) was purified with Protino Ni-NTA Agarose kit according to the manufacturer's instructions (Macherey-Nagel), and desalted with PD10 columns (GE Healthcare, Chicago, Illinois, USA). Protein concentration was determined with the Bradford test (Bradford, 1976), and its purity checked on 12.5% SDS-PAGE.

### ***MicroScale Thermophoresis***

Recombinant EF1515 was marked onto its His-Tag using the His-Tag Labeling Kit RED-tris-NTA 2<sup>nd</sup> generation Monolith (Nanotemper Technologies, München, Germany) following the manufacturer's recommendations, and diluted in ES-Buffer (10 mM Tris pH8.0, 40 mM NaCl, 10 mM KCl, 1 mM MgCl<sub>2</sub>, 0.05% Tween-80) at a final concentration of 25 nM. Before the assay, RNA were heated for 5 min at 70°C and slowly cooled down at room temperature to allow a proper formation of secondary structures. RNA were diluted in a concentration range, mixed with the labeled protein, inserted into capillaries (Nanotemper Technologies) and introduced into the Monolith NT.115 device (Nanotemper Technologies). Data analyses were performed with NTAnalysis software V2.3 (Nanotemper Technologies).

### ***Virulence study on a *Galleria mellonella* model***

Infection assays were performed on *G. mellonella* larva as previously described (Benachour *et al.*, 2012). Bacteria were inoculated with a dose of  $2 \times 10^6$  UFC to 15 caterpillars per strain for each experiment. At least three experiments were performed for each condition. *G. mellonella* survival was followed from 16h post-infection and during 24h.

### ***API 50 CH assays***

Metabolism of carbohydrates was investigated using the API 50 CH strips (Biomérieux, Marcy-l'Etoile, France) and the API 50CHL medium according to the manufacturer's recommendations.

### ***Extraction of total proteins***

Cultures of strains V19 and  $\Delta 5'1515$  were performed to OD<sub>600</sub> 0.5. Cells were washed with Tris 50 mM pH8.0, 1 mM PMSF, and lysed with FastPrep (MP Biomedicals). Lysed cells were centrifuged and the supernatant was retrieved. Ten microliters of Nuclease Mix (GE Healthcare, Chicago, Illinois, USA) per milliliter of supernatant were added and the solution was incubated for 45 min at room temperature. Protein precipitation was performed overnight after addition of 4 volumes of a solution of 80% acetone and 20% Tris 50 mM pH8, 20 mM DTT, 5mM TCEP. Proteins were quantified with the Bradford method (Bradford, 1976).

### ***Sample preparation, digestion, iTRAQ labeling and Offgel fractionation***

Proteins were reduced, alkylated, digested with trypsin, labeled with 114 and 117 iTRAQ chemistry and then combined according to the manufacturer's instructions (AB Sciex, USA). After labeling, the pooled peptide samples were fractionated using a 3100 OFFGel fractionator according to the manufacturer's instructions for Peptide fractionation (Agilent Technologies,

Japan) using 13cm Immobiline DryStrip with pH 4-7 (GE Healthcare). After focusing, 12 samples were collected and reduced by evaporation in a vacuum concentrator.

### ***Sample preparation for Mass Spectrometry analysis***

For nano-LC fragmentation, protein or peptide samples were desalted and concentrated onto a 10 $\mu$ L C18 Omix tips (Agilent) before analysis.

The chromatography step was performed on a nano-LC system (Prominence, Shimadzu). Peptides were concentrated on a Zorbax 5 x 0.3mm C18 precolumn (Agilent) and separated onto a Zorbax 150 x 75 $\mu$ m C18 column (Agilent). Mobile phases consisted of 0.1% trifluoroacetic acid, 99.9% water (v/v) (A) and 0.1% trifluoroacetic acid in 99.9% ACN (v/v) (B). The nanoflow rate was set at 300 nl/min, and the gradient profile was as follows: constant 2% B for 5 min, from 2 to 5% B in 1 min, from 5 to 32% B in 144 min, from 32 to 70% B in 10 min, from 70 to 90% B in 5 min and return to 2% B in 10min. The 300 nl/min volume of the peptide solution was mixed with 1.2  $\mu$ L/min volumes of solutions of 5 mg/ml CHCA matrix prepared in a diluent solution of 50% ACN with 0.1% TFA. Twenty second fractions were spotted by an AccuSpot spotter (Shimadzu) on a stainless steel Opti-TOF™ 384 targets.

### ***Mass Spectrometry analysis***

MS experiments were carried out on a Sciex 5800 proteomics analyzer equipped with TOF TOF ion optics and an OptiBeam™ on-axis laser irradiation with 1000 Hz repetition rate. The system was calibrated before analysis with a mixture of des-Arg-Bradykinin, Angiotensin I, Glu1-Fibrinopeptide B, ACTH (18-39), ACTH (7-38) and mass precision was better than 50 ppm for reflectron mode. A laser intensity of 3400 was typically employed for ionizing. MS spectra were acquired in the positive reflector mode by summarizing 1000 single spectra (5  $\times$  200) in the mass range from 700 to 4000 Da. MS/MS spectra from the twenty most intense



ions were acquired in the positive MS/MS reflector mode by summarizing a maximum of 2500 single spectra ( $10 \times 250$ ) with a laser intensity of 4300. For the tandem MS experiments, the acceleration voltage applied was 1 kV and air was used as the collision gas. Gas pressure medium was selected as settings.

### ***Peptide sequencing and protein precursor identification***

The fragmentation pattern was used to determine the sequence of the peptide. Database searching was performed using the Mascot 2.6.1 program (Matrix Science). Two databases were used: a homemade list of well known contaminants (keratins and trypsin; 21 entries) and an *E. faecalis* V583 protein database (including 3264 entries).

The variable modifications allowed were as follows: methyl methanethiosulfonate modification of cysteines, K-acetylation, methionine oxidation, and dioxidation. “Trypsin” was selected and two miscleavages were also allowed. “iTraQ 4plex” was selected from “Quantitation”. Mass accuracy was set to 300 ppm and 0.6 Da for MS and MS/MS mode respectively. Mascot data were then transferred to Proline validation software (<http://www.profiroteomics.fr/proline/>) for data filtering according to a significance threshold of  $<0.05$  and the elimination of protein redundancy on the basis of proteins being evidenced by the same set or a subset of peptides. Validated proteins were quantified using Protein Pilot with the Paragon algorithm (Sciex) and with the following settings: trypsin digestion, methyl methanethiosulfonate modification of cysteines, iTRAQ 4-plex labeled peptides, default iTRAQ isotope correction settings, quantification, bias correction, biological modifications and thorough ID parameters selected. The detected protein threshold (unused protein score and confidence of results) was set to 95% and false discovery rate analysis was enabled.

## ***Bioinformatic analyses***

FindTerm, SoftBerry (Li, 2011) program was used for terminator prediction, with -11 energy threshold. Multalin (Corpet, 1988) and Clustal Omega (Sievers *et al.*, 2011) tools were used for DNA sequences alignments. To predict secondary structures, we used KineFold (Xayaphoummine *et al.*, 2005) software, with co-transcriptional fold simulation, pseudoknots allowed and non-crossing entanglements parameters. Signal peptides and protein localization were predicted with the Phobius web tool (Käll *et al.*, 2004), and protein predicting domains were investigated with the InterPro software (Mitchell *et al.*, 2018).

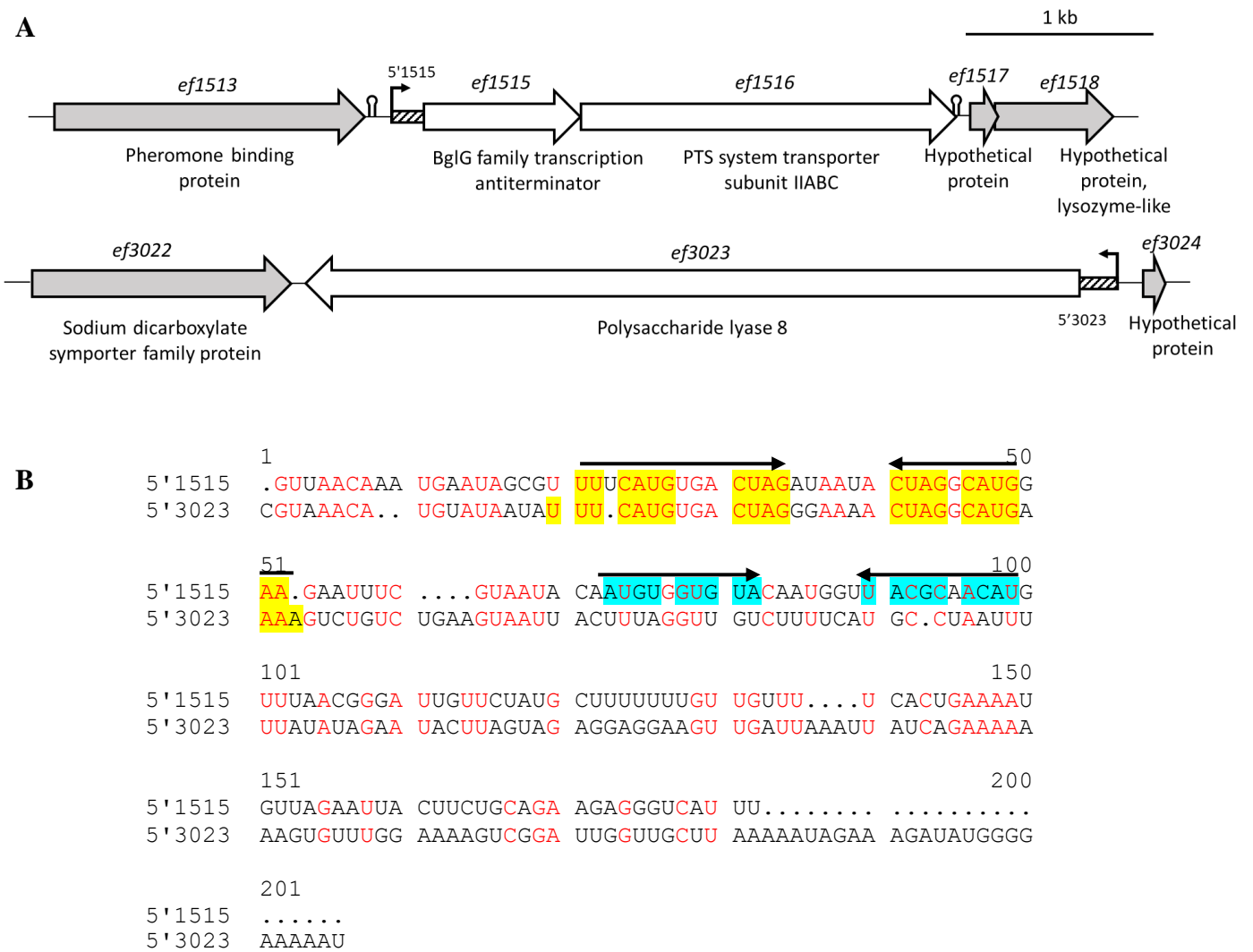
## **Results**

### ***Genetic environment of the ef1515-ef1516 locus***

The genomic environment of the *ef1515-ef1516* operon was firstly investigated to identify potential regulation mechanism.

According to the DOOR database, *ef1515* and *ef1516* genes belong to the same operon (Dam *et al.*, 2007; Mao *et al.*, 2009), and are respectively annotated as a Bgl family antiterminator and a N-acetylglucosamine (NAG) specific EIIABC PTS subunit.

To determine the transcription start site, a 5'RACE assay was achieved with ef1515R (as SP1), ef1515\_SP2 and ef1515\_SP3 primers (Table S1), and revealed the existence of a 172 bp long 5'UTR upstream of the predicted initiation codon, renamed 5'1515. The presence of this 5'UTR was confirmed by RT-PCR assays with forward primer (src85\_Forward), whose sequence corresponds to the nucleotides 49 to 71 of 5'1515, and reverse primer (ef1515R), whose sequence corresponds to the nucleotide 142 to 161 of *ef1515* ORF (Figure 1 and data not shown).



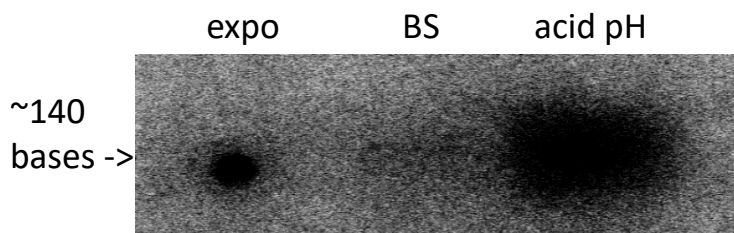
**Figure 1:** A: Genetic environment of *ef1515-ef1516* operon and *ef3023* gene in *E. faecalis* V19. 5'1515 and 5'3023 (5'UTR) are represented with a hatched bar. B: Alignment of 5'1515 and 3'3023 by Multalin and Clustal Omega programs, and their putative secondary structures detected by Kinefold software. Identical bases are represented in red. The different putative secondary structures are with arrows and highlighted with different colors. The blue highlighted region corresponds to the putative terminator on 5'1515.

Considering that EF1515 is a Bgl family antiterminator, and that in *E. coli*, the BglG protein regulates the expression of its own operon by unrolling a transcription terminator located in its 5'UTR (Amster-Choder, 2005), the Kinefold web tool was used to predict secondary structures in 5'1515. This allowed the identification of a potential terminator (free-energy: -32 kcal/mol, Figure 1B).

To study this UTR, northern blot assay was performed using total RNA extracted in glucose M17 in three different conditions: exponential phase, acid pH (pH 5.8) and bile salts 0.06%, conditions where 5'1515 is overexpressed according to previous RNA-Seq assay (Salze et al., submitted). RNA were hybridized with a labeled 5'1515-specific DNA synthesized by PCR using src85\_Forward and SRC85\_SP1 primers (Table S1) and corresponding to the nucleotide 49 to 110 of 5'1515.

The experiment showed the existence of a RNA of approximately 140 bases (Figure 2), confirming that 5'1515 contains a terminator that cause a premature transcription termination.

This assay also showed that this small RNA covered part of 5'1515 is strongly overexpressed in acid pH condition.



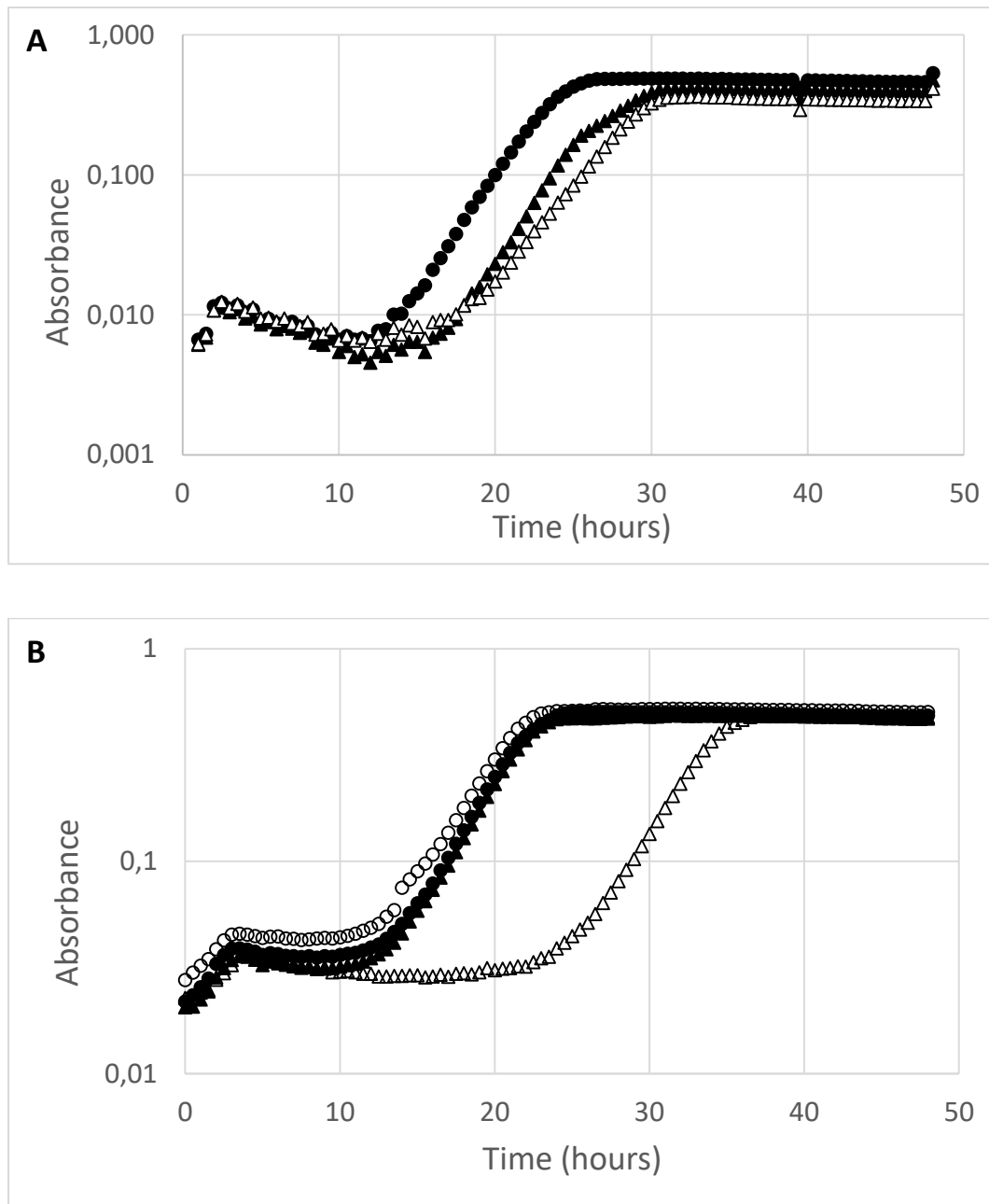
**Figure 2:** Northern blot assay targeting 5'1515, performed with RNA extracted from *E. faecalis* V19 strain in exponential growth phase (expo), after 30 min exposure to bile salts (BS) or pH adjusted at 5.8 (acid pH). Labeled probe was synthesized by PCR using src85\_Forward and SRC85\_SP1 primers.

### ***Role of ef1515-ef1516 in carbohydrate transport***

To characterize the *ef1515-ef1516* operon,  $\Delta 5'1515$ ,  $\Delta ef1515$  and  $\Delta ef1516$  mutants were constructed as described in the Material and Methods section. An overexpressing strain harboring a pAGEnt derivative and named 5'1515+, was also constructed to determine if 5'1515 had a regulatory RNA function.

The strains  $\Delta ef1515$  and  $\Delta ef1516$  were grown in M17cc MOPS supplemented with glucose or ribose, representing a PTS and a non-PTS transported sugar, respectively. We also performed a growth test in M17cc MOPS supplemented with NAG, considering EF1516 predicted function. Whatever the carbohydrate used, each strain grew at the same rate (data not shown), suggesting that this PTS protein does not transport NAG, or that there is another transporter that compensates for the loss of a functional EF1516 protein in the  $\Delta ef1516$  strain.

To test if EF1516 transports another sugar, we screened a large variety of sugar using API 50 CHL assays. These analyzes showed a difference of acidification in the presence of D-gluconate for the 5'1515+ strain compared to its control (empty pAGEnt containing V19 strain). Growth of the constructed strains was then followed into M17cc medium supplemented with D-gluconate, and showed a delayed exponential growth phase for  $\Delta ef1515$  and  $\Delta ef1516$  strains compared to V19 (Figure 3, A), that was even longer for the 5'1515+ strain (Figure 3, B). These results suggest that the *ef1515-ef1516* operon plays a role in the D-gluconate uptake. Moreover, we can not exclude that 5'1515 may have another distinct activity when the transcription is prematurely terminated.



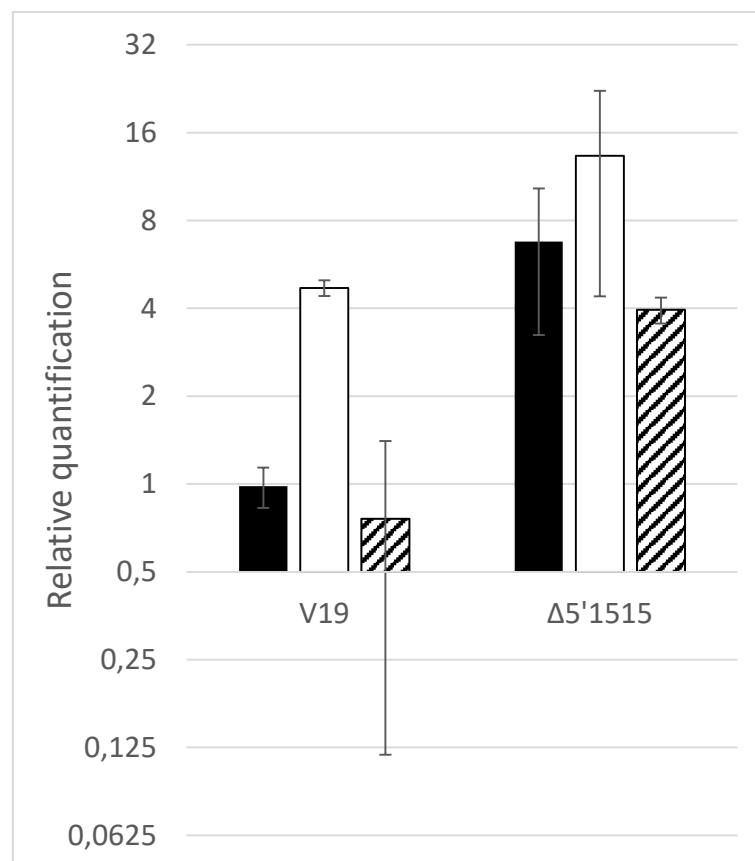
**Figure 3:** Growth of *E. faecalis* studied strains in M17cc MOPS supplemented with 0.5% D-Gluconate: (A) V19 strain (black circles),  $\Delta ef1515$  (black triangles), and  $\Delta ef1516$  (empty triangles), (B) V19 strain (black circles),  $\Delta 5'1515$  strain (empty circles), pAGEnt (black triangles), and  $5'1515+$  (empty triangles).

### *Induction of ef1515-ef1516 is sugar-dependent*

In order to understand how  $5'1515$  affects the expression of *ef1515-ef1516* operon, a  $\Delta 5'1515$  mutant was constructed. RT-qPCR analyzes were then performed to evaluate the expression of

*ef1515* and *ef1516* genes in this mutant and in the wild type strain grown in the presence of ribose, NAG or D-gluconate as sole carbon source.

The *ef1515* gene showed an overexpression in the  $\Delta 5'1515$  strain for all conditions (factors of  $6.7 \pm 2.5$ ,  $2.8 \pm 1.7$ , and  $8.4 \pm 7.6$ , for ribose, NAG and D-gluconate, respectively) (Figure 4) compared to the V19 strain and this was also the case for *ef1516* (data not shown). This result was confirmed at the proteomic level with iTRAQ experiment, where EF1516 was shown to be more abundant in the  $\Delta 5'1515$  background (1.7 fold higher than in V19 in ribose condition,  $p$ -value=0.0017). These results suggest that the  $5'1515$  could have a repressing effect on the expression of *ef1516-ef1516* operon, potentially through the transcription terminator mentioned above.



**Figure 4:** qPCR assay targeting *ef1515* and performed on RNA extracted from *E. faecalis* V19 or  $\Delta 5'1515$  strains grown on ribose (black), NAG (white) or D-gluconate (hatched) as sole carbon source.

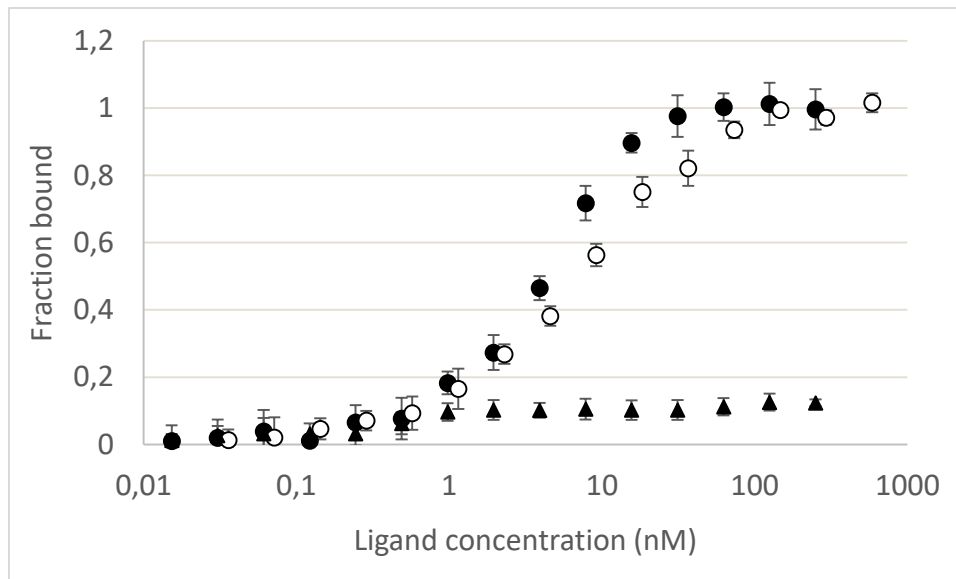
Surprisingly, we also observed that in V19, *ef1515* was overexpressed ( $4.9 \pm 1.05$  fold) in the NAG condition, but not in the presence of D-gluconate ( $0.84 \pm 0.78$  fold), comparatively to the ribose condition. This result seems inconsistent with the  $\Delta ef1515$  and  $\Delta ef1516$  growth defect observed in D-gluconate but not in NAG. It can be explained by the fact that the repression of *ef1516-ef1516* could be linked to an increase of the 5'1515 small RNA abundance. This increasing could lead to the growth phenotype similar to that of 5'1515+. This would support the hypothesis of another function of 5'1515.

### ***EF1515 regulates its own expression***

To look deeper in the regulation mechanism involving 5'1515, the direct interaction between the EF1515 protein and the 5'1515 RNA was studied by MicroScale Thermophoresis. A recombinant His-tagged EF1515 protein (EF1515-6xHis) was purified after cloning the EF1515 coding sequence into the pQE70 plasmid. This protein was labelled on the His-tag and incubated in presence of the 5'1515 RNA produced *in vitro*.

From this experiment, we observed a dose-response binding of EF1515-6xHis on 5'1515 RNA, with a Kd of 1.42 nM (Figure 5). A sRNA (SRC65, Muller *et al.*, 2015) was used as negative control and did not show any interaction with EF1515-6xHis. These results indicate that EF1515 has a high affinity for its 5'UTR and strongly suggest that it regulates both its own expression and that of EF1516 by interacting with 5'1515 and thus probably releasing the transcription terminator.





**Figure 5:** Binding of EF1515-6xHis to 5'1515 (black circles) or 5'3023 (empty circles) RNA (ligand). The same experiment was carried out using the SRC65 sRNA (Muller *et al.*, 2015) as negative control (black triangles). Error bars represent the standard error in triplicate measurements.

### *EF1515 directly regulates the ef3023 gene expression*

In order to find other targets of EF1515, the BLAST tool (Madden, 2003) was used to detect homologies with 5'1515 on the *E. faecalis* genome sequence. A sequence located upstream of the *ef3023* gene encoding a polysaccharide lyase shares 48% sequence identity (Clustal Omega software, based on sequences from transcription start to downstream gene start codon) with 5'1515 (including 88% sequence identity on 32 nucleotides) (Figure 1, B), with a terminator located at the 3' end of this sequence, detected by the FindTerm tool. This suggests that *ef3023* possesses a 5'UTR (named 5'3023) on which EF1515 could potentially bind to regulate the expression of the polysaccharide lyase. It is to be noted that a putative sRNA corresponding to the 5'3023 sequence was previously identified during RNA-Seq study (Salze *et al.*, under review), and corresponds to the sequence between 2,906,072 and 2,905,955 position on V583 strain genome (complementary strand).

The interaction between EF15156xHis and the 5'3023 RNA was then studied by MST (Figure 5). This assay highlighted a binding of the protein EF15156xHis on the 5'3023 RNA, but with a lower affinity than on 5'1515 RNA (5.16 nM vs 1.42 nM).

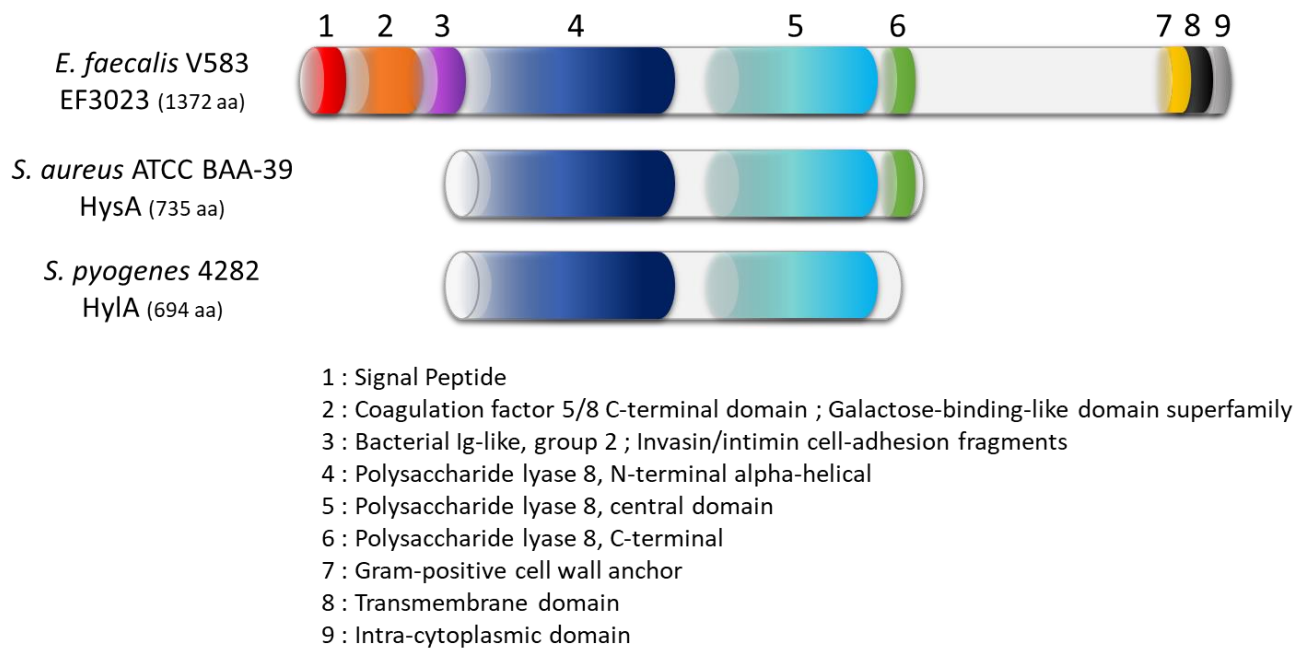
This result is consistent with the hypothesis that EF1515 also regulates the expression of *ef3023*, suggesting that EF1515, EF1516 and EF3023 belong to the same regulon.

### ***Role of ef3023***

EF3023 is annotated as a polysaccharide lyase, and is therefore part of the enzymes that degrade uronic acid-containing polysaccharides. This group of enzymes is divided into 29 families, and EF3023 is annotated in the 8th family which targets uronic acid-containing polysaccharides such as some glycosaminoglycans that are components of the extracellular matrix (hyaluronate, chondroitin or xanthan) (CAZy database, Lombard *et al.*, 2014). For now, the function and substrate of EF3023 are unknown, although it is annotated as hyaluronate lyase by KEGG database (Kanehisa *et al.*, 2017).

Using the BLAST tool, we compared EF3023 with two polysaccharide lyases from *Staphylococcus aureus* ATCC BAA-39 (HysA protein) and *Streptococcus pyogenes* 4282 (HylA protein) that share strong homology (29% identity, 64% cover and 28% identity, 46% cover, respectively). Firstly, EF3023 is twice bigger than the other two. Secondly, the Phobius and InterPro web tools showed that homologies between the three enzymes are narrowed to the three characteristic polysaccharide lyase domains (Figure 6). Indeed, the regions that are absent in *S. aureus* and *S. pyogenes* enzymes correspond to a signal peptide, a coagulation factor 5/8 C-terminal domain, an Ig-like domain, a Gram-positive LPxTG cell-wall anchor domain (commonly found in gram-positive bacterial surface proteins), a transmembrane domain and a small intra-cytoplasmic domain (Figure 6). This strongly suggests that the EF3023 protein is

mostly extra-cytoplasmic, and the coagulation and Ig-like domain are characteristic of the Microbial Surface Component Recognizing Adhesive Matrix Molecules (MSCRAMM) family, a class of adhesion proteins.

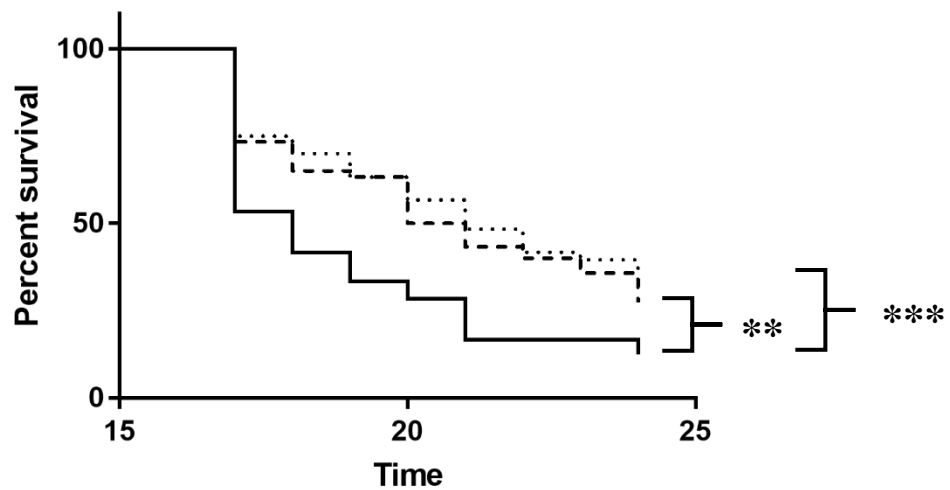


**Figure 6:** Signal peptide and domain prediction on EF3023, according to the Phobius and Interpro softwares. Proteins homologous to EF3023, HysA from *S. aureus* ATCC BAA-39 and HylA from *S. pyogenes* 4282, are also represented.

After growth tests in M17ccMOPS supplemented with 0.5% hyaluronic acid, we found that *E. faecalis* V19 is unable to grow with this compound as a substrate. According to previous studies, *S. aureus* and *S. pyogenes* can grow with this sugar as sole carbon source (Smith and Willett, 1968; Marion *et al.*, 2012). This could mean that EF3023 targets another polysaccharide, that its hyaluronidase function is defective or that another element of the corresponding metabolic pathway is absent or non-functional.

### ***Role in virulence***

To determine whether the genes studied here are involved in the virulence of *E. faecalis*, we infected *G. mellonella* larvae with V19,  $\Delta 5'1515$ ,  $\Delta ef1515$ ,  $\Delta ef1516$  or  $\Delta ef3023$  strains and monitored their survival (Figure 7). The experiment showed that the deletion of  $\Delta ef1515$  or  $\Delta ef1516$  does not affect larval mortality. However, larvae infected by  $\Delta 5'1515$  and  $\Delta ef3023$  strains showed a significant increased in survival relative to the parental V19 strain.



**Figure 7:** Survival curve of *Galleria mellonella* infected by *E. faecalis* V19 (line),  $\Delta 5'1515$  (dashed line) and  $\Delta ef3023$  (dotted line) strains. The experiment was performed at least three times using 15 caterpillars per strain and per test. \*\*:  $p$ -value  $<0.01$  comparatively to the V19 strain. \*\*\*:  $p$ -value  $<0.001$  comparing to the V19 strain (log-rank test).

### **Discussion**

For its development, *E. faecalis* needs to produce energy from carbohydrates fermentation. Thus, it has to use specific mechanisms to adapt to a new environment and to a new type of carbohydrate. In this study, we showed that the *ef1515-ef1516* operon is involved in sugar metabolism and is controlled by an antiterminator system. We observed that this operon possesses a 5'UTR, named 5'1515, which represses the operon expression probably through a

hairpin structure that causes a termination of the transcription. Indeed, a small transcript was observed in Northern blot assay and the whole operon was overexpressed in absence of 5'1515. We proved that the BglG family antiterminator EF1515 was able to bind this 5'UTR. In fact, the regulation of this operon probably works on the same model as the *E. coli bgl* system (Amster-Choder, 2005), where the PTS transporter gene (*ef1516* in our case) is transcribed with the antiterminator (EF1515), phosphorylates it in absence of the specific sugar, and consequently deactivates it.

We noted that the EF1516 PTS transporter was annotated as specific for NAG. We indeed observed an induction of the operon in presence of this sugar. However, neither the growth of the *ef1515* mutant nor that of  $\Delta$ *ef1516* was affected with NAG as sole source of carbon. We therefore deduced that NAG could enter the cell by another transporter.

To determine if EF1516 could transport other sugars, we screened a wide range of carbohydrates and we observed that the growth of the *ef1515* or *ef1516* mutants was affected when cultured with D-gluconate. Surprisingly, the operon expression was not induced in the presence of this sugar: it has to be noted that *E. faecalis* possess PTS shown to be D-gluconate specific (EF3136, EF3137, EF3138 and EF3139, Brockmeier *et al.*, 2009) that could be induced in the presence of this carbohydrate.

However, another hypothesis is that the phenotype observed in D-gluconate is not caused by a direct effect of the *ef1515-ef1516* operon, but by the small RNA resulting from a transcription termination at the 3'end of the 5'UTR. We indeed observed that the growth of *E. faecalis* was even more affected when 5'1515 was overexpressed. Its action could be through *trans*-coding sRNA activity, or protein sequestration (Waters and Storz, 2009).

The small RNA corresponding to the 5'1515 sequence might also be involved in virulence, as a lower larval mortality was observed when *G. mellonella* was infected by the  $\Delta$ 5'1515 strain,

but not when infected by  $\Delta ef1515$  or  $\Delta ef1516$  strain. This phenotype could be related to the growth phenotype. Indeed, D-gluconate metabolism can be linked to virulence, as this sugar was previously shown to be preferentially used by *E. coli* during colonization (Chang et al., 2004).

We have also found that EF1515 not only regulates the expression of its own operon, but also that of the *ef3023* gene. This latter gene possesses a 5'UTR, named 5'3023, which shares homology with 5'1515. In addition, we have proved that EF1515 binds 5'3023 which suggests that EF1515 have a regulatory function on the expression of *ef3023*. RT-qPCR assays using RNA from  $\Delta 5'1515$  strain did not show any variation of *ef3023* expression. Knowing that the *ef1515-ef1516* operon is overexpressed in  $\Delta 5'1515$  strain, we can say that the overexpression of the antiterminator does not affect *ef3023* expression.

*B. subtilis* possess 3 others Bgl/SacY RNA-binding antiterminators, and it has been shown that there are sequence similarities between the different RNAs that bind to the antiterminator in this bacterium. These sequences form a hairpin structure with two internal loops, which can correspond to antiterminator binding bases (Hübner et al., 2011). A previous study also showed that these structures form spontaneously (Clerte et al., 2013). We predicted that in their homologous regions, 5'1515 and 5'3023 are able to form similar hairpin structures with one internal loop and we strongly suspect this structure to be the EF1515 binding site (highlighted in yellow in the Figure 1B). There is another Bgl anti-terminator in *E. faecalis* V583, EF2599. This does not appear to be part of an operon, but its upstream gene, *ef2598*, is annotated as a  $\beta$ -glucoside-specific PTS transporter. It is to be noted that EF2599 and its binding region share a very strong homology with the BglG protein from *E. coli* (Mahadevan et al., 1987) and its 5'UTR. However, this 5'UTR doesn't share any homology with 5'1515.

The EF3023 protein is annotated as a polysaccharide lyase 8, and possesses characteristic domains of this group of enzymes that are known to specifically target hyaluronate. We observed that *E. faecalis* is unable to grow with hyaluronate as the only carbon source, even if EF3023 shares strong homologies with the known hyaluronate lyases *HysA* and *Hyl* respectively from *S. aureus* and *S. pyogenes*. *E. faecalis* protein is twice bigger than the other two, and possesses characteristic domains from another class of proteins: the MSCRAMM family. Moreover, Sillanpää *et al.* have observed that EF3023 possesses a cell-wall anchor domain and a ligand-binding site including an Ig-like domain, and aligned it with other MSCRAMM proteins from *E. faecalis* (Sillanpää *et al.*, 2004; Sillanpää *et al.*, 2009). However, it appeared that the EF3023 protein does not fit the typical model of MSCRAMM, as its Ig-folded region is shorter than the others and as a lower binding potential (<50%) than those of other MSCRAMM proteins (Sillanpää *et al.*, 2004). Therefore, EF3023 is probably a cell-wall protein with catalytic functions, although we have not identified its substrate yet.

We observed that EF3023 might be involved in *E. faecalis* virulence as its absence leads to a defect of the bacterium's ability to kill *Galleria mellonella* larvae. This agrees with previous studies in *E. faecalis* and *S. aureus*, where EF3023 and its homolog were suspected to be virulence factors (Makris *et al.*, 2004; Tsigrelis *et al.*, 2006).

### **Concluding remarks**

In this study, a regulation network involving EF1515 protein and two 5'UTR was highlighted. The fixation of this protein on 5'UTR located upstream *ef1515* and *ef3023* genes was evidenced. The regulated proteins, EF1515 itself, EF1516 and EF3023, play a role in both virulence and carbohydrate metabolism in *E. faecalis*. In depth study of its sugar metabolism may lead to the identification of new therapy targets against this hospital-acquired pathogen.

*Note* : Afin de privilégier la clarté du manuscrit, les données supplémentaires précédemment citées sont disponibles en Annexe 2.

## Acknowledgment

We are grateful to Lynn E. Hancock for providing the pLT06 plasmid.

## References

- Amster-Choder, O. (2005) The *bgl* sensory system: a transmembrane signaling pathway controlling transcriptional antitermination. *Current Opinion in Microbiology* **8**: 127–134.
- Aranda, P.S., LaJoie, D.M., and Jorcyk, C.L. (2012) Bleach gel: A simple agarose gel for analyzing RNA quality. *ELECTROPHORESIS* **33**: 366–369.
- Arias, C.A., and Murray, B.E. (2012) The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat Rev Microbiol* **10**: 266–278.
- Aymerich, S., and Steinmetz, M. (1992) Specificity determinants and structural features in the RNA target of the bacterial antiterminator proteins of the BglG/SacY family. *Proc Natl Acad Sci USA* **89**: 10410–10414.
- Benachour, A., Ladjouzi, R., Le Jeune, A., Hébert, L., Thorpe, S., Courtin, P., *et al.* (2012) The lysozyme-induced peptidoglycan N-acetylglucosamine deacetylase PgdA (EF1843) is required for *Enterococcus faecalis* virulence. *J Bacteriol* **194**: 6066–6073.
- Bizzini, A., Zhao, C., Budin-Verneuil, A., Sauvageot, N., Giard, J.-C., Auffray, Y., and Hartke, A. (2010) Glycerol is metabolized in a complex and strain-dependent manner in *Enterococcus faecalis*. *J Bacteriol* **192**: 779–785.
- Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248–254.
- Brock, T.D., Madigan, M.T., and Martinko, J.M. (2007) *Brock, Biologie des micro-organismes*. Pearson Education, Paris.
- Brockmeier, A., Skopnik, M., Koch, B., Herrmann, C., Hengstenberg, W., Welti, S., and Scheffzek, K. (2009) Activity of the *Enterococcus faecalis* EIIA<sub>gnt</sub> PTS component and its strong interaction with EIIB<sub>gnt</sub>. *Biochemical and Biophysical Research Communications* **388**: 630–636.
- Chang, D.-E., Smalley, D.J., Tucker, D.L., Leatham, M.P., Norris, W.E., Stevenson, S.J., *et al.* (2004) Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proceedings of the National Academy of Sciences* **101**: 7427–7432.



- Clerte, C., Declerck, N., and Margeat, E. (2013) Competitive folding of anti-terminator/terminator hairpins monitored by single molecule FRET. *Nucleic Acids Res* **41**: 2632–2643.
- Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* **16**: 10881–10890.
- Crutz, A.M., Steinmetz, M., Aymerich, S., Richter, R., and Coq, D.L. (1990) Induction of levansucrase in *Bacillus subtilis*: an antitermination mechanism negatively controlled by the phosphotransferase system. *J Bacteriol* **172**: 1043–1050.
- Dam, P., Olman, V., Harris, K., Su, Z., and Xu, Y. (2007) Operon prediction using both genome-specific and general genomic information. *Nucleic Acids Res* **35**: 288–298.
- Deutscher, J., Francke, C., and Postma, P.W. (2006) How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol Mol Biol Rev* **70**: 939–1031.
- Harwani, D., Zangoui, P., and Mahadevan, S. (2012) The  $\beta$ -glucoside (*bgl*) operon of *Escherichia coli* is involved in the regulation of *oppA*, encoding an oligopeptide transporter. *J Bacteriol* **194**: 90–99.
- Houman, F., Diaz-Torres, M.R., and Wright, A. (1990) Transcriptional antitermination in the *bgl* operon of *E. coli* is modulated by a specific RNA binding protein. *Cell* **62**: 1153–1163.
- Huang, F., Spangler, J.R., and Huang, A.Y. (2017) In vivo cloning of up to 16 kb plasmids in *E. coli* is as simple as PCR. *PLOS ONE* **12**: e0183974.
- Hübner, S., Declerck, N., Diethmaier, C., Le Coq, D., Aymerich, S., and Stülke, J. (2011) Prevention of cross-talk in conserved regulatory systems: identification of specificity determinants in RNA-binding anti-termination proteins of the BglG family. *Nucleic Acids Res* **39**: 4360–4372.
- Idelson, M., and Amster-Choder, O. (1998) SacY, a transcriptional antiterminator from *Bacillus subtilis*, is regulated by phosphorylation in vivo. *J Bacteriol* **180**: 660–666.
- Käll, L., Krogh, A., and Sonnhammer, E.L.L. (2004) A combined transmembrane topology and signal peptide prediction method. *J Mol Biol* **338**: 1027–1036.
- Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**: D353–D361.
- Li, R.W. (ed.) (2011) *Metagenomics and its applications in agriculture, biomedicine, and environmental studies*. Nova Science Publisher's, Hauppauge, N.Y.
- Linares, D.M., Perez, M., Ladero, V., Rio, B. del, Redruello, B., Martin, M.C., et al. (2014) An agmatine-inducible system for the expression of recombinant proteins in *Enterococcus faecalis*. *Microbial cell factories* **13**: 1–9.
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**: D490–495.

- Madden, T. (2003) *The BLAST Sequence Analysis Tool*. National Center for Biotechnology Information (US), . <https://www.ncbi.nlm.nih.gov/books/NBK21097/>. Accessed November 21, 2018.
- Madsen, K.T., Skov, M.N., Gill, S., and Kemp, M. (2017) Virulence factors associated with *Enterococcus Faecalis* infective endocarditis: A Mini Review. *Open Microbiol J* **11**: 1–11.
- Mahadevan, S., Reynolds, A.E., and Wright, A. (1987) Positive and negative regulation of the *bgl* operon in *Escherichia coli*. *J Bacteriol* **169**: 2570–2578.
- Makris, G., Wright, J.D., Ingham, E., and Holland, K.T. (2004) The hyaluronate lyase of *Staphylococcus aureus* – a virulence factor? *Microbiology* **150**: 2005–2013.
- Mao, F., Dam, P., Chou, J., Olman, V., and Xu, Y. (2009) DOOR: a database for prokaryotic operons. *Nucleic Acids Res* **37**: D459–463.
- Marion, C., Stewart, J.M., Tazi, M.F., Burnaugh, A.M., Linke, C.M., Woodiga, S.A., and King, S.J. (2012) *Streptococcus pneumoniae* can utilize multiple sources of hyaluronic acid for growth. *Infection and Immunity* **80**: 1390–1398.
- Mitchell, A.L., Attwood, T.K., Babbitt, P.C., Blum, M., Bork, P., Bridge, A., *et al.* (2018) InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res* (in press).
- Muller, C., Cacaci, M., Sauvageot, N., Sanguinetti, M., Rattei, T., Eder, T., *et al.* (2015) The intraperitoneal transcriptome of the opportunistic pathogen *Enterococcus faecalis* in mice. *PLOS ONE* **10**: e0126143.
- Murray, B.E. (1990) The life and times of the *Enterococcus*. *Clin Microbiol Rev* **3**: 46–65.
- Paulsen, I.T., Banerjee, L., Myers, G.S.A., Nelson, K.E., Seshadri, R., Read, T.D., *et al.* (2003) Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* **299**: 2071–2074.
- Rincé, A., Le Breton, Y., Verneuil, N., Giard, J.-C., Hartke, A., and Auffray, Y. (2003) Physiological and molecular aspects of bile salt response in *Enterococcus faecalis*. *International Journal of Food Microbiology* **88**: 207–213.
- 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**: e23948.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**: 539–539.
- Sillanpaa, J., Nallapareddy, S.R., Houston, J., Ganesh, V.K., Bourgoigne, A., Singh, K.V., *et al.* (2009) A family of fibrinogen-binding MSCRAMMs from *Enterococcus faecalis*. *Microbiology* **155**: 2390–2400.
- Sillanpaa, J., Xu, Y., Nallapareddy, S.R., Murray, B.E., and Höök, M. (2004) A family of putative MSCRAMMs from *Enterococcus faecalis*. *Microbiology* **150**: 2069–2078.

- Smith, R.F., and Willett, N.P. (1968) Rapid plate method for screening hyaluronidase and chondroitin sulfatase-producing microorganisms. *Appl Microbiol* **16**: 1434–1436.
- Thurlow, L.R., Thomas, V.C., and Hancock, L.E. (2009) Capsular polysaccharide production in *Enterococcus faecalis* and contribution of CpsF to capsule serospecificity. *J Bacteriol* **191**: 6203–6210.
- Tsigrelis, C., Singh, K.V., Coutinho, T.D., Murray, B.E., and Baddour, L.M. (2006) Vancomycin-resistant *Enterococcus faecalis* endocarditis: Linezolid failure and strain characterization of virulence factors. *Journal of Clinical Microbiology* **45**: 631–635.
- Waters, L.S., and Storz, G. (2009) Regulatory RNAs in bacteria. *Cell* **136**: 615–628.
- Xayaphoummine, A., Bucher, T., and Isambert, H. (2005) Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots. *Nucleic Acids Research* **33**: W605–W610.
- Zhao, C., Hartke, A., La Sorda, M., Posteraro, B., Laplace, J.-M., Auffray, Y., and Sanguinetti, M. (2010) Role of methionine sulfoxide reductases A and B of *Enterococcus faecalis* in oxidative stress and virulence. *Infect Immun* **78**: 3889–3897.

### 3. Résultats complémentaires

#### ***La régulation de l'opéron ef1515-ef1516 ne dépend pas de la répression catabolique***

Etant donné le rôle de l'opéron dans le transport des sucres, nous avons cherché à déterminer si l'expression de l'opéron *ef1515-ef1516* était sous la dépendance de la répression catabolique.

Une séquence très proche d'une boîte *cre* caractéristique chez *E. faecalis* (WTGWAARCGYWWWC) (Suárez *et al.*, 2011) a été détectée sur 5'1515, cette séquence (ATGAATAGCGTTTTTC) diffère du consensus décrit par l'insertion d'un T.

Afin de déterminer si cette boîte *cre* potentielle est effective ou non, une expérience de retard sur gel a été réalisée dans le but de déterminer si la protéine CcpA était capable de se fixer sur l'ADN 5'1515 (Figure 22). L'interaction a été effectuée en présence de la protéine Hpr phosphorylée sur sa sérine 46 (HPrP) qui est le co-facteur de CcpA, permettant sa liaison à l'ADN (Jones *et al.*, 1997).

Un retard dans la migration de l'ADN a été observé sur le gel dans les puits tests (mise en présence de 5'1515, CcpA et HPrP) et les puits témoins négatifs (témoins sans HprP ou avec de l'ADN compétiteur), ce qui signifie qu'il y aurait une fixation, mais qu'elle serait non spécifique.

HPrP		+	+	+	+	+						+	+		+
CcpA			1/10	1/20	1/40	1/80	1/10	1/20	1/40	1/80		1/10	1/20		1/10
ADN <sup>32</sup> P	5'1515	5'1515	5'1515	5'1515	5'1515	5'1515	5'1515	5'1515	5'1515	5'1515	T-	T-	T-	T+	T+

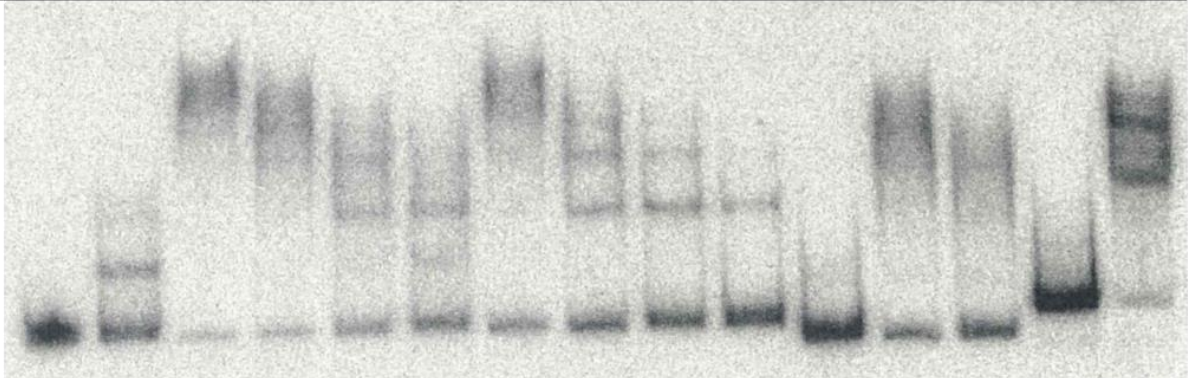


Figure 22 : Gel retard permettant d'étudier la fixation de CcpA sur 5'1515 en présence de son co-facteur HprP. Les témoins négatifs sans 5'1515 (5'UTR seule, ADN témoin seul et avec CcpA et HprP) et les témoins positifs sont représentés. T- : fragment interne du gène gusA d'E. coli ; T+ : région promotrice du gène ef0292.

### **Etude d'un facteur de virulence potentiel situé dans la même région génomique que l'opéron ef1515-ef1516**

Au cours de cette étude, nous nous sommes également intéressés à un gène situé dans la région génomique de l'opéron *ef1515-ef1516* : *ef1518*.

Ce gène code pour une protéine hypothétique contenant un domaine « lysozyme-like ». De telles protéines sont connues pour hydrolyser les liaisons  $\beta$ 1-4 du peptidoglycane, lui-même contenant du N-acetyl-glucosamine. Un lien avec l'opéron *ef1515-ef1516* est donc envisageable.

Un mutant de ce gène a été construit selon les mêmes méthodes que celles décrites dans le précédent article et au moyen des oligonucléotides ef1518-1b, ef1518\_2, ef1518\_3, ef1518\_4b, ef1518\_5, ef1518\_6, 1518\_Seq\_for et 1518\_Seq\_rev (voir Chapitre III.2 et Tableau 3).

Le mutant  $\Delta ef1518$  n'a montré aucun défaut de croissance par rapport à la souche sauvage, mais il semble être affecté dans sa virulence chez le modèle *G. mellonella* (Figure 23). Aucune variation de l'expression d'*ef1518* en absence de 5'1515 n'a pu être mise en évidence par RT-qPCR, ce qui signifierait que le gène *ef1518* n'est probablement pas régulé par l'antitermineur EF1515.

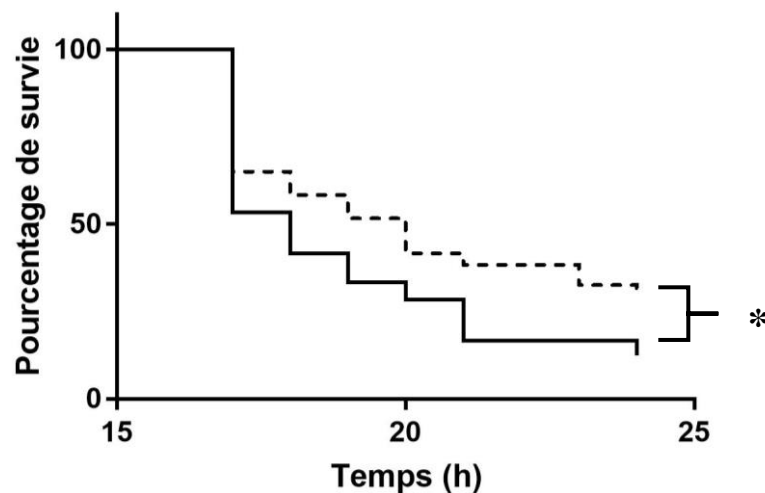


Figure 23 : Survie de larves *Galleria mellonella* infectées par *E. faecalis* V19 (ligne) ou par le mutant  $\Delta ef1518$  (pointillés). \* : différence significative entre les deux courbes ( $p$ -value < 0,05).

L'étude de la 5'UTR 5'1515, de l'opéron *ef1515-ef1516*, d'*ef3023* et d'*ef1518* a montré l'implication de ces régions génomiques dans la virulence d'*E. faecalis*. L'opéron *ef1515-ef1516* joue un rôle dans le métabolisme du NAG mais aussi dans celui du D-gluconate. La fonction exacte d'EF3023 et celle d'EF1518 restent encore à démontrer, mais leur implication certaine dans la virulence d'*E. faecalis* fait que leur étude mérite d'être approfondie. L'importance de 5'1515 dans la régulation de son opéron a été démontrée. Cette régulation s'opère *via* la liaison d'EF1515 sur 5'1515, et EF1515 agirait alors probablement en séparant des structures secondaires terminatrices de transcription. Cette action d'EF1515 s'effectue également sur la 5'UTR (5'3023) du gène *ef3023*. Nous avons donc ici deux exemples de

régulation par un ARN régulateur agissant en *cis*, mais il existe d'autres mécanismes impliquant des ARN de ce type.

## **Chapitre 4 : Caractérisation d'une hélicase à motif DEAD régulée par une région 5' non traduite (5'UTR).**

Dans ce chapitre, nous allons étudier l'action d'une 5'UTR sur la régulation de l'expression d'une hélicase à motif DEAD, fortement soupçonnée d'être impliquée dans le dégradosome chez *E. faecalis*. Nous nous intéresserons à cette fonction de la protéine, mais également à son implication dans la virulence et la réponse au stress. Les résultats sont présentés sous la forme d'un article qui est actuellement soumis pour publication dans Journal of Bacteriology. Des résultats complémentaires qui ne seront pas inclus dans l'article sont présentés en fin de chapitre.

1. Publication 3 : Study of RNA metabolism actors in *Enterococcus faecalis*



## Study of RNA metabolism actors in *Enterococcus faecalis*

Marine Salze<sup>a</sup>, Cécile Muller<sup>a</sup>, Benoit Bernay<sup>b</sup>, Thomas Clamens<sup>c</sup>, Olivier Lesouhaitier<sup>c</sup>, and  
Alain Rincé<sup>a</sup>

<sup>a</sup> Normandie Univ, UNICAEN, Unité de Recherche Risques Microbiens U2RM, Caen, France

<sup>b</sup> Normandie Univ, UNICAEN, SFR ICORE 4206, Proteogen platform, Caen, France

<sup>c</sup> Normandie Univ, University of Rouen, LMSM EA 4312 Laboratoire de Microbiologie  
Signaux et Microenvironnement, Evreux, France

### Abstract

RNase and DEAD-box helicases are key elements in RNA processing and are present in most organisms. They have been well-studied in bacteria like *Escherichia coli* and *Bacillus subtilis* but not in the hospital acquired pathogen *Enterococcus faecalis*. In this study, we got interested in the function of the four RNases (J1, J2, Y and III) and the three DEAD-box helicases (CshA, CshB, CshC) of this bacterium, and their involvement in bacterial stress response or virulence. RNase J1 was shown to be essential whereas, RNases J2 and III play a role in cold, oxidative and bile salts response, and RNase Y interferes in general fitness. It was also established that RNases J2, Y and III mutants have a reduced virulence in the *Galleria mellonella* infection model. Concerning DEAD-box helicases, all of them are involved in cold shock response. In our study conditions  $\Delta cshA$  was the most stress impacted mutant, and we focused on the DEAD-box helicase CshA. We demonstrated that CshA regulates its own expression by binding to a 5'Untranslated Region (5'UTR). We also showed that this protein would be involved in the degradosome function thanks to its privileged interaction with the enolase.

## **Importance**

*Enterococcus faecalis* is a bacteria able to cope with a wide range of environmental conditions that makes it a redoubtable opportunistic pathogen. This study demonstrates the importance of proteins involved in RNA metabolism, *i.e.* RNases J2, Y and III, as well as DEAD-box helicases CshA, CshB and CshC, in the general stress response or the virulence properties of the bacteria. RNase J1 was shown to be essential in *E. faecalis*. CshA, which autoregulates by a mechanism involving its binding to a 5' untranslated region, plays a prominent role in general fitness. This DEAD-box helicase interacts with the enolase through its C-terminus, suggesting a role of CshA in the degradosome.

## **Introduction**

*Enterococcus faecalis* is a gram-positive, facultative anaerobic and ubiquitous bacterium commonly found in mammal's gastrointestinal tract. This bacterium is also an opportunistic pathogen which is one of the major causes of hospital-acquired infections. It causes mainly urinary tract infections, but also bacteremia, endocarditis and meningitis (Babady, 2016). If several virulence factors, such as the hemolysin Cyl, the gelatinase GelE, or the general stress protein Gls24 (Arias and Murray, 2012; Madsen *et al.*, 2017) have been described, the ability of *E. faecalis* to adapt to a wide range of environments and its high resistance to stress are also important traits in the infection process. These require a tight control of the gene expression for a quick and efficient adaptation. Thus mRNA turnover is crucial for bacteria, and this role is performed by a multi-protein complex called the degradosome. In the gram negative bacteria *Escherichia coli*, the main proteins are the ribonuclease RNase E, a central enzyme of this complex, the polynucleotide phosphorylase (PNPase), the helicase RhlB and the glycolytic

enzyme enolase (Khemici and Carpousis, 2003). RNases, by their inherent endo- and/or exonuclease activities, are key enzymes of the degradosome.

In the majority of Gram positive bacteria, no RNase E homolog is found. In *Bacillus subtilis* and *Staphylococcus aureus*, the degradosome is centered around RNase Y, and the other main proteins are RNase J1, RNase J2, PNPase, the phosphofructokinase PfkA, the enolase, and the DEAD-box helicase CshA (Lehnik-Habrink *et al.*, 2010; Roux *et al.*, 2011). The complex is initiated by RNase Y, which is functionally homologous to RNase E, even if the two enzymes do not share important sequence similarity (Lehnik-Habrink *et al.*, 2011). RNase Y binds successively to the PNPase, RNase J1 and J2, and then to the other proteins of the complex (Lehnik-Habrink *et al.*, 2012).

RNases J1 and J2 are paralog enzymes that exert endo- and exo-ribonuclease activities. In *B. subtilis*: J1 is essential and is involved in many biological processes, while J2 is non-essential and has a weaker ribonuclease activity (Jamalli *et al.*, 2014). In *E. faecalis*, the RNase J2 was shown to play a role in virulence, being required for regulation of Ebp pili, but also involved in the regulation of virulence genes like *gls24* (Gao *et al.*, 2017).

Other RNases play critical roles in bacterial RNA processing and decay, like the RNase III, which is a double-strand specific endoribonuclease that regulates its own synthesis. Even so, its affiliation to the degradosome is still not defined. Contrary to RNases J1, Y and E, RNase III is able to process RNA with a triphosphate-end (Johnson *et al.*, 2011; Lioliou *et al.*, 2012). In *E. faecalis*, this RNase was shown to be involved in RNA-mediated regulation by cleaving small RNA-mRNA complexes (Johnson *et al.*, 2011).

DEAD-box RNA helicases are also important for RNA processing. They are indeed involved in almost every RNA metabolism from their biosynthesis to their degradation (Cordin *et al.*, 2006), and are able to modify the RNA-RNA and short DNA-RNA interactions (Khemici and

Linder, 2016). In the *B. subtilis* degradosome, the action of the RNA helicase CshA is required for depletion function: it unrolls secondary structures in the RNA, making it available for the RNases action (Lehnik-Habrink *et al.*, 2012). In *B. subtilis*, like all DEAD-box helicases, CshA is composed of two Rec-A-like domains, and a C-terminal domain essential for its interaction with the other degradosome proteins (Lehnik-Habrink *et al.*, 2010).

Recently, reports have been made about the involvement of DEAD-box helicases in gene regulation, stress resistance or virulence. In *Helicobacter pylori*, the degradosome DEAD-box helicase RphA was found to be involved in resistance to cold shock and in colonization (El Mortaji *et al.*, 2018). In *Pseudomonas aeruginosa*, the helicase DeaD regulates the translation of the *exsA* gene transcript by relaxing mRNA secondary structures (Intile *et al.*, 2015).

In this work, we studied the involvement of the main RNases (J1, J2, III and Y) and DEAD-box helicases (CshA, B and C) in *E. faecalis* physiology and virulence. We showed that CshA plays a critical role, for this reason we were particularly interested in the action and regulation of this DEAD-box helicase. We studied its interactions with degradosome proteins, and highlighted the enolase as its privileged partner. Finally, we showed that these seven proteins exert central functions, and are required for fitness and colonization in the pathogen *E. faecalis*.

## **Material and Methods**

### ***Bacterial strains and growth conditions***

In this study, we used *E. faecalis* V19 (Zhao *et al.*, 2010) as a reference strain, which corresponds to a plasmid-cured derivative of the V583 strain (Paulsen *et al.*, 2003). Cultures were achieved in BHI medium, supplemented or not with 0.06% bile salts, or 1 mM H<sub>2</sub>O<sub>2</sub>, or with pH adjusted at 5.7, and incubated at 37°C or 15°C. For growth experiments without

glucose, cultures were performed into M17 supplemented with ribose 0.5% (RM17). Chloramphenicol (15 µg/ml) was added for *E. faecalis* strains containing the pNZ273 or pMAD plasmids or derivatives. The *E. coli* strains TOP10, M15 pRep4 and DH5α were used respectively for *in vitro* RNA production, recombinant protein production and translational fusions. These *E. coli* strains were grown in LB medium at 37°C with agitation, supplemented with kanamycin (25 µg/ml) for the M15 pRep4 strain and pTOPO-containing strains, ampicillin (100 µg/ml) for strains containing pQE30 or pMAD derivatives or chloramphenicol (10 µg/ml) for strains containing the pNZ273 plasmid.

### ***Molecular techniques***

5'RACE-PCR were performed using the 5'/3' RACE kit, 2<sup>nd</sup> generation (Roche, Bâle, Switzerland), with poly-A or poly-C tailing. PCR were performed using Phusion High-Fidelity DNA Polymerase (ThermoFisher, Waltham, Massachusetts, USA), or GoTaq DNA Polymerase (Promega, Madison, Wisconsin, USA). PCR products were purified when necessary using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel, Düren, Germany) and plasmid extractions were achieved with the NucleoSpin Plasmid kit (Macherey-Nagel) according to the manufacturer's recommendations.

### ***Construction of E. faecalis strains***

*E. faecalis* V19 was used for the construction of all mutant strains ( $\Delta rnjA$ ,  $\Delta rnjB$ ,  $\Delta rnc$ ,  $\Delta rny$ ,  $\Delta cshA$ ,  $\Delta cshB$ ,  $\Delta cshC$ ) and *E. coli* DH5α for all cloning purposes. *E. faecalis* competent cells were produced as previously described (Bae *et al.*, 2002) except that bacterial treatment was performed with 10 µg/ml lysozyme and of 10 U mutanolysin. All deletion strains were generated with the same approach: the flanking regions of the gene were PCR amplified with primer pairs 1/2 (upstream fragment) and 3/4 (downstream fragment) that include restriction

endonuclease sites (All primers used in this study are presented in Table S2). The fragments were digested with the corresponding endonucleases and cloned in tandem into the pMAD vector (Arnaud *et al.*, 2004). The deletion of the gene on the genome of the *E. faecalis* V19 strain was then obtained by double crossing-over, as previously described (Aké *et al.*, 2011), and checked by PCR (with primers pairs 5/6 of each gene) and sequencing.

To prove that *rnjA* is essential in *E. faecalis*, the entire gene or an internal fragment were amplified using pAgent0846F/pAgent0846R and pUCB0846F/pUCB0846R primer pairs and were cloned into the plasmids pAGent (Linares *et al.*, 2014) and pUCB300 (Frère *et al.*, 1993), respectively, using the *NcoI* and *PstI* cloning sites. Once pAGent-*rnjA* was introduced into *E. faecalis* V19 strain, cells were cultivated in M17 supplemented with ribose 0.5% to avoid catabolic repression on the *PaguB* promoter and chloramphenicol (15 µg/ml). Selection of *E. faecalis* clones containing pUCB300- $\Delta$ *rnjA* was achieved with erythromycin (50 µg/ml).

For the construction of the TAP-tag fused to the *cshA* C-terminus, primer pairs 846TAPF1\_Bam/846TAPR2\_Bam (upstream fragment) and 846TAPF3\_XmaI/846TAPR4\_NcoI (downstream fragment) were used to amplify flanking region of the chosen site for insertion of TAP at the C-terminal of CshA (from V19 chromosome), and CBPF/PROTAR to amplify the TAP-tag (from the plasmid pILL825C), (Stingl *et al.*, 2008). The three fragments were sequentially introduced into pMAD using the restriction endonuclease *BamHI*, *XmaI* and *NcoI*. The construction was then inserted into the *E. faecalis* V19 chromosome, using the same method as the one used for the mutant constructions, giving the CshA-TAP strain.

### ***Constructions in the pNZ273 reporter plasmid***

pNZ273 plasmid was amplified by PCR with P0846\_FP2 and P0846\_RP2, for C3, and with P0846\_FP2\_C4\_C5 and P0846\_RP2 for C4 and C5 (Figure 4B). The inserts were amplified

with P0846\_FP1 and P0846\_RP1 for C3, and P0846\_FP1 and P0846\_RP1\_C4\_C5 for C4 and C5, giving overlapping tails compatibles with the plasmid. The assembly was performed *in vitro* using the NEBuilder HiFi DNA Assembly Cloning Kit (New England BioLabs, Ipswich, Massachusetts, USA), and the resulting DNA was used to transform *E. coli* NEB-5 $\alpha$ . For C5, a PCR was performed on C4 construction using 0846-2 and 0846-3 primers (Figure 4B) and the two ends of the product were digested by *Bam*HI and ligated. After replication of the plasmids in *E. coli*, the constructions were extracted and introduced into *E. faecalis* V19 (wild type) and  $\Delta$ *cshA* strains. All steps were controlled by PCRs with pNZ273\_MCS\_for and pNZ273\_MCS\_rev primers.

### ***$\beta$ -glucuronidase (GUS) activity assay***

The pNZ273 plasmid used for translational fusions includes the reporter gene *gusA*, encoding the  $\beta$ -glucuronidase (Platteuw *et al.*, 1994). Strains were cultured at 37°C in GM17 medium until OD<sub>600</sub> 0.7, then centrifuged. The pellets were resuspended in an equivalent volume of phosphate buffer, and OD<sub>600</sub> was measured. Five hundred  $\mu$ L (v) of the suspension was then mixed with 37.5  $\mu$ L of lysis buffer (Chloroform 66.7%, SDS 0.03%) and incubated at 37°C for 5 min. One hundred  $\mu$ L of p-nitrophenyl- $\beta$ -D-glucuronide (4 mg/mL) were added, and the mix was incubated at 37°C for 10 min (t). Two hundred and fifty  $\mu$ L of Na<sub>2</sub>CO<sub>3</sub> 1M were next added to stop the reaction, the solution was centrifuged 5 min, and the supernatant was used to measure the OD<sub>405</sub>. The GUS activity was expressed in  $\mu$ M/min/OD<sub>600</sub> and was determined using the following formula:  $GUS = (96 \times OD_{405}) / (t \times OD_{600})$ .

### ***Total RNA extraction***

Total RNA extractions were performed from 10 mL of culture for RT-qPCR, and 100 mL for Northern blot assay, of bacteria grown at 37°C until OD<sub>600</sub> 0.5. For stressing conditions, the

cultures were grown in RM17 medium until OD<sub>600</sub> 0.3 and centrifuged before being resuspended in stress-inducing medium (RM17 supplemented with 0.06% bile salts, or with pH adjusted at 5.7) and incubated 30 min at 37°C.

### ***RT-qPCR assay***

RNA extracted in stress conditions and RNA previously extracted from mouse infection (Muller *et al.*, 2015) were analyzed. The QuantiTect® Reverse Transcription kit (Qiagen) was used for RT reactions, using 1 µg of RNA and random primers, and qPCR assays were performed with the GoTaq qPCR Master Mix (Promega).

Reactions were loaded into a C1000™ Thermal Cycler (Bio-Rad, Hercules, California, USA) detection system (Bio-Rad) and the qPCR were performed with the following conditions: 95°C for 3 min, followed by 40 cycles at 95°C for 15sec, 60°C for 1 min. The transcription level of each gene was normalized to that of the *gyrA* reference gene. Genomic DNA from V19 cells was used to determine PCR efficiency.

### ***Northern blotting***

Northern blot assays were performed on 10 µg RNA as previously described (Shioya *et al.*, 2011). Hybridized membranes were exposed for one night to a storage phosphor screen (Packard Instrument Company, Canberra, Australia) and were read with a Cyclone apparatus (Packard Instrument Company).

### ***Tandem affinity purification and identification of interactome proteins***

*E. faecalis* V19 and CshA-TAP cells were collected in late exponential phase (OD = 0.8, 2 liters) and incubated 20 min at 37°C with 0.6% of formaldehyde for crosslinking or directly harvested and then washed once in 1% KCl, and stored at -80°C. The cells were lysed using the



FastPrep apparatus (MP Biomedicals). CshA interactome was purified as previously described (Redko *et al.*, 2016).

### ***Recombinant CshA production and purification***

The gene encoding CshA was amplified with pQE-846-F and pQE-846-R primers. The PCR product and the pQE30 plasmid (Qiagen, Venlo, Netherlands) were digested by *Bam*HI and *Pst*I restriction endonucleases, ligated and used to transform *E. coli* M15 pRep4. The strain was then grown at 37°C in Terrific Broth medium (Kram and Finkel, 2015) containing 25 µg/mL kanamycin and 100 µg/mL ampicillin until OD<sub>600</sub> 0.5, and the protein production was induced with 0.5 mM IPTG for 4 hours. Recombinant CshA (CshA-6xHis) was purified using the Protino Ni-NTA Agarose kit according to the manufacturer's instructions (Macherey-Nagel), and desalted with PD10 columns (GE Healthcare, Chicago, Illinois, USA). Proteins were quantified with the Bradford Method (Bradford, 1976), and analyzed by SDS-PAGE (12.5%). Nano-LC-MSMS analysis was performed on a NanoElute ultrahigh pressure nano flow chromatography system directly connected to a TIMS-TOF pro mass spectrometer with a modified nano electrospray ion source (Bruker Daltonics). Peptides were first concentrated onto a C18 pepmap 100 (5mm x 300µm i.d.) precolumn (Thermo Scientific) and separated at 50°C onto a reversed phase Reprosil column (25cm x 75µm i.d.; Ionopticks). Mobile phases consisted of 0.1% formic acid, 99.9% water (v/v) (A) and 0.1% formic acid in 99.9% ACN (v/v) (B). The nanoflow rate was set at 400 nl/min, and the gradient profile was as follows: from 2 to 15% B (60 min), followed by an increase to 25% B (30 min) and then to 37% (10 min), followed by a washing step at 95% B and reequilibration. MS spectra were acquired in the positive mode in the mass range from 100 to 1700 m/z in PASEF mode with exclusion of single charged peptides.

Database searching was performed using the Mascot 2.6.1 program (Matrix Science). An *E. faecalis* V583 protein database was used. The modifications allowed were as follows: C-Carbamidomethyl and M-oxidation. “Trypsin” was selected and two miscleavages were also allowed. Mass accuracy was set to 20 ppm and 0.05 Da for MS and MS/MS mode, respectively.

### ***RNA in vitro production***

DNA regions coding the RNA of interest were amplified using primers topo0846\_FP1/topo0846\_RP1 and topo65\_FP1/topo65\_RP1 for 5' *cshA* and SRC65, respectively, and the pTOPO plasmid (Zero Blunt TOPO™, Invitrogen, Carlsbad, California, USA) using primers topo\_FP2 and topo\_RP2. Inserts were independently ligated with the plasmid amplicon using the NEBuilder HiFi DNA Assembly Cloning Kit (New England BioLabs) and used to transform *E. coli* TOP10. After plasmid extraction and linearization with *XhoI* or *SpeI* for 5' *cshA* and SRC65, respectively, corresponding RNA were produced using the MAXIScript™ T7 *in vitro* Transcription Kit (Invitrogen).

### ***Microscale thermophoresis***

Recombinant CshA was firstly marked onto the His-Tag by using the His-Tag Labeling Kit RED-tris-NTA 2<sup>nd</sup> generation Monolith (Nanotemper Technologies, München, Germany) following the manufacturer's recommendations, and diluted at a final concentration of 30 nM in reaction buffer (20 mM Tris HCl pH 7.4, 150 mM KCl, 1 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, 0.01% Tween-80). Test and control RNAs were firstly heated for 5 min at 95°C and cooled down at room temperature for approximately 15 min to allow a proper formation of secondary structures. They were then diluted at 250 nM in reaction buffer. Labeled recombinant CshA and a series of dilution of tested RNA or control RNA ranged from 0.03 to 250 nM were then mixed together in standard capillaries (Monolith NT.115 Capillary) and incubated 5 min at room

temperature. Thermophoresis was measured at room temperature using a Monolith NT.115 instrument (Nanotemper Technologies). Data analyses were performed with NTAnalysis software V2.3 (Nanotemper Technologies).

### ***Virulence study on the *Galleria mellonella* model***

*G. mellonella* larval infections were performed as previously described (Benachour *et al.*, 2012), with an inoculum of  $2 \cdot 10^8$  bacteria (OD 0.8) per larvae. Fifteen caterpillars were used by analysis and the experiments were repeated at least 3 times.

### ***Bioinformatic analyses***

For terminator prediction, we used the program FindTerm, SoftBerry (Li, 2011), using default parameters. Secondary structures were predicted using the KineFold (Xayaphoummine *et al.*, 2005) tool with co-transcriptional fold simulation, pseudoknots allowed and non-crossing entanglements parameters.

## **Results**

### ***Phylogeny of RNases and DEAD-box helicases in Enterococci***

We initiated the study of the 4 main RNases (J1, J2, Y and III) and 3 DEAD-box helicases in *E. faecalis* V583 (Table 1 and Figure S1). For each protein, we used the coordinates of the genes from NCBI to determine the protein sequence, except for CshA for which we used the TTG start codon on position 805,171 on the V583 genome (TTG). This latter is indeed associated with a ribosome binding site motif (AGGAGG) and conserved in enterococci contrary to the initiation codon indicated in the NCBI annotated sequence.

The Uniprot database was used to compare the sequence of these 7 proteins with those of 212 entries for enterococci, including 54 strains of *E. faecalis*. Each protein was found in 97% to

100% of enterococcal strains (Table S2). To date and based on available data, we observed that the RnjB and CshA are present and conserved in all enterococci indexed in the database.

**Table 1:** Features of the RNases J1, J2, Y and III, and the DEAD-box helicases CshA, B, C of *E. faecalis* V19.

Protein name	Gene name (ORF number in V583)	Protein size by number of amino-acids (and kDa)	Characteristics
RNase J1	<i>rnjA</i> (EF2924)	561 (62.9 kDa)	39% identity with RNase J2
RNase J2	<i>rnjB</i> (EF1185)	562 (61.2 kDa)	Role in virulence (mice), fitness, bile salts resistance. Involved in pilin, <i>gls24</i> and <i>pyrR</i> expression (in OG1RF strain) (Gao <i>et al.</i> , 2010; Gao <i>et al.</i> , 2017).
RNase Y	<i>rny</i> (EF3170)	518 (58.1 kDa)	
RNase III	<i>rnc</i> (EF3097)	230 (26.2 kDa)	
CshA helicase	<i>cshA</i> (EF0846)	515 (57.2 kDa)	DEAD motif, C-terminal extension, 39% identity with CshB, 30% identity with CshC.
CshB helicase	<i>cshB</i> (EF1377)	449 (51.2 kDa)	DEAD motif, 29% identity with CshC
CshC helicase	<i>cshC</i> (EF1009)	433 (49.0 kDa)	DEVD motif

Among other genera, the 7 proteins were also found in *Carnobacterium*, *Clostridioides*, *Tetragenococcus*, *Vagococcus* genus and in *Listeria monocytogenes*, with more than 50% of identity, (Table S2). Also, the RNases J1 and J2 genes, *rnjA* and *rnjB* respectively, share homologies with *B. subtilis* enzymes genes (70% and 47% sequence identity respectively).

In *E. faecalis* V583, the 3 DEAD-box helicases have similar sequences (Figure 1) especially in specific conserved domains of these proteins, except for CshA that have a C-terminal extension of more than 65 amino-acids compared to CshB and CshC.

	<b>Q</b>	<b>I</b>	
CshA	-LKFKELG <sup>Q</sup> LAP <sup>Q</sup> ELL <sup>Q</sup> SAVERAG <sup>Q</sup> FEEATPI <sup>Q</sup> SETIPLALAGKDVIG <sup>Q</sup> QAQTGTGKTA <sup>Q</sup> AFGLPM		59
CshB	MPSFKQFQ <sup>Q</sup> FPFINEALAEKGFEEPT <sup>Q</sup> EVQ <sup>Q</sup> EKLIPIIKKGKSVIG <sup>Q</sup> SQTGSGKTH <sup>Q</sup> TFFLLPL		60
CshC	-MRFLE- <sup>Q</sup> LP <sup>Q</sup> EAW <sup>Q</sup> E <sup>Q</sup> W <sup>Q</sup> AS <sup>Q</sup> GFNEPSTI <sup>Q</sup> EKSFTPLREGENVLG <sup>Q</sup> ISPTGTGKTLAYMLPL		58
	* : : . **:* : :*. : : * : . : * : * : * : * : *		
	<b>IA</b>	<b>IB</b>	
CshA	LNKIDADNRVIQGLVIA <sup>IA</sup> PTRELAI <sup>IA</sup> QTQEELFRLGR--DKKIRVQAVYGGADIGRQIRGLK		117
CshB	MDKVKPTIDEVQIVITAPSRELANQ <sup>IA</sup> IYQEAQQLARFSQPEIRVSNFVGGTDKQRQLNKLK		120
CshC	LLTVEKG-QGNQLLI <sup>IA</sup> APSQELAMQ <sup>IA</sup> IAEVARTWAK--PLQLKVQTLIGGANVSRQIDKLK		115
	: . . . * : : * : * : * * * * : : . : : * : . * : * : * : *		
	<b>IC</b>	<b>II</b>	
CshA	-DRPHIVVGT <sup>IC</sup> PGRLLDHINRRTLKLETVETLVL <sup>IC</sup> <u>DEAD</u> EMLNMGFLEDIEKIISQVPEQRQ		176
CshB	HQQPHVVIGT <sup>IC</sup> PGRILDMMNEQALKVHTAF <sup>IC</sup> AFV <sup>IC</sup> <u>DEAD</u> MTLDMGF <sup>IC</sup> LAEVDQIAGRLPEKLQ		180
CshC	-KRPEVLIGT <sup>IC</sup> PGRILELMKNKKVKAQLLKTIVM <sup>IC</sup> <u>DEVD</u> QLFQEEELSLTKQIL <sup>IC</sup> THTPTEYQ		174
	. : * : . : * : * : * : * : * : . : * : * : * : * : * : * : *		
	<b>III</b>		
CshA	TLLFSATMPPAIKNIGVKFMKQPEHVKIKAKEMTADLIDQYYVRSKDFEKFDIMTRLLDV		236
CshB	MLVFSATIPEKLRPFLK <sup>III</sup> KYLENPVIEHIKPKAVISETIDNWLISTKGKNSNKIIYQLLTI		240
CshC	LVFYSATADRVVNQAQ-SLSQKLQVIDVTEEDTSAGQVAHYFIRLAPRKKADYLRLLAHT		233
	. : * : * * : : . : : : : : : : : : : : . : . : * :		
	<b>IV</b>	<b>IVA</b>	<b>V</b>
CshA	QTPELTIVFGRTKRRVDELARGLEARGYKAEGIHG <sup>IV</sup> DL <sup>IV</sup> SQ <sup>IV</sup> KRMSVLR <sup>IV</sup> SFKNGNLDILVAT		296
CshB	GHPYLAIVFANTKQRVDEITDYLKDQGLKVAKIHG <sup>IV</sup> DITPRERKRVMRQVQNL <sup>IV</sup> DYQYVVAT		300
CshC	-EAFRSMVFFNQVADLGAAEEKLVYENVPAIGLASDQSKQLRKLAI <sup>IV</sup> DQFKAERVKLLLT		292
	: : * * . : . * . . : . * : : * : . : . : . : * :		
	<b>VA</b>	<b>VI</b>	
CshA	DVAARGLDISGVTHVYNYDIPQDPESYVHRI <sup>VI</sup> GR <sup>VI</sup> TGRAGKGGMSVTFVTPNEMG-YLHVIE		355
CshB	DLAARGIDIEGVSHVINA <sup>VI</sup> AEVPHELDFFIHRVGR <sup>VI</sup> TGRNGLNGTAITLYSPADDE-AITQIE		359
CshC	DIAARGLDFTGVPYVNVVDVPLTEESYIHRSGRVGRMGAQGAVITFINDGTKRDYQRLMK		352
	* : * * * : * : * : * : * : * : * : * : * : * : * : * : * : *		
CshA	NLTKKRMTPLRPPTEKEAFKQLSAAVETIESDLEENGLDKYLESAEALLEKYSATDLAA		415
CshB	QLGVSF----KPK---EIKNGEIVE---TYDR---NRRTKRE-----KSREELDP		396
CshC	QLAIPY-----QEIFLYDGALHEQPKAKD-----ELEKKNVVKD <sup>VI</sup> TYSEKKSTKERKP		399
	: * : : . * : : : : * : * : *		
CshA	LLLKTVAKDPADAVP-----VKITP-----ERPLPSNKK-----GFNKN---GRGGGGG		456
CshB	TLIGLVKKKKKKIKPGYKKKINWAIEDSNKKK <sup>VI</sup> KIERRQKDRSARKAKKNSSN-----		449
CshC	ETVNSVVKEPQET-----RKKKKK <sup>VI</sup> KNKTKNKGARKAKK-----		433
	: * * . : : . * : *		
CshA	NNRNRNRGGNGKGGSYRGNKNHKDGDRNYKNSDRKNKDQRGGK <sup>VI</sup> RHNDKKRGFVIRENAK		515
CshB	-----		449
CshC	-----		433

**Figure 1:** Alignment of *E. faecalis* DEAD-box helicases. The alignment was performed using CLUSTAL O (1.2.4). The conserved motifs of helicases (Cordin *et al.*, 2006) are highlighted and the DEA/VD motifs are underlined (motif II). “\*”: fully conserved amino-acid. “:”: conservation between groups of strongly similar properties. “.”: conservation between groups of weakly similar properties.

### ***Mutant constructions: rnjA is essential for E. faecalis***

The first step to investigate the role of the different selected actors was to construct a deletion mutant for the RNases J1, J2, III and Y, as well as for the 3 DEAD-box helicases CshA, CshB and CshC. The genes were deleted as described in Material and Methods. We obtained a mutant for *rnjB*, *rnc*, *rny*, *csaA*, *csbA* and *csbC*, demonstrating that these functions are independently dispensable in *E. faecalis* V19. Different attempts to construct a *rnjA* mutant in *E. faecalis* V19 or JH2-2 were unsuccessful suggesting that this gene is probably essential for *in vitro* growth. In order to demonstrate that *rnjA* is essential, *E. faecalis* strain V19 was first transformed by stably replicating plasmid pAGent (carrying resistance to chloramphenicol), expressing *rnjA* thanks to the *PaguB* agmatine inducible promoter. In parallel, we cloned an intern fragment of the *rnjA* gene in the suicide plasmid pUCB300 (pUCB- $\Delta$ *rnjA* – resistance to erythromycin), firstly in *E. coli* DH5 $\alpha$  and then we introduced pUCB300 or pUCB- $\Delta$ *rnjA* in *E. faecalis* V19 in presence of erythromycin and/or chloramphenicol, and with or without agmatine. The insertion of the pUCB- $\Delta$ *rnjA* plasmid in the chromosomal copy of *rnjA* was monitored thanks to the antibiotic pressure. No erythromycin resistant clones were obtained with the empty plasmid pUCB300 in the presence or absence of agmatine that showed no recombination at another place on the chromosome. When the V19 strain carrying the pAGent-*rnjA* was transformed with pUCB- $\Delta$ *rnjA*, we obtained between 200 and 300 clones per  $\mu$ g of DNA in the presence of agmatine, and no recombinant without the inducer. This demonstrated that *rnjA* is an essential gene in *E. faecalis* V19.

### ***Roles in growth and stress response***

To determine if mutations have any effect on *E. faecalis* growth, the different mutants were grown in Brain Heart Infusion (BHI) medium at 37°C. The growth of the mutants was not

significantly different from that of the wild-type strain, with the exception of  $\Delta cshA$  which had a doubling time of 50.3 min *versus* 42.3 min for the V19 strain (Table 2).

Given the importance of *E. faecalis* stress resistance abilities, the growth of the V19 strain and that of the 7 mutants were tested in different stress conditions, *i.e.* acid pH (pH 5.7), bile salts (0.06%), oxidative stress (1 mM H<sub>2</sub>O<sub>2</sub>), and low temperature (15°C) (Table). Significant difference in the doubling time was observed for  $\Delta rnjB$ ,  $\Delta rnc$  and  $\Delta rny$  compared to the V19 strain for low temperature, oxidative stress and bile salts exposure, respectively. The deletion of the *cshB* and *cshC* genes did also alter growth of *E. faecalis* at 15°C, or when facing the oxidative stress for *cshC* only and the *cshA* mutation had an effect in all stress conditions studied. Consequently, CshA, CshB and CshC participate in resistance to stressful conditions, with CshA having the widest role in our conditions.

**Table 2:** Doubling time (in min) of V19,  $\Delta rnjB$ ,  $\Delta rnc$ ,  $\Delta rny$ ,  $\Delta cshA$ ,  $\Delta cshB$  and  $\Delta cshC$  strains in BHI and in presence of different stress. Standard errors of the mean (three independent replicates) doubling time are indicated, and the asterisks correspond to significant differences compared to the WT strain in same condition culture ( $p < 0.01$ ).

Growth conditions	V19	$\Delta rnjB$	$\Delta rnc$	$\Delta rny$	$\Delta cshA$	$\Delta cshB$	$\Delta cshC$
BHI, 37°C	42.3 (±3.2)	43.4 (±3.3)	42.2 (±3.5)	47.9 (±4.6)	50.3 (±5.8)*	41.2 (±2.9)	40.1 (±3.9)
pH 5.7, 37°C	59.8 (±4.5)	55.9 (±4.9)	60.5 (±5.1)	75.2 (±6.1)	79.1 (±6.2)*	59.1 (±3.9)	59.2 (±3.5)
0.06% Bile Salts, 37°C	56.7 (±3.5)	54.2 (±3.7)	63.4 (±5.5)	71.7 (±5.6)*	75.1 (±4.7)*	58.3 (±5.9)	59.8 (±4.3)
1 mM H <sub>2</sub> O <sub>2</sub> , 37°C	58.7 (±4.2)	60.1 (±5.6)	66.6 (±5.0)*	64.4 (±5.9)	80.0 (±6.1)*	60.5 (±5.3)	68.7 (±4.2)*
BHI, 15°C	270.5 (±8.2)	392.3 (±9.2)*	295.2 (±7.2)	295.2 (±7.3)	527.8 (±10.1)*	458 (±11.3)*	504.3 (±9.0)*

To determine if these stress conditions induce the expression of the studied genes, RT-qPCR assays were performed. The cold exposure led to an overexpression of all 7 genes (Table 2). Moreover, *rny* and *cshA* are induced in all the tested conditions, suggesting a central role of the RNase Y and CshA in cell adaption during stress exposure, while RNase J2, RNase III, CshB and CshC seem to have a more stress-specific role.

**Table 3:** Significant fold changes in the expression of *rnjA*, *rnjB*, *rnc*, *rny*, *cshA*, *cshB*, *cshC*, in V19 strain exposed to 30 min of stress or during mice infection, and compared to exponential growth cells at 37°C in BHI. The level of expression was measured by RT-qPCR and compared to the expression of the housekeeping gene *gyrA*. NS: non-significant, in bold: induced expression.

Stress conditions	<i>rnjA</i>	<i>rnjB</i>	<i>rnc</i>	<i>rny</i>	<i>cshA</i>	<i>cshB</i>	<i>cshC</i>
pH5.7, 37°C	-14.2	-21.0	NS	<b>9.2</b>	<b>2.5</b>	NS	NS
0.06% Bile Salts, 37°C	<b>3.5</b>	<b>5.2</b>	<b>6.3</b>	<b>5.3</b>	<b>4.2</b>	NS	NS
1 mM H <sub>2</sub> O <sub>2</sub> , 37°C	-29.8	-28.3	<b>3.4</b>	<b>10.0</b>	<b>7.3</b>	-2.2	NS
BHI, 15°C	<b>5.7</b>	<b>2.6</b>	<b>4.2</b>	<b>5.9</b>	<b>12.2</b>	<b>9.3</b>	<b>5.3</b>
mice peritonitis	<b>16.2</b>	<b>12.5</b>	<b>2.5</b>	<b>9.1</b>	<b>6.3</b>	NS	NS

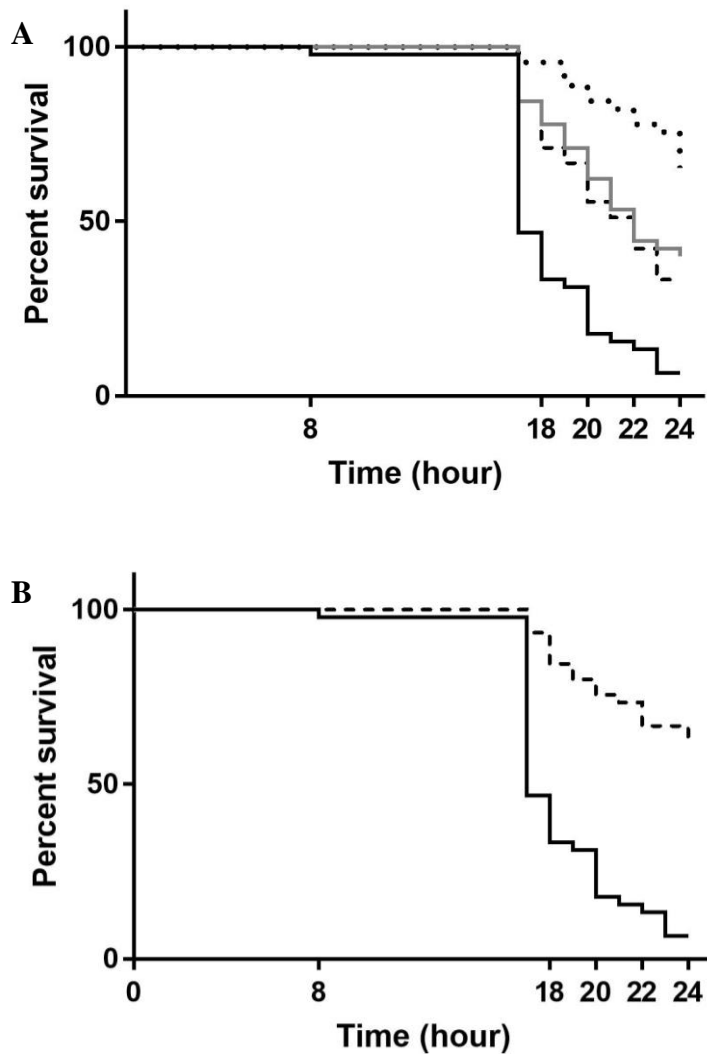
### ***Roles in virulence***

The expression of the RNases and helicases was also determined during mice infection. RNA extracted from bacterial cells recovered after mice peritonitis (Muller *et al.*, 2015) were used for RT-qPCR experiments. The *rnjA*, *rnjB*, *rny*, *rnc* and *cshA* genes showed an induced expression (Table 3) that suggests a role of the encoded proteins in the opportunism of *E. faecalis*.

In order to determine if the studied proteins are required for virulence, we tested the ability for the different mutants to kill the *G. mellonella* larvae (Figure 2). The monitoring of caterpillars survival showed that  $\Delta cshA$  and  $\Delta rny$  mutants were highly deficient in their infectious abilities ( $p$ -value <0.0001). The  $\Delta rnjB$  and  $\Delta rnc$  exhibited an intermediary phenotype between the



precedent group and the V19 strain ( $p$ -value  $<0.0001$ ). It can be suggested that part of the virulence ability of the V19 strain is linked to its capacity to survive stress encountered in the host. Both  $\Delta cshC$  and  $\Delta cshB$  mutants showed phenotypes of virulence similar to that of the V19 strains (data not shown).



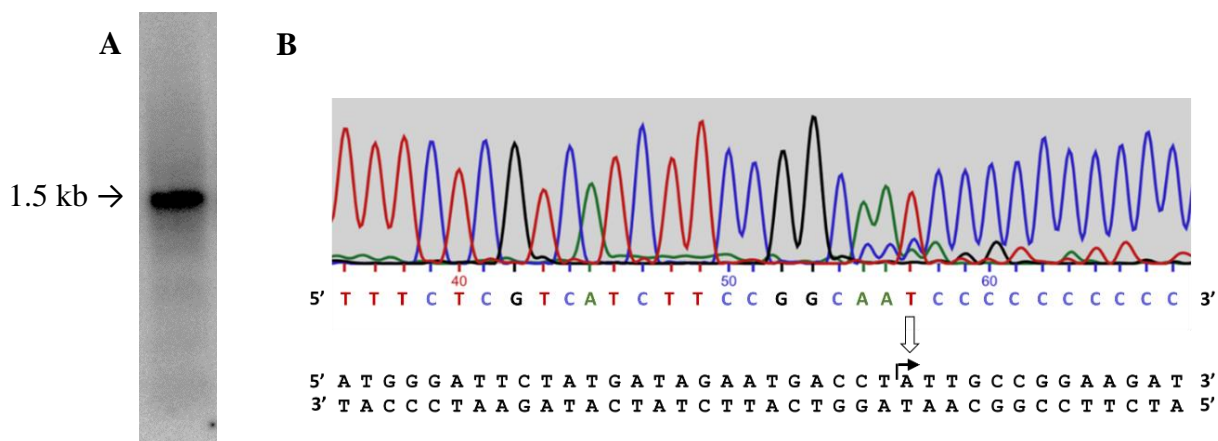
**Figure 2:** Survival of *G. mellonella* inoculated with *E. faecalis* V19,  $\Delta rnjB$ ,  $\Delta rnc$ ,  $\Delta rny$  and  $\Delta cshA$  strains. Worm larvae were inoculated with  $2 \cdot 10^8$  CFU of bacteria grown in BHI at 37°C. Survival was monitored between 8 and 26 h post-infection at 37°C using 15 larvae per strain. A: V19 (black line),  $\Delta rnjB$  (grey line),  $\Delta rnc$  (dashed line),  $\Delta rny$  (dotted line), B: V19 (black line),  $\Delta cshA$  (dashed line). The assays were performed in triplicates, and log-rank tests were used to evaluate results significance.

### *Genetic environment of cshA, discovery of the 5'UTR 5'cshA*

The physiological studies described above demonstrated that CshA played an important role in both the stress response and in the virulence of *E. faecalis* which prompted its characterization. In a previous study (Muller *et al.*, 2015), we highlighted the existence of a putative sRNA (named SRC43) located between EF0845 (*murF*) and EF0846 (*cshA*) genes.

RNA hybridization with a probe specific for this sRNA (Figure 3A) showed a transcript whose size (approximately 1.5 kb) is identical to that of the mRNA hybridized with a *cshA* specific probe (data not shown). In addition, *in silico* analyzes using Findterm software did not reveal any transcription terminator upstream of *cshA*. A 5'RACE-PCR assay confirmed that the 5' end of the putative sRNA (Figure 3B) is identical to that of *cshA* (data not shown). Considering these results, we concluded that this putative sRNA is actually a 5'UTR (Untranslated Region) of *cshA*, renamed 5'*cshA* (Figure 4).

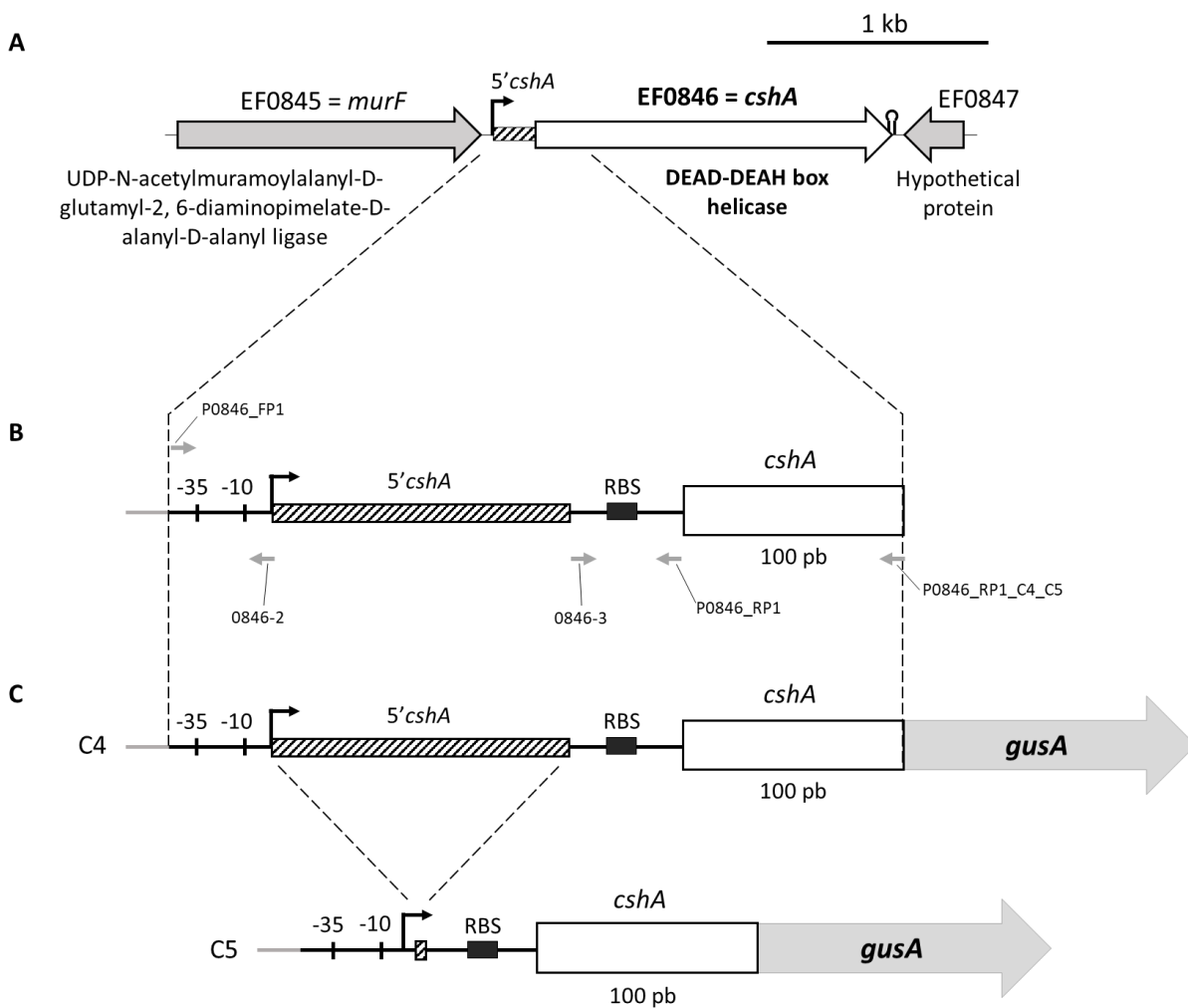
The +1 of transcription for *cshB* and the *ef1008-cshC* operon were also identified by 5'RACE-PCR, and positioned 36 bp and 25 bp before ATG for each transcription unit, respectively (data not shown), indicating the absence of 5'UTR for these helicase genes.



**Figure 3:** Characterization of the *cshA* mRNA. (A) Northern blot targeting the gene encoding the putative SRC43 sRNA upstream of *cshA*, performed with total RNA of the *E. faecalis* V19 strain. (B) 5'RACE-PCR sequencing spectrum of SRC43 obtained after poly-C tailing of cDNA. The corresponding transcriptional start site (+1) is indicated by an arrow.

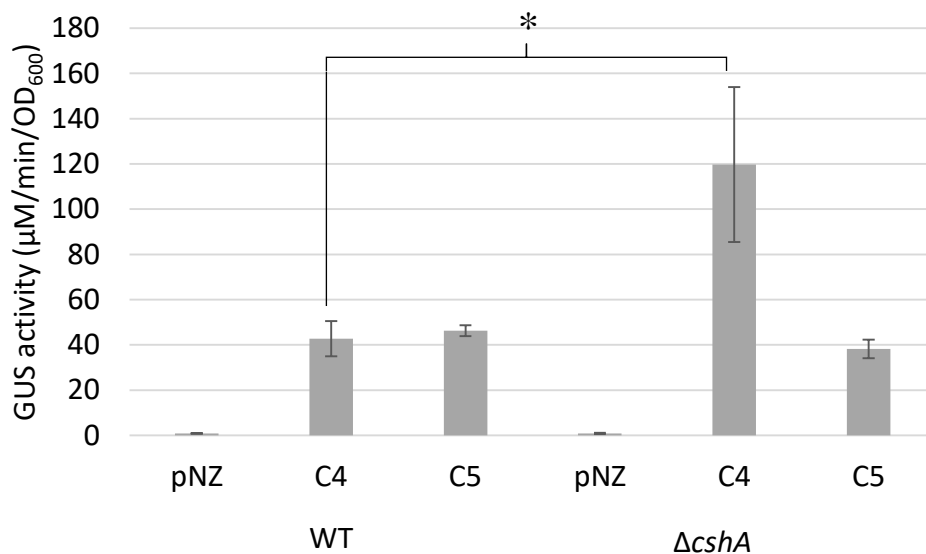
### *CshA* regulates its own expression

To understand the role of this 5'UTR in *cshA* regulation, we used translational fusion constructions. Three translational fusions were constructed in the pNZ273 vector, named C3 to C5 (Figure 4), with C3 and C4 containing the 5'UTR, contrary to C5. C4 and C5 include the 100 first base pairs of the *cshA* open reading frame in order to make sure to include all potential regulator elements.



**Figure 4:** (A) Genetic environment of the gene encoding the CshA DEAD-box helicase (EF0846), (B) location of oligonucleotides used for fusion constructions into the pNZ273 plasmid, and (C) C4 and C5 constructions inserted into the pNZ273 plasmid. The C4 construction contains the promoter and the entire 5'UTR of *cshA* (*5'cshA* (striped), the Ribosome Binding Site (RBS) and the first one hundred base pairs of the *cshA* ORF (white) cloned into the pNZ273 plasmid (grey). The C5 construction is the same as C4, but with a deletion of *5'cshA*.

These constructions were used to transform *E. faecalis* V19 and  $\Delta cshA$  strains and an empty pNZ273 plasmid was also introduced in these two tested strains to be used as negative control. C3 produced a very weak signal ( $4.5 \mu\text{M}/\text{min}/\text{OD}_{600}$ ) (data not shown) compared to C4 that produced a GUS activity of  $42.7 \mu\text{M}/\text{min}/\text{OD}_{600}$  (Figure 5). This suggests that the first hundred base pairs of *cshA* play a role in the expression of the gene.



**Figure 5:**  $\beta$ -glucuronidase (GUS) assay with the C4 and C5 constructions in the V19 and  $\Delta cshA$  strains. pNZ: empty plasmid, negative control. Data are the mean of three independent assays. \*: statistically significant in comparison to the same construction in the V19 strain, p-value < 0.05 (Kruskal-Wallis test).

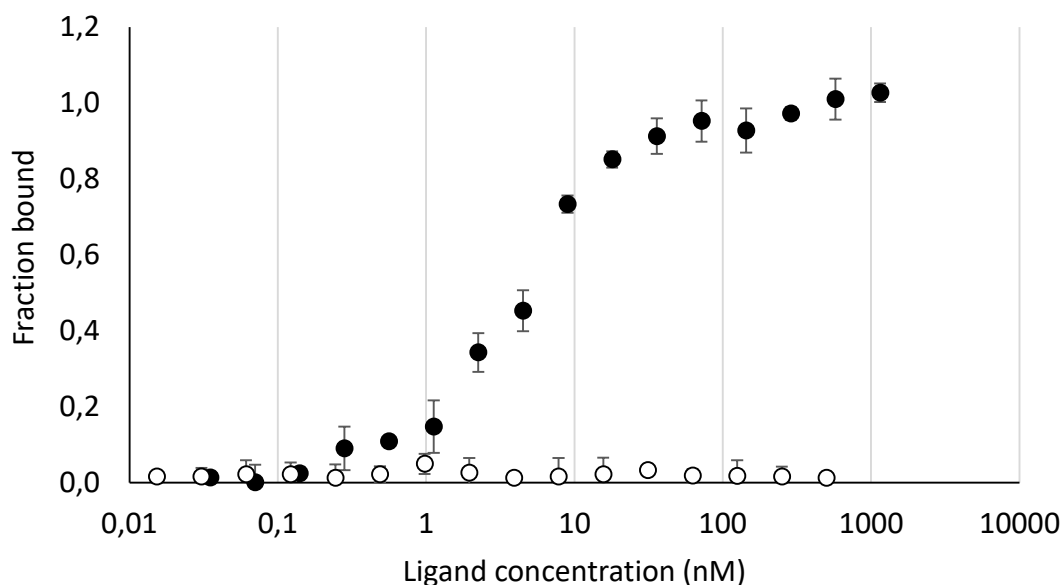
We also detected a very strong activity of  $119.8 \mu\text{M}/\text{min}/\text{OD}_{600}$  with C4 into  $\Delta cshA$  strain (Figure 5). This result indicates that in the absence of CshA, 5' *cshA* allows better expression of downstream elements, suggesting a transcriptional or post-transcriptional auto-regulation of *cshA* via its 5'UTR. This result was confirmed by a RT-qPCR assay targeting *gusA* (Table S2), which shows that the amount of the mRNA corresponding to the C4 construct was 1.7 times higher in the  $\Delta cshA$  background than in the V19 strain.

Considering these experiments, CshA would have a repressing activity on its own expression. We can therefore hypothesize that CshA acts *via* its helicase activity, by relaxing secondary structures in the 5'UTR RNA, allowing the formation of others structures that terminate the transcription or block the translation which could indirectly affect the stability or RNAs.

### ***CshA binds to its 5'UTR***

According to our hypothesis CshA would be able to bind to the 5'*cshA* RNA. To prove it, we used the MicroScale Thermophoresis (MST) technology. Recombinant CshA (CshA-6xHis) was purified (using pQE30) and labelled on the His-tag. RNA corresponding to a region covering the 5'UTR and the 100 first bases of the *cshA* ORF (equivalent of the insert of the C4 construction) was synthesized using *in vitro* transcription (see Material and Methods). The previously identified SRC65 sRNA (Shioya *et al.*, 2011) which does not share any sequence homology with 5'*cshA*, was also synthesized and used as negative control.

The MST experiment revealed an increase of the CshA-6xHis bound fraction following the concentration of the 5'*cshA* RNA, in contrast to the negative control (Figure 6).



**Figure 6:** Dose-response curve of MicroScale Thermophoresis (MST) showing the binding of CshA to its ligand 5'*cshA* RNA (black circles). Negative control (white circles) was performed with CshA and the SRC65 sRNA that does not share homology with 5'*cshA*.

This dose-response interaction between CshA and 5' *cshA* RNA made it possible to determine a  $K_D$  of 1.6 nM, showing a high and specific affinity of CshA for its 5'UTR RNA.

### ***Analysis of the CshA interactome***

To identify CshA complexing proteins, we constructed a fusion of the *cshA* gene with the TAP tag (Tandem Affinity Purification) (Stingl *et al.*, 2008), composed of the sequence of the protein A, a TEV protease site, and the calmodulin binding domain as described in Material and Methods. Thus CshA-TAP served as a bait protein to allow the identification of its interacting partners, after purification with two successive affinity columns, SDS-PAGE and analysis by mass spectrometry. The first assay showed that CshA interacting proteins are mainly the enolase and ribosomal proteins of the 50S and 30S subunit (data not shown).

Given the RNA binding property of CshA, the same experiment was achieved with a supplementary step of incubation with RNaseA after cell lysis (in order to remove RNA and therefore identify CshA's direct partners). The interactome analysis showed that only the enolase was still bound to CshA. These data provide evidence that CshA interaction with ribosomal subunits is dependent on RNA encountered in the cell. Thus, the helicase has a favoured partner that is the enolase.

### **Discussion**

In this study, we have explored expression and functions of key enzymes of the bacterial RNA metabolism: RNases J1, J2, III, Y and DEAD-box RNA helicases CshA, CshB and CshC. We found that these 7 enzymes are highly conserved in enterococci, but also in other Gram-positive bacteria. This suggests that these proteins play a prominent role in responding to environmental changes among bacteria in the *Enterococcus* genus.

As for *B. subtilis* (Mathy *et al.*, 2007), we have found that the RNase J1 is also essential in *E. faecalis*. RNase Y appeared to be important for the general fitness of the bacteria, as its

depletion leads to a growth rate defect. It is induced in a wide range of stress and during mice peritonitis, showing its key role into stress adaptation. Moreover, the virulence of *E. faecalis* lacking this RNase was strongly defected. This is in agreement with previous data obtained in *Streptococcus pyogenes*, where RNase Y mediates the expression of the exotoxin B, a major virulence factor (Broglia *et al.*, 2018). Both RNase J2 and III play a role in the response to cold, oxidative, and bile salts stress. They also have an effect on virulence, although this is less pronounced than with RNase Y. This agrees with previous studies; indeed, in *E. faecalis*, the RNase J2 was found to be involved in response to bile salts stress and important for infection (Gao *et al.*, 2017), and in *S. aureus*, RNase III plays a role in virulence factors repression and stress adaptation through modulation of the expression of the stress response protein CspA (Lioliou *et al.*, 2012).

Involved in every RNA process, DEAD-box helicases are key enzymes in prokaryotic metabolism (Cordin *et al.*, 2006). They are not always essential, but their depletion can affect bacterial fitness (El Mortaji *et al.*, 2018). Thereby, removal of *cshA* in the genome caused a reduction of the growth rate, and a deficiency in the bacterial virulence as observed in the *Galleria mellonella* assay. These two characteristics are likely connected, as a defect in growth leads to a less efficient colonization. This suggest that CshA plays a central role in growth, and therefore in virulence. These data are consistent with that of RT-qPCR assays revealing the overexpression of *cshA* in mice peritonitis. Finally, our data also showed that *cshA* is overexpressed in a wide range of stress (acid, cold, oxidative and bile salts presence), highlighting its role into the general stress response. We noted that it was more expressed during the cold stress than in the other conditions. In contrast, the other two helicases (CshB and CshC) were overexpressed only in cold stress. Therefore, these two proteins are most likely specific to this stress. These results agree with the study of CshA, CshB and CshC in *B. cereus* (Pandiani *et al.*, 2011), but also with other previous studies that demonstrated the role of DEAD-helicases

in cold stress resistance (Resch *et al.*, 2010; El Mortaji *et al.*, 2018). Also, CshC does not seem to have an effect on *E. faecalis* virulence.

The existence of a 5'UTR (5'*cshA*) upstream of the *cshA* ORF was highlighted with *in silico* and experimental studies. We also demonstrated that CshA is able to bind this 5'UTR. Moreover in the absence of CshA, gene downstream of 5'*cshA* has a higher expression as compared to the control conditions. This means that CshA represses its own expression, in order to maintain a proper amount of the protein. When CshA interacts with 5'*cshA*, it may form or unwind secondary structures, ending the transcription or inhibiting the translation. No 5'UTR were evidenced for *cshB* and *cshC* (data not shown), which means that cold stress regulation of these helicase is independent of a 5'UTR.

DEAD-box helicases were previously described as capable of promoting the initiation of translation through destabilization of mRNA structures in prokaryotes (Kaberdin and Bläsi, 2013; Intile *et al.*, 2015), but no role of reduction of the translation efficiency has been described to date. RNA helicase autoregulation has been reported in cyanobacteria, and this occurs indirectly through undefined transcriptional and post-transcriptional mechanisms (Rosana *et al.*, 2012). This was also highlighted recently in *Caulobacter crescentus*, where the DEAD-box helicase RhIE might regulate its own expression through its 5'UTR, but as a translation inducer (Aguirre *et al.*, 2017). It is to be noted that this 5'UTR does not share any sequence identity with 5'*cshA*. So our results may be the first sighting of a direct auto-repressing activity of a DEAD-box helicase.

We also studied the role of CshA as part of the degradosome through interactome assays. We showed that the enolase was the preferred protein with which CshA interacts in the absence of RNA. Such an interaction has already been demonstrated before in *B. subtilis* (Lehnik-Habrink *et al.*, 2010), but to our knowledge not in *E. faecalis*. In *B. subtilis*, all components of the degradosome interacts with CshA (Lehnik-Habrink *et al.*, 2010), but we did not witnessed such



interaction here for *E. faecalis*. This could be explained by the fact that in order to maintain expression of CshA under the control of its 5'UTR, the TAP tag was added at the C-terminus of the protein. As it was previously suggested that this end is involved in the degradosome interaction in *S. aureus* (Giraud *et al.*, 2015), this may be why we detected only few proteins in the interactome study. This could also mean that enolase is interacting with CshA through another part of the helicase.

In summary, our data demonstrate that proteins of RNA metabolism play important roles in environmental changes response and virulence of *E. faecalis*. These findings provide new perspectives on the molecular mechanism that lead to the opportunistic behavior of this bacteria.

### **Acknowledgment**

We are grateful to Hilde De Reuse for providing the pILL825C plasmid.

*Note* : Afin de privilégier la clarté du manuscrit, les données supplémentaires précédemment citées sont disponibles en Annexe 3.

## References

- Aguirre, A.A., Vicente, A.M., Hardwick, S.W., Alvelos, D.M., Mazzon, R.R., Luisi, B.F., and Marques, M.V. (2017) Association of the cold-shock DEAD-box RNA helicase RhIE to the RNA degradosome in *Caulobacter crescentus*. *J Bacteriol* **JB.00135-17**.
- Aké, F.M.D., Joyet, P., Deutscher, J., and Milohanic, E. (2011) Mutational analysis of glucose transport regulation and glucose-mediated virulence gene repression in *Listeria monocytogenes*. *Molecular Microbiology* **81**: 274–293.
- Arias, C.A., and Murray, B.E. (2012) The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat Rev Microbiol* **10**: 266–278.
- Arnaud, M., Chastanet, A., and Débarbouillé, M. (2004) New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. *Appl Environ Microbiol* **70**: 6887–6891.
- Babady, N.E. (2016) Hospital-associated infections. *Microbiology Spectrum* **4** <http://www.asmscience.org/content/journal/microbiolspec/10.1128/microbiolspec.DMIH2-0003-2015>. Accessed June 27, 2017.
- Bae, T., Kozłowicz, B., and Dunny, G.M. (2002) Two targets in pCF10 DNA for PrgX binding: their role in production of Qa and prgX mRNA and in regulation of pheromone-inducible conjugation11. *Journal of Molecular Biology* **315**: 995–1007.
- Benachour, A., Ladjouzi, R., Le Jeune, A., Hébert, L., Thorpe, S., Courtin, P., *et al.* (2012) The lysozyme-induced peptidoglycan N-acetylglucosamine deacetylase PgdA (EF1843) is required for *Enterococcus faecalis* virulence. *J Bacteriol* **194**: 6066–6073.
- Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248–254.
- Broglià, L., Materne, S., Lécivain, A.-L., Hahnke, K., Rhun, A.L., and Charpentier, E. (2018) RNase Y-mediated regulation of the streptococcal pyrogenic exotoxin B. *RNA Biology* **15**: 1336–1347.
- Cordin, O., Banroques, J., Tanner, N.K., and Linder, P. (2006) The DEAD-box protein family of RNA helicases. *Gene* **367**: 17–37.
- El Mortaji, L., Aubert, S., Galtier, E., Schmitt, C., Anger, K., Redko, Y., *et al.* (2018) The sole DEAD-box RNA helicase of the gastric pathogen *Helicobacter pylori* is essential for colonization. *MBio* **9**.
- Frère, J., Novel, M., and Novel, G. (1993) Molecular analysis of the *Lactococcus lactis* subspecies *lactis* CNRZ270 bidirectional theta replicating lactose plasmid pUCL22. *Mol Microbiol* **10**: 1113–1124.

- Gao, P., Pinkston, K.L., Bourgoigne, A., Murray, B.E., Hoof, A. van, and Harvey, B.R. (2017) Functional studies of *E. faecalis* RNase J2 and its role in virulence and fitness. *PLoS One* **12** <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5383250/>. Accessed October 30, 2018.
- Gao, P., Pinkston, K.L., Nallapareddy, S.R., Hoof, A. van, Murray, B.E., and Harvey, B.R. (2010) *Enterococcus faecalis* rnjB is required for pilin gene expression and biofilm formation. *J Bacteriol* **192**: 5489–5498.
- Giraud, C., Hausmann, S., Lemeille, S., Prados, J., Redder, P., and Linder, P. (2015) The C-terminal region of the RNA helicase CshA is required for the interaction with the degradosome and turnover of bulk RNA in the opportunistic pathogen *Staphylococcus aureus*. *RNA Biol* **12**: 658–674.
- Intile, P.J., Balzer, G.J., Wolfgang, M.C., and Yahr, T.L. (2015) The RNA helicase DeaD stimulates ExsA translation to promote expression of the *Pseudomonas aeruginosa* type III secretion system. *J Bacteriol* **197**: 2664–2674.
- Jamalli, A., Hébert, A., Zig, L., and Putzer, H. (2014) Control of expression of the RNases J1 and J2 in *Bacillus subtilis*. *J Bacteriol* **196**: 318–324.
- Johnson, C.M., Haemig, H.H.A., Chatterjee, A., Wei-Shou, H., Weaver, K.E., and Dunny, G.M. (2011) RNA-mediated reciprocal regulation between two bacterial operons is RNase III dependent. *MBio* **2**.
- Kaberdin, V.R., and Bläsi, U. (2013) Bacterial helicases in post-transcriptional control. *Biochim Biophys Acta* **1829**: 878–883.
- Khemici, V., and Carpousis, A.J. (2003) The RNA degradosome and poly(A) polymerase of *Escherichia coli* are required in vivo for the degradation of small mRNA decay intermediates containing REP-stabilizers. *Molecular Microbiology* **51**: 777–790.
- Khemici, V., and Linder, P. (2016) RNA helicases in bacteria. *Curr Opin Microbiol* **30**: 58–66.
- Kram, K.E., and Finkel, S.E. (2015) Rich medium composition affects *Escherichia coli* survival, glycation, and mutation frequency during long-term batch culture. *Appl Environ Microbiol* **81**: 4442–4450.
- Lehnik-Habrink, M., Lewis, R.J., Mäder, U., and Stülke, J. (2012) RNA degradation in *Bacillus subtilis*: an interplay of essential endo- and exoribonucleases. *Mol Microbiol* **84**: 1005–1017.
- Lehnik-Habrink, M., Newman, J., Rothe, F.M., Solovyova, A.S., Rodrigues, C., Herzberg, C., et al. (2011) RNase Y in *Bacillus subtilis*: a Natively disordered protein that is the functional equivalent of RNase E from *Escherichia coli*. *J Bacteriol* **193**: 5431–5441.
- Lehnik-Habrink, M., Pförtner, H., Leonie, R., Pietack Nico, Herzberg Christina, and Stülke Jörg (2010) The RNA degradosome in *Bacillus subtilis*: identification of CshA as the major RNA helicase in the multiprotein complex. *Molecular Microbiology* **77**: 958–971.
- Li, R.W. (ed.) (2011) *Metagenomics and its applications in agriculture, biomedicine, and environmental studies*. Nova Science Publisher's, Hauppauge, N.Y.

- Linares, D.M., Perez, M., Ladero, V., Rio, B. del, Redruello, B., Martin, M.C., *et al.* (2014) An agmatine-inducible system for the expression of recombinant proteins in *Enterococcus faecalis*. *Microbial cell factories* **13**: 1–9.
- Lioliou, E., Sharma, C.M., Caldelari, I., Helfer, A.-C., Fechter, P., Vandenesch, F., *et al.* (2012) Global regulatory functions of the *Staphylococcus aureus* endoribonuclease III in gene expression. *PLoS Genet* **8**: e1002782.
- Madsen, K.T., Skov, M.N., Gill, S., and Kemp, M. (2017) Virulence factors associated with *Enterococcus faecalis* infective endocarditis: a mini review. *Open Microbiol J* **11**: 1–11.
- Mathy, N., Bénard, L., Pellegrini, O., Daou, R., Wen, T., and Condon, C. (2007) 5'-to-3' exoribonuclease activity in bacteria: Role of RNase J1 in rRNA maturation and 5' stability of mRNA. *Cell* **129**: 681–692.
- Muller, C., Cacaci, M., Sauvageot, N., Sanguinetti, M., Rattei, T., Eder, T., *et al.* (2015) The intraperitoneal transcriptome of the opportunistic pathogen *Enterococcus faecalis* in mice. *PLOS ONE* **10**: e0126143.
- Pandiani, F., Chamot, S., Brillard, J., Carlin, F., Nguyen-the, C., and Broussolle, V. (2011) Role of the five RNA helicases in the adaptive response of *Bacillus cereus* ATCC 14579 cells to temperature, pH, and oxidative stresses. *Appl Environ Microbiol* **77**: 5604–5609.
- Paulsen, I.T., Banerjee, L., Myers, G.S.A., Nelson, K.E., Seshadri, R., Read, T.D., *et al.* (2003) Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* **299**: 2071–2074.
- Platteeuw, C., Simons, G., and Vos, W.M. de (1994) Use of the *Escherichia coli* beta-glucuronidase (*gusA*) gene as a reporter gene for analyzing promoters in lactic acid bacteria. *Appl Environ Microbiol* **60**: 587–593.
- Redko, Y., Galtier, E., Arnion, H., Darfeuille, F., Sismeiro, O., Coppée, J.-Y., *et al.* (2016) RNase J depletion leads to massive changes in mRNA abundance in *Helicobacter pylori*. *RNA Biol* **13**: 243–253.
- Resch, A., Večerek, B., Palavra, K., and Bläsi, U. (2010) Requirement of the CsdA DEAD-box helicase for low temperature riboregulation of *rpoS* mRNA. *RNA Biology* **7**: 796–802.
- Rosana, A.R.R., Chamot, D., and Owtrim, G.W. (2012) Autoregulation of RNA helicase expression in response to temperature stress in *Synechocystis* sp. PCC 6803. *PLOS ONE* **7**: e48683.
- Roux, C.M., DeMuth, J.P., and Dunman, P.M. (2011) Characterization of components of the *Staphylococcus aureus* mRNA degradosome holoenzyme-like complex. *J Bacteriol* **193**: 5520–5526.
- 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**: e23948.

Stingl, K., Schauer, K., Ecobichon, C., Labigne, A., Lenormand, P., Rousselle, J.-C., *et al.* (2008) *In vivo* interactome of *Helicobacter pylori* urease revealed by tandem affinity purification. *Molecular & Cellular Proteomics* **7**: 2429–2441.

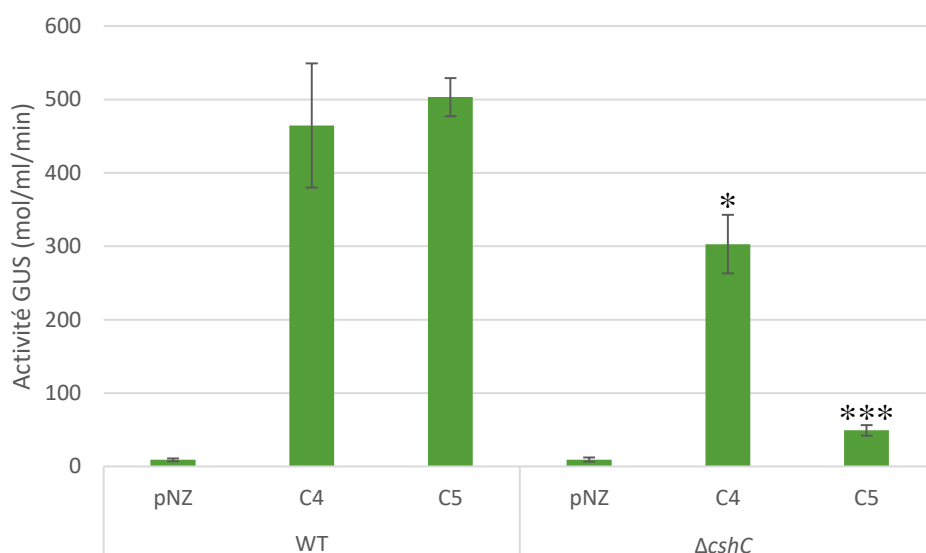
Xayaphoummine, A., Bucher, T., and Isambert, H. (2005) Kinofold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots. *Nucleic Acids Research* **33**: W605–W610.

Zhao, C., Hartke, A., La Sorda, M., Posteraro, B., Laplace, J.-M., Auffray, Y., and Sanguinetti, M. (2010) Role of methionine sulfoxide reductases A and B of *Enterococcus faecalis* in oxidative stress and virulence. *Infect Immun* **78**: 3889–3897.

## 2. Résultats complémentaires

### ***Influence de CshC dans la régulation de CshA***

Lors de nos expériences de fusions traductionnelles avec la région promotrice, la 5'UTR et la boîte RBS de l'hélicase CshA, nous avons également transformé la souche  $\Delta cshC$  avec les constructions C4 et C5. Cela nous a permis de voir si cette hélicase pouvait avoir une influence sur l'expression de CshA, au moyen du dosage de la  $\beta$ -glucuronidase (Figure 24). Comme lors de la première expérience, le plasmide vide a également été introduit dans la souche  $\Delta cshC$  pour servir de témoin négatif.



*Figure 24 : Résultats des dosages de la  $\beta$ -glucuronidase avec les constructions C4 et C5, transformées dans les souches V19 et  $\Delta cshC$ . pNZ : plasmide vide utilisé comme contrôle négatif. Dosage de l'activité GUS réalisée sur 10 minutes. \* : statistiquement significatif en comparaison avec la même construction chez la souche V19 ( $p < 0,05$ ); \*\*\* :  $p < 0,001$ .*

La construction C4 présente une activité GUS légèrement réduite ( $303 \pm 40$  mol/ml/min,) dans la souche  $\Delta cshC$  par rapport à la même construction chez la souche sauvage ( $465 \pm 85$  mol/ml/min,).

La construction C5 a quant à elle montré une activité significativement réduite ( $49 \pm 7$  mol/ml/min,) par rapport à la souche sauvage ( $503 \pm 26$  mol/ml/min.). Cela laisse supposer que CshC est impliquée dans la régulation de l'expression de *csmA*. De plus, nous remarquons que cette répression est présente lorsque CshC et la 5'UTR sont absentes. Nous pouvons donc en conclure que ces deux éléments sont importants pour l'expression de *csmA*, et que leurs actions régulatrices sont indépendantes.

Ces résultats ont été confirmés par RT-qPCR ciblant le gène *gusA* (oligonucléotides *gusA-G* et *gusA-D*, Tableau 3), avec un profil similaire aux résultats de dosage de la  $\beta$ -glucuronidase (C4 :  $\Delta\text{cshC}/V19 = 0,46$  ; C5 :  $\Delta\text{cshC}/V19 = 0,15$ ).

Afin de déterminer si CshB ou CshC étaient régulées par un mécanisme similaire à celui de CshA, la présence d'une 5'UTR a été recherchée par des expériences de 5'RACE-PCR (oligos *ef1377-SP1* et *ef1377-SP2* pour CshB, et *ef1009-SP1*, *ef1009-SP2* et *ef1009-SP3* pour CshC, voir Tableau 3). Il est apparu que *cskB* ne possède pas de 5'UTR. De son côté, *cskC* est co-transcrit avec le gène qui le précède, *ef1008*. Des expériences de 5'RACE sur le gène *ef1008* sont en cours.

### **Régulation potentielle du gène *ef0845 (murF)* par 5'*csmA***

Lors de l'étude de l'environnement génomique de 5'*csmA*, nous avons réalisé une expérience de 3'RACE-PCR sur le gène en amont, *ef0846*, afin de voir si ce gène n'était pas co-transcrit avec la 5'UTR, 5'*csmA*. En effet, nous avons noté l'absence *in silico* de terminateur entre ces deux régions.

Nous avons été surpris de voir que l'extrémité 3' du gène *ef0845* se trouvait au niveau de la fin prédite par le RNA-Seq du petit ARN putatif que nous avons identifié comme étant 5'*csmA*. Il

pourrait donc être possible que 5' *cshA*, en plus d'être la 5'UTR du gène *ef0846* (*cshA*), fasse également partie de la 3'UTR du gène *ef0845* (*murF*).

Le gène *ef0845* code une UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate--D-alanyl-D-alanyl ligase, enzyme impliquée dans la synthèse de la paroi bactérienne. En effet, elle catalyse la dernière étape de la synthèse du précurseur du peptidoglycane, l'UDP-N-acetylmuramoyl-pentapeptide. Une expérience de microscopie électronique, a montré chez une souche mutante de 5' *cshA* la présence de nodules à la surface des bactéries, absents chez la souche sauvage (Figure 25).

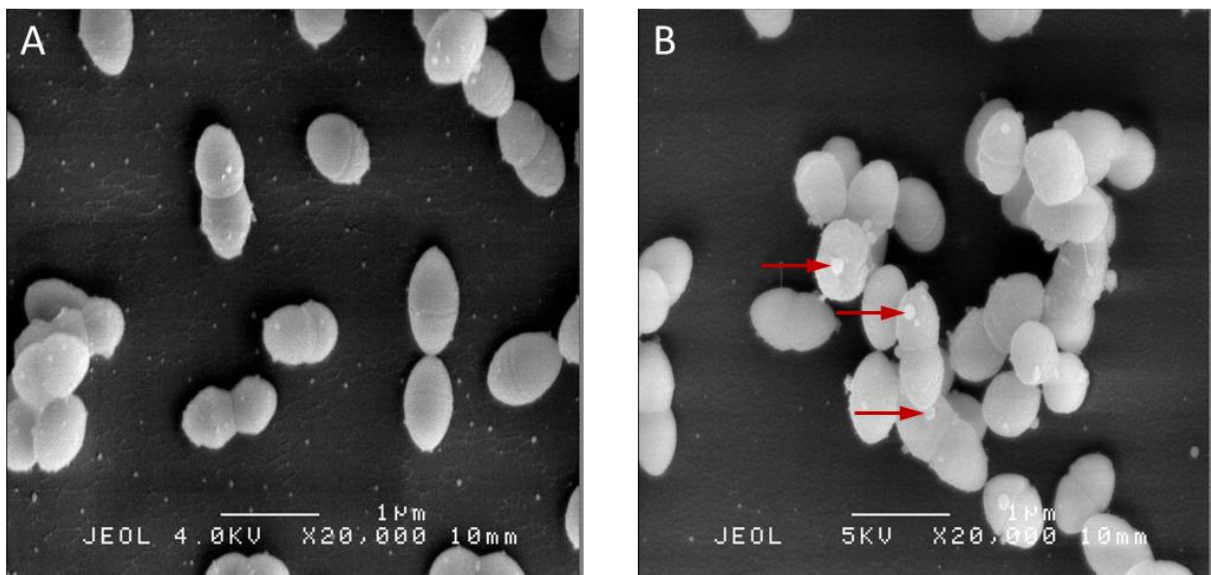


Figure 25 : Observation en microscopie électronique à balayage de la souche V19 (A) et de la souche  $\Delta 5'$ *cshA* (B).

La présence de ces nodules indiquerait potentiellement une modification de la paroi, ce qui irait dans le sens d'une régulation du gène *ef0845* par une 3'UTR contenant la séquence de 5' *cshA*.



*Discussion générale et  
perspectives*

Au cours de cette étude, nous avons détecté 65 petits ARN putatifs induits *in vivo* et/ou en conditions de stress, dont 25 qui n'avaient pas été détectés jusqu'à présent. Nous avons choisi d'en étudier trois parmi tous ceux qui étaient fortement exprimés *in vivo*, en présence de sels biliaires et en pH acide.

Parmi ces trois sARN putatifs, seul SRC65 était véritablement un sARN. Nous avons pu montrer son implication dans la réponse au stress. En effet, SRC65 est induit en condition de stress (sels biliaires, urine, sérum, pH acide et stress oxydatif), et il joue un rôle dans le déclenchement de la phase exponentielle de croissance : sa surexpression provoque un allongement de la phase de latence lors de la croissance bactérienne. Des études ont montré qu'un accroissement de la phase de latence peut être lié à une augmentation de la tolérance aux antibiotiques tels que l'ampicilline (Levin-Reisman *et al.*, 2017). Il serait donc intéressant d'étudier la tolérance aux antibiotiques des souches mutantes et de surexpression du sARN. Nous avons également souligné une homologie importante entre SRC65 et un autre sARN, SRC90. Ces deux sARN pourraient donc posséder des fonctions redondantes et partager la ou les mêmes cibles. Nous avons identifié plusieurs de leurs cibles potentielles *in silico*. Parmi celles-ci, *ef3251* code une protéine induite en sels biliaires (données déposées sur la base de données ArrayExpress par Solheim *et al.*, 2007), vraisemblablement sécrétée mais dont la fonction demeure inconnue. Le gène *ef0022* a été identifié comme cible potentielle de SRC65 uniquement. Il code pour la sous-unité EIID d'un PTS spécifique du mannose. Nous avons montré qu'il est moins exprimé chez le mutant  $\Delta$ SRC65 comparativement à la souche sauvage. Cependant, de précédentes études ont montré qu'*ef0022* est réprimé en présence de sang (données déposées sur la base de données ArrayExpress par Vebø *et al.*, 2009) et qu'*ef0021*, situé sur le même opéron, est réprimé en urine (données déposées sur la base de données ArrayExpress par Shankar *et al.*, 2002; Vebø *et al.*, 2010). Par l'approche iTRAQ, nous avons également détecté des protéines liées à la traduction (protéines ribosomales, protéines de

synthèse de l'ARNt...) dont l'abondance était différente chez le mutant délété du gène codant SRC65 et chez la souche sauvage V19. En accord avec le phénotype de croissance de la souche surexprimant SRC65, il n'est pas impossible que les gènes codant ces protéines ayant des fonctions primordiales dans la cellule soient des cibles du sARN.

Le régulon de SRC65 n'est pas encore entièrement caractérisé, mais nous avons proposé des candidats crédibles. Sa relation avec SRC90 reste également à approfondir. Les manipulations pour construire un mutant SRC90 ont été infructueuses jusqu'à présent. Il serait pourtant profitable d'étudier le comportement d'un tel mutant et surtout celui d'un double mutant SRC65/SRC90. Cela permettrait probablement de mieux cibler le(s) mécanisme(s) dans le(s)quel ces deux sARN interviennent.

Nous avons ensuite concentré nos efforts sur les deux autres ARN régulateurs qui nous fournissaient des résultats plus avancés et qui se sont révélés être des 5'UTR.

Nous avons mis en évidence que la 5'UTR 5'1515 contrôlait l'expression de l'opéron *ef1515-ef1516*, probablement *via* la formation de structures secondaires qui permettent la génération d'un terminateur de la transcription. La protéine EF1515, un antitermineur de la famille BglG/SacY, est capable de se lier à 5'1515 : nous supposons donc qu'elle y exerce une activité antiterminatrice suivant un mécanisme commun à d'autres protéines de la famille BglG/SacY découvertes chez d'autres espèces (Idelson et Amster-Choder, 1998; Amster-Choder, 2005). Ainsi, la protéine EF1516 contrôlerait l'activité d'EF1515 par phosphorylation/déphosphorylation et aurait donc un rôle de senseur et de régulateur, en plus de sa fonction première. Celle-ci est annotée comme sous unité EIIABEC d'un PTS spécifique du NAG, et l'opéron est effectivement induit en présence de ce sucre. Cependant, aucune différence de croissance n'a été observée pour les mutants  $\Delta 5'1515$ ,  $\Delta ef1515$  ou  $\Delta ef1516$  par rapport à la souche V19, ce qui laisse supposer que si le NAG est bien transporté par ce PTS,

ce transporteur ne constitue pas la seule voie d'entrée du NAG dans la cellule. Un autre transporteur hypothétique reste encore à identifier. Des expériences de transcriptomique ont également montré qu'*ef1516* était réprimé en présence d'urine et de sang (données déposées sur la base de données ArrayExpress par Shankar *et al.*, 2002; Vebø *et al.*, 2010 et Vebø *et al.*, 2009). De plus, cet opéron n'est probablement pas contrôlé par la répression catabolique mais il est possible qu'il existe un lien via la protéine HPr : cette protéine intervient dans la répression catabolique, mais également chez *E. coli* dans la régulation de l'antitermineur BglG. Cette régulation s'effectue par la participation d'HPr à la cascade de phosphorylation des PTS, mais aussi par activation directe de BglG. BglG doit d'une part être déphosphorylée par le PTS BglF, et d'autre part phosphorylée par HPr sur un site différent afin de pouvoir jouer son rôle d'antitermineur (Amster-Choder, 2005).

Nous avons relevé que la phase exponentielle de croissance se déclenchait plus tardivement pour les mutants  $\Delta ef1515$  et  $\Delta ef1516$  lorsqu'ils sont cultivés en présence de D-gluconate par rapport à la souche V19 sauvage. Cependant, l'opéron ne semble pas être induit en présence de ce sucre. En revanche, le prolongement de la phase de latence est encore plus important chez la souche 5'1515+. De plus, la souche  $\Delta 5'1515$  est affectée dans sa virulence, mais pas  $\Delta ef1515$  et  $\Delta ef1516$ . Cela pourrait signifier que la terminaison prématurée de la transcription due à la présence d'un terminateur situé sur 5'1515 conduit à la formation d'un ARN de petite taille avec une activité indépendante impliquant le métabolisme du D-gluconate. L'éventuelle fonction de cet ARN issu de la terminaison de la transcription mériterait donc d'être étudiée.

Nous avons détecté une autre cible régulée par l'antitermineur EF1515, le gène *ef3023*. EF1515 se fixe sur la 5'UTR (5'3023) de ce gène et y exercerait une régulation selon le même mécanisme que pour 5'1515.

La protéine EF3023 nous a intrigués par le fait qu'elle possède à la fois des domaines enzymatiques (polysaccharide lyase 8) et des domaines caractéristiques des protéines d'adhésion (ancrage dans la paroi, domaine Ig-like). Une étude chez *E. faecalis* a montré que les domaines Ig-like de cette protéine présentent un potentiel d'adhésion plus faible que ceux des autres protéines d'adhésion de même nature (Sillanpaa *et al.*, 2004). De plus, bien qu'EF3023 soit annotée hyaluronate lyase par la base de données KEGG, *E. faecalis* semble être incapable de pousser en présence d'acide hyaluronique comme seule source de carbone. Le substrat exact de cette protéine reste donc encore à découvrir, sachant que la chondroïtine et le xanthane sont des candidats crédibles. En effet, la famille des polysaccharide lyases 8 contient aussi les enzymes dégradant ces polysaccharides. Nous avons cependant démontré qu'EF3023 était bien impliquée dans la virulence d'*E. faecalis*, ce qui est conforme à une précédente étude (Tsigrelis *et al.*, 2006). Il pourrait être intéressant d'examiner les capacités d'adhésion de la souche  $\Delta ef3023$  sur cellules eucaryotes, en utilisant par exemple des lignées de cellules épithéliales de colon (Lebreton *et al.*, 2012) ou de rein (Cacaci *et al.*, 2018).

Nous avons également étudié le gène *ef1518* situé dans la même région génomique que l'opéron *ef1515-ef1516*. Ce gène code une protéine porteuse d'un domaine « lysozyme-like » et de telles protéines sont souvent associées à la virulence des bactéries (Mushegian *et al.*, 1996; Del Giudice *et al.*, 2013). Cependant, il est réprimé plus de 12 fois en modèle de péritonite chez la souris (Muller *et al.*, 2015). Etant donné que, à l'instar du lysozyme, les protéines « lysozyme-like » hydrolysent les liaisons  $\beta$ 1-4 du peptidoglycane, et que ce dernier est composé entre autres de N-acetyl-glucosamine, nous nous sommes demandés s'il ne pouvait pas y avoir un lien entre EF1518 et EF1516, annotée comme spécifique du NAG. Nous avons pu observer que, si EF1518 ne semble pas impliquée dans la croissance bactérienne *in vitro* (pas de phénotype de croissance), son absence provoque un défaut de virulence de la bactérie. Il a aussi été montré précédemment qu'*ef1518* était surexprimé en présence de sels biliaires (données déposées sur

la base de données ArrayExpress par Solheim *et al.*, 2007). Des expériences de RT-qPCR ciblant l'ARNm du gène *ef1518* chez le mutant  $\Delta 5'1515$  n'ont montré aucune variation d'expression comparativement à la souche sauvage.

Nous savons donc que la protéine EF1515 régule l'expression d'au moins deux gènes. Il est tout à fait possible qu'elle cible aussi d'autres régions sur le chromosome. Afin d'identifier un tel régulon, il peut être possible de mettre en place une expérience de co-immunoprécipitation *in vivo* d'EF1515 et de ses ARN associés, couplée par exemple à une étape de « crosslinking ». Le rayonnement UV permettra de créer des liaisons covalentes entre les protéines et l'acide nucléique associé et le complexe EF1515-ARN cible ainsi formé pourra être récupéré et analysé par séquençage.

L'autre 5'UTR étudiée, 5'*cshA*, régule l'expression d'une hélicase à motif DEAD codée par le gène *ef0846*, ou *cshA*. Nous avons montré que l'absence de CshA provoquait une augmentation de l'expression d'une protéine rapporteuse dans un contexte de fusion traductionnelle. Il a aussi été montré que cette protéine se fixe à sa propre 5'UTR. Il peut en être déduit que CshA est capable de s'auto-réprimer, probablement au travers de sa capacité à défaire des structures secondaires de sa 5'UTR. Il est possible que CshA permette la rupture des structures tiges-boucles se formant spontanément sur 5'*cshA* et que cela aboutisse à l'apparition d'autres structures qui bloqueraient le RBS. De plus, les expériences de dosage de la  $\beta$ -glucuronidase ont montré que les 100 premières paires de base de l'ORF *cshA* jouaient un rôle crucial dans sa régulation. En effet, leur délétion entraîne un défaut majeur dans l'expression de la protéine rapporteuse.

Des résultats complémentaires ont montré que l'expression de *cshA* était réduite de manière très marquée en l'absence d'une autre hélicase à motif DEAD, CshC, et de 5'*cshA*. Cela montre

que *cshA* est soumis à d'autres modes de régulation, et que ceux-ci impliquent 5'*cshA* ainsi qu'une action inductrice de CshC.

D'autres résultats laissent supposer que la région chromosomique située entre *ef0845* et *cshA* correspondrait à la fois à la 3'UTR du premier gène et à la 5'UTR du deuxième. EF0845, ou MurF, est impliquée dans la synthèse de la paroi, et la délétion de 5'*cshA* a entraîné la formation de nodules à la surface des bactéries. Il est donc possible que ce phénotype soit lié à une modification de l'expression de MurF, et cet aspect reste encore à explorer.

Concernant le rôle de CshA, nous avons observé sa surexpression dans un large éventail de stress (stress acide, stress thermique froid, stress oxydatif, présence de sels biliaries). Nous avons également mis en évidence un défaut de croissance et de virulence d'*E. faecalis* en l'absence de CshA. Tout cela montre le rôle central de cette hélicase dans la réponse au stress, la virulence, et le fitness de la bactérie.

Il est connu que chez *B. subtilis*, CshA est un composant du dégradosome (Lehnik-Habrink *et al.*, 2010). Nous avons pu démontrer que chez *E. faecalis*, CshA interagit en priorité avec l'énolase, une autre protéine du dégradosome, probablement *via* son extrémité N-terminale. Sa ou ses fonctions exactes restent encore à déterminer. Nous ne savons effectivement pas si elle intervient dans la séparation des brins d'ARN, dans le sens 5'-3', 3'-5' ou les deux, ni si elle possède une activité de liaison des brins d'ARN. Cela pourrait être vérifié en mettant CshA en présence d'ARN double brin et en observant la quantité d'ARN simple brin qui en découle ou inversement en présence d'ARNs simple brins complémentaires marqués pour visualiser la formation de double brin, comme cela a été effectué par Chamot *et al.*, 2005.

Parallèlement à ce travail, nous avons également étudié les RNases, protéines clés du dégradosome, ainsi que les autres hélicases à motif DEAD d'*E. faecalis*. La RNase Y s'est démarquée par le fait qu'elle est induite dans toutes les conditions de stress testées, et que sa

délétion entraîne un défaut de croissance et de virulence chez la bactérie. Elle est donc probablement impliquée dans le fitness de la bactérie. La RNase J1 est essentielle chez *B. subtilis* et nos résultats suggèrent que c'est aussi le cas chez *E. faecalis*. Les RNases J2 et III sont toutes les deux impliquées dans la virulence, la réponse aux stress froid, oxydatif et aux sels biliaires.

Quant aux deux autres hélicases à motif DEAD, CshB et CshC, elles semblent être des protéines spécifiques de la réponse au stress froid, ce qui est une caractéristique fréquemment retrouvée lors d'études sur les hélicases à motif DEAD (Resch *et al.*, 2010; Pandiani *et al.*, 2011; El Mortaji *et al.*, 2018).

Ce travail a pu montrer que les trois ARN étudiés, sélectionnés d'après leur induction en condition d'infection, possédaient des modes d'action très diversifiés (Figure 26). En effet, nous avons trouvé : un sARN (SRC65) agissant en *trans* probablement en redondance avec un autre sARN, une 5'UTR (5'1515) réprimant l'expression de son opéron *via* la formation d'un terminateur de transcription, et enfin une autre 5'UTR (5'*cshA*) participant à l'auto-régulation du gène co-transcrit, mais aussi à la régulation par d'autres acteurs comme par exemple CshC.

On remarquera que les deux 5'UTR étudiées fonctionnent par deux mécanismes différents : 5'1515 a une action répressive avec la terminaison prématurée de la transcription de l'opéron *ef1515-ef1516*. En revanche, l'influence de 5'*cshA* dépend probablement de l'action de différents régulateurs (CshA semble être répresseur, et CshC semble être inducteur).

L'ensemble des travaux présentés dans ce manuscrit a permis de mettre en lumière la grande diversité des modes de régulation par les ARN régulateurs chez *E. faecalis*.



## Chapitre II

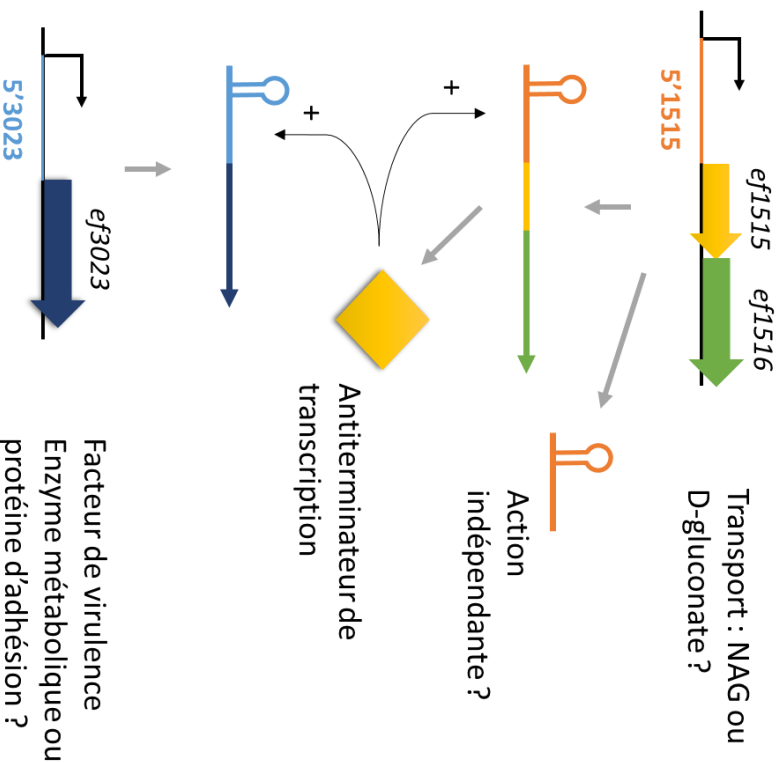
Etude du rôle et de l'action d'un petit ARN régulateur agissant en trans



Réponse générale au stress  
Déclenchement de la phase  
exponentielle de croissance

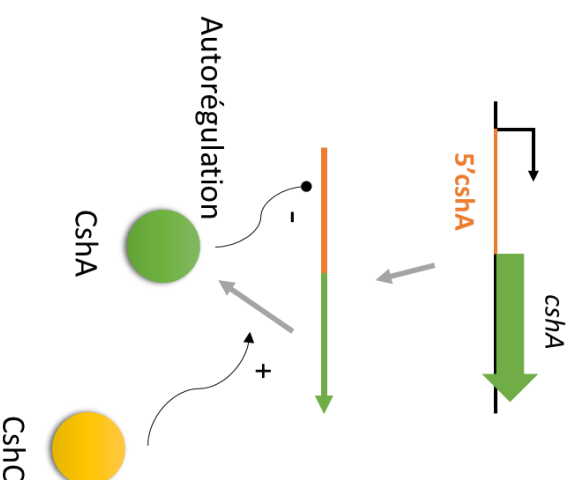
## Chapitre III

Etude des mécanismes de régulation d'une 5'UTR possédant un terminateur de transcription



## Chapitre IV

Caractérisation d'une hélicase à motif DEAD régulée par une région 5' non traduite (5'UTR)



Réponse générale au stress  
Fitness  
Interaction avec l'énolase

Figure 26 : Schéma récapitulatif des résultats concernant les trois ARN régulateurs étudiés.

# Références

- Aguirre, A.A., Vicente, A.M., Hardwick, S.W., Alvelos, D.M., Mazzon, R.R., Luisi, B.F., and Marques, M.V. (2017) Association of the cold-shock DEAD-box RNA helicase RhlE to the RNA degradosome in *Caulobacter crescentus*. *J Bacteriol* **JB.00135-17**.
- Amster-Choder, O. (2005) The *bgl* sensory system: a transmembrane signaling pathway controlling transcriptional antitermination. *Current Opinion in Microbiology* **8**: 127–134.
- Andrewes, F.W. and Horder, T.J. (1906) A study of streptococci pathogenic for man. *Lancet*, **2**, 708–713.
- Aranda, P.S., LaJoie, D.M., and Jorcyk, C.L. (2012) Bleach gel: A simple agarose gel for analyzing RNA quality. *ELECTROPHORESIS* **33**: 366–369.
- Arias, C.A., and Murray, B.E. (2012) The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat Rev Microbiol* **10**: 266–278.
- Arnaud, M., Chastanet, A., and Débarbouillé, M. (2004) New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. *Appl Environ Microbiol* **70**: 6887–6891.
- Attaiech, L., Glover, J.N.M., and Charpentier, X. (2017) RNA Chaperones Step Out of Hfq's Shadow. *Trends in Microbiology* **0** [http://www.cell.com/trends/microbiology/abstract/S0966-842X\(17\)30018-5](http://www.cell.com/trends/microbiology/abstract/S0966-842X(17)30018-5). Accessed February 10, 2017.
- Babady, N.E. (2016) Hospital-Associated Infections. *Microbiology Spectrum* **4** <http://www.asmscience.org/content/journal/microbiolspec/10.1128/microbiolspec.DMIH2-0003-2015>. Accessed June 27, 2017.
- Bae, T., Kozłowicz, B., and Dunny, G.M. (2002) Two targets in pCF10 DNA for PrgX binding: their role in production of Qa and prgX mRNA and in regulation of pheromone-inducible conjugation11. *Journal of Molecular Biology* **315**: 995–1007.
- Boisset, S., Geissmann, T., Huntzinger, E., Fechter, P., Bendridi, N., Possedko, M., *et al.* (2007) *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev* **21**: 1353–1366.
- Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248–254.
- Brinster, S., Furlan, S., and Serror, P. (2007) C-terminal WxL domain mediates cell wall binding in *Enterococcus faecalis* and other gram-positive bacteria. *J Bacteriol* **189**: 1244–1253.
- Brock, T.D., Madigan, M.T., and Martinko, J.M. (2007) *Brock, Biologie des micro-organismes*. Pearson Education, Paris.

- Cacaci, M., Giraud, C., Leger, L., Torelli, R., Martini, C., Posteraro, B., *et al.* (2018) Expression profiling in a mammalian host reveals the strong induction of genes encoding LysM domain-containing proteins in *Enterococcus faecium*. *Sci Rep* **8**: 12412.
- Carpousis, A.J. (2007) The RNA degradosome of *Escherichia coli*: an mRNA-degrading machine assembled on RNase E. *Annu Rev Microbiol* **61**: 71–87.
- Caswell, C.C., Gaines, J.M., Ciborowski, P., Smith, D., Borchers, C.H., Roux, C.M., *et al.* (2012) Identification of two small regulatory RNAs linked to virulence in *Brucella abortus* 2308. *Molecular Microbiology* **85**: 345–360.
- Chamot, D., Colvin, K.R., Kujat-Choy, S.L., and Owtrim, G.W. (2005) RNA structural rearrangement via unwinding and annealing by the cyanobacterial RNA helicase, CrhR. *J Biol Chem* **280**: 2036–2044.
- Charpentier, E. (2015) CRISPR-Cas9: how research on a bacterial RNA-guided mechanism opened new perspectives in biotechnology and biomedicine. *EMBO Mol Med* **7**: 363–365.
- Cordin, O., Banroques, J., Tanner, N.K., and Linder, P. (2006) The DEAD-box protein family of RNA helicases. *Gene* **367**: 17–37.
- Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* **16**: 10881–10890.
- Dam, P., Olman, V., Harris, K., Su, Z., and Xu, Y. (2007) Operon prediction using both genome-specific and general genomic information. *Nucleic Acids Res* **35**: 288–298.
- Davis, M.C., Kesthely, C.A., Franklin, E.A., and MacLellan, S.R. (2017) The essential activities of the bacterial sigma factor. *Can J Microbiol* **63**: 89–99.
- DebRoy, S., Gao, P., Garsin, D.A., Harvey, B.R., Kos, V., Nes, I.F., and Solheim, M. (2014a) Transcriptional and Post Transcriptional control of enterococcal gene regulation. In *enterococci: From commensals to leading causes of drug resistant infection*. Gilmore, M.S., Clewell, D.B., Ike, Y., and Shankar, N. (eds). Massachusetts Eye and Ear Infirmary, Boston. <http://www.ncbi.nlm.nih.gov/books/NBK190422/>. Accessed February 16, 2016.
- DebRoy, S., Gebbie, M., Ramesh, A., Goodson, J.R., Cruz, M.R., Hoof, A. van, *et al.* (2014b) A riboswitch-containing sRNA controls gene expression by sequestration of a response regulator. *Science* **345**: 937–940.
- Del Giudice, M.G., Ugalde, J.E., and Czubener, C. (2013) A lysozyme-like protein in *Brucella abortus* is involved in the early stages of intracellular replication. *Infect Immun* **81**: 956–964.
- Del Papa, M.F., and Perego, M. (2008) Ethanolamine activates a sensor histidine kinase regulating its utilization in *Enterococcus faecalis*. *J Bacteriol* **190**: 7147–7156.
- El Mortaji, L., Aubert, S., Galtier, E., Schmitt, C., Anger, K., Redko, Y., *et al.* (2018) The sole DEAD-box RNA helicase of the gastric pathogen *Helicobacter pylori* is essential for colonization. *MBio* **9**.
- El Mouali, Y., and Balsalobre, C. (2018) 3'untranslated regions: regulation at the end of the road. *Curr Genet* .

- Fouquier d'Hérouel, 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 Res* **39**: e46.
- Frankenberg, L., Brugna, M., and Hederstedt, L. (2002) *Enterococcus faecalis* heme-dependent catalase. *J Bacteriol* **184**: 6351–6356.
- Fritzenwanker, M., Chakraborty, A., Hain, T., Zimmermann, K., and Domann, E. (2016) Draft genome sequences of the probiotic *Enterococcus faecalis* Symbioflor 1 clones DSM16430 and DSM16434. *Genome Announc* **4**.
- Galinier, A., Haiech, J., Kilhoffer, M.C., Jaquinod, M., Stülke, J., Deutscher, J., and Martin-Verstraete, I. (1997) The *Bacillus subtilis* crh gene encodes a HPr-like protein involved in carbon catabolite repression. *Proc Natl Acad Sci USA* **94**: 8439–8444.
- Galloway-Peña, J.R., Liang, X., Singh, K.V., Yadav, P., Chang, C., La Rosa, S.L., *et al.* (2015) The identification and functional characterization of WxL proteins from *Enterococcus faecium* reveal surface proteins involved in extracellular matrix interactions. *J Bacteriol* **197**: 882–892.
- Gao, W., Howden, B.P., and Stinear, T.P. (2017) Evolution of virulence in *Enterococcus faecium*, a hospital-adapted opportunistic pathogen. *Curr Opin Microbiol* **41**: 76–82.
- Giard, J.-C., Hartke, A., Flahaut, S., Benachour, A., Boutibonnes, P., and Auffray, Y. (1996) Starvation-induced multiresistance in *Enterococcus faecalis* JH2-2. *Curr Microbiol* **32**: 264–271.
- Giard, J.C., Laplace, J.M., Rincé, A., Pichereau, V., Benachour, A., Leboeuf, C., *et al.* (2001) The stress proteome of *Enterococcus faecalis*. *Electrophoresis* **22**: 2947–2954.
- Gilmore, M.S., Clewell, D.B., Ike, Y., and Shankar, N. (eds) (2014) *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*. Massachusetts Eye and Ear Infirmary, Boston. <http://www.ncbi.nlm.nih.gov/books/NBK190424/>. Accessed June 21, 2018.
- Gottesman, S., and Storz, G. (2011) Bacterial small RNA regulators: Versatile roles and rapidly evolving variations. *Cold Spring Harbor Perspectives in Biology* **3**: a003798–a003798.
- Hartke, A., Giard, J.-C., Laplace, J.-M., and Auffray, Y. (1998) Survival of *Enterococcus faecalis* in an oligotrophic microcosm: Changes in morphology, development of general stress resistance, and analysis of protein synthesis. *Appl Environ Microbiol* **64**: 4238–4245.
- Heurlier, K., Williams, F., Heeb, S., Dormond, C., Pessi, G., Singer, D., *et al.* (2004) Positive control of swarming, rhamnolipid synthesis, and lipase production by the posttranscriptional RsmA/RsmZ system in *Pseudomonas aeruginosa* PAO1. *J Bacteriol* **186**: 2936–2945.
- Idelson, M., and Amster-Choder, O. (1998) SacY, a transcriptional antiterminator from *Bacillus subtilis*, is regulated by phosphorylation in vivo. *J Bacteriol* **180**: 660–666.
- Innocenti, N., Golumbeanu, M., Hérouël, A.F. d', 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.

- Jamalli, A., Hébert, A., Zig, L., and Putzer, H. (2014) Control of expression of the RNases J1 and J2 in *Bacillus subtilis*. *J Bacteriol* **196**: 318–324.
- Jia, X., Zhang, J., Sun, W., He, W., Jiang, H., Chen, D., and Murchie, A.I.H. (2013) Riboswitch Control of Aminoglycoside Antibiotic Resistance. *Cell* **152**: 68–81.
- Johnson, C.M., Haemig, H.H.A., Chatterjee, A., Wei-Shou, H., Weaver, K.E., and Dunny, G.M. (2011) RNA-mediated reciprocal regulation between two bacterial operons is RNase III dependent. *MBio* **2**.
- Jones, B.E., Dossonnet, V., Küster, E., Hillen, W., Deutscher, J., and Klevit, R.E. (1997) Binding of the catabolite repressor protein CcpA to its DNA target is regulated by phosphorylation of its corepressor HPr. *J Biol Chem* **272**: 26530–26535.
- Kaberdin, V.R., and Bläsi, U. (2013) Bacterial helicases in post-transcriptional control. *Biochim Biophys Acta* **1829**: 878–883.
- Käll, L., Krogh, A., and Sonnhammer, E.L.L. (2004) A combined transmembrane topology and signal peptide prediction method. *J Mol Biol* **338**: 1027–1036.
- Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**: D353–D361.
- Kery, M.B., Feldman, M., Livny, J., and Tjaden, B. (2014) TargetRNA2: identifying targets of small regulatory RNAs in bacteria. *Nucleic Acids Res* **42**: W124–W129.
- Khemic, V., and Carpousis, A.J. (2003) The RNA degradosome and poly(A) polymerase of *Escherichia coli* are required in vivo for the degradation of small mRNA decay intermediates containing REP-stabilizers. *Molecular Microbiology* **51**: 777–790.
- Khemic, V., and Linder, P. (2016) RNA helicases in bacteria. *Curr Opin Microbiol* **30**: 58–66.
- Kim, J.-Y., Awji, E.G., Park, N.-H., Park, J.-Y., Kim, J.-C., Lee, S.-P., *et al.* (2017) Probiotic properties and adsorption of *Enterococcus faecalis* PSCT3-7 to vermiculite. *J Vet Sci* **18**: 95–99.
- Kortmann, J., and Narberhaus, F. (2012) Bacterial RNA thermometers: molecular zippers and switches. *Nat Rev Microbiol* **10**: 255–265.
- Laalami, S., and Putzer, H. (2011) mRNA degradation and maturation in prokaryotes: the global players. *BioMolecular Concepts* **2** <https://www.degruyter.com/view/j/bmc.2011.2.issue-6/bmc.2011.042/bmc.2011.042.xml>. Accessed January 18, 2019.
- Ladjouzi, R., Bizzini, A., Schaik, W. van, Zhang, X., Rincé, A., Benachour, A., and Hartke, A. (2015) Loss of antibiotic tolerance in Sod-deficient mutants is dependent on the energy source and arginine catabolism in enterococci. *J Bacteriol* **197**: 3283–3293.
- Lancefield, R.C. (1933) A Serological differentiation of human and other groups of hemolytic streptococci. *J Exp Med* **57**: 571–595.

- Latorre, M., Quenti, D., Travisany, D., Singh, K.V., Murray, B.E., Maass, A., and Cambiazo, V. (2018) The role of Fur in the transcriptional and iron homeostatic response of *Enterococcus faecalis*. *Front Microbiol* **9**: 1580.
- Lebreton, F., Schaik, W. van, Sanguinetti, M., Posteraro, B., Torelli, R., Le Bras, F., *et al.* (2012) AsrR is an oxidative stress sensing regulator modulating *Enterococcus faecium* opportunistic traits, antimicrobial resistance, and pathogenicity. *PLoS Pathog* **8**: e1002834.
- Lehnik-Habrink, M., Lewis, R.J., Mäder, U., and Stülke, J. (2012) RNA degradation in *Bacillus subtilis*: an interplay of essential endo- and exoribonucleases. *Mol Microbiol* **84**: 1005–1017.
- Lehnik-Habrink, M., Newman, J., Rothe, F.M., Solovyova, A.S., Rodrigues, C., Herzberg, C., *et al.* (2011) RNase Y in *Bacillus subtilis*: a natively disordered protein that is the functional equivalent of RNase E from *Escherichia coli*. *J Bacteriol* **193**: 5431–5441.
- Lehnik-Habrink, M., Pförtner, H., Leonie, R., Pietack Nico, Herzberg Christina, and Stülke Jörg (2010) The RNA degradosome in *Bacillus subtilis*: identification of CshA as the major RNA helicase in the multiprotein complex. *Molecular Microbiology* **77**: 958–971.
- Lenz, D.H., Mok, K.C., Lilley, B.N., Kulkarni, R.V., Wingreen, N.S., and Bassler, B.L. (2004) The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell* **118**: 69–82.
- Levin-Reisman, I., Ronin, I., Gefen, O., Braniss, I., Shores, N., and Balaban, N.Q. (2017) Antibiotic tolerance facilitates the evolution of resistance. *Science* **355**: 826–830.
- Li, R.W. (ed.) (2011) *Metagenomics and its applications in agriculture, biomedicine, and environmental studies*. Nova Science Publisher's, Hauppauge, N.Y.
- Linares, D.M., Perez, M., Ladero, V., Rio, B. del, Redruello, B., Martin, M.C., *et al.* (2014) An agmatine-inducible system for the expression of recombinant proteins in *Enterococcus faecalis*. *Microbial cell factories* **13**: 1–9.
- Linder, P., and Jankowsky, E. (2011) From unwinding to clamping — the DEAD box RNA helicase family. *Nature Reviews Molecular Cell Biology* **12**: 505–516.
- Lioliou, E., Sharma, C.M., Caldelari, I., Helfer, A.-C., Fechter, P., Vandenesch, F., *et al.* (2012) Global regulatory functions of the *Staphylococcus aureus* endoribonuclease III in gene expression. *PLoS Genet* **8**: e1002782.
- Luo, M.L., Leenay, R.T., and Beisel, C.L. (2016) Current and future prospects for CRISPR-based tools in bacteria. *Biotechnol Bioeng* **113**: 930–943.
- Maccallum W.G., and Hastings T.W. (1899) A case of acute endocarditis caused by *Micrococcus zymogenes* (nov. spec.), with a description of the microorganism. *J Exp Med*. 1899 Sep 1;4(5-6):521-34.
- Madden, T. (2003) *The BLAST Sequence Analysis Tool*. National Center for Biotechnology Information (US), . <https://www.ncbi.nlm.nih.gov/books/NBK21097/>. Accessed November 21, 2018.

- Madsen, K.T., Skov, M.N., Gill, S., and Kemp, M. (2017) Virulence factors associated with *Enterococcus Faecalis* infective endocarditis: a mini review. *Open Microbiol J* **11**: 1–11.
- Mao, F., Dam, P., Chou, J., Olman, V., and Xu, Y. (2009) DOOR: a database for prokaryotic operons. *Nucleic Acids Res* **37**: D459-463.
- Marciniak, B.C., Pabijaniak, M., Jong, A. de, Dühring, R., Seidel, G., Hillen, W., and Kuipers, O.P. (2012) High- and low-affinity cre boxes for CcpA binding in *Bacillus subtilis* revealed by genome-wide analysis. *BMC Genomics* **13**: 401.
- Martin, J.D., and Mundt, J.O. (1972) Enterococci in insects. *Appl Microbiol* **24**: 575–580.
- Maugat, S., Pontiès, V., Colomb-Cotinac, M., Soing-Altrach, S., Subiros, M., Bernet, C., *et al.* (2018) Bilan 2001-2017 des signalements externes d'infections nosocomiales. Part des signalements impliquant une bactérie multirésistante, hautement résistante-émergente ou un *Clostridium difficile*. [http://invs.santepubliquefrance.fr/beh/2018/25-26/2018\\_25-26\\_2.html](http://invs.santepubliquefrance.fr/beh/2018/25-26/2018_25-26_2.html).
- Mendonça, M., Moreira, G.M.S.G., Conceição, F.R., Hust, M., Mendonça, K.S., Moreira, Â.N., *et al.* (2016) Fructose 1,6-bisphosphate aldolase, a novel immunogenic surface protein on *Listeria* Species. *PLoS ONE* **11**: e0160544.
- Meyer, F.M., Jules, M., Mehne, F.M.P., Le Coq, D., Landmann, J.J., Görke, B., *et al.* (2011) Malate-mediated carbon catabolite repression in *Bacillus subtilis* involves the HPrK/CcpA pathway. *J Bacteriol* **193**: 6939–6949.
- Michaux, C. (2013) Identification et caractéristique fonctionnelle de petits ARN non codants chez *Enterococcus faecalis* et analyse d'une protéine "RNA-binding.", Thèse de doctorat, Université de Caen-Normandie.
- 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. *Infection and Immunity* **82**: 3599–3611.
- Michaux, C., Martini, C., Shioya, K., Lecheheb, S.A., Budin-Verneuil, A., Cosette, P., *et al.* (2012) CspR, a cold shock RNA-binding protein involved in the long-term survival and the virulence of *Enterococcus faecalis*. *J Bacteriol* **194**: 6900–6908.
- Michaux, C., Verneuil, N., Hartke, A., and Giard, J.-C. (2014b) Physiological roles of small RNA molecules. *Microbiology* **160**: 1007–1019.
- Miller, W.R., Murray, B.E., Rice, L.B., and Arias, C.A. (2016) Vancomycin-resistant enterococci: therapeutic challenges in the 21st century. *Infectious Disease Clinics of North America* **30**: 415–439.
- Mitchell, A.L., Attwood, T.K., Babbitt, P.C., Blum, M., Bork, P., Bridge, A., *et al.* (2018) InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res* .
- Mizuno, T., Chou, M.Y., and Inouye, M. (1984) A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proc Natl Acad Sci USA* **81**: 1966–1970.

- Monedero, V., Poncet, S., Mijakovic, I., Fieulaine, S., Dossonnet, V., Martin-Verstraete, I., *et al.* (2001) Mutations lowering the phosphatase activity of HPr kinase/phosphatase switch off carbon metabolism. *EMBO J* **20**: 3928–3937.
- Monedero, V., Revilla-Guarinos, A., and Zúñiga, M. (2017) Physiological role of two-component signal transduction systems in food-associated lactic acid bacteria. *Adv Appl Microbiol* **99**: 1–51.
- Morita, T., Maki, K., and Aiba, H. (2005) RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev* **19**: 2176–2186.
- Muller, C., Cacaci, M., Sauvageot, N., Sanguinetti, M., Rattei, T., Eder, T., *et al.* (2015) The intraperitoneal transcriptome of the opportunistic pathogen *Enterococcus faecalis* in mice. *PLOS ONE* **10**: e0126143.
- Murray, B.E. (1990) The life and times of the *Enterococcus*. *Clin Microbiol Rev* **3**: 46–65.
- Mushegian, A.R., Fullner, K.J., Koonin, E.V., and Nester, E.W. (1996) A family of lysozyme-like virulence factors in bacterial pathogens of plants and animals. *Proc Natl Acad Sci USA* **93**: 7321–7326.
- Nami, Y., Abdullah, N., Haghshenas, B., Radiah, D., Rosli, R., and Yari Khosroushahi, A. (2014) A newly isolated probiotic *Enterococcus faecalis* strain from vagina microbiota enhances apoptosis of human cancer cells. *J Appl Microbiol* **117**: 498–508.
- Newman, J.A., Hewitt, L., Rodrigues, C., Solovyova, A.S., Harwood, C.R., and Lewis, R.J. (2012) Dissection of the network of interactions that links RNA processing with glycolysis in the *Bacillus subtilis* degradosome. *Journal of Molecular Biology* **416**: 121–136.
- Oliva, G., Sahr, T., and Buchrieser, C. (2015) Small RNAs, 5' UTR elements and RNA-binding proteins in intracellular bacteria: impact on metabolism and virulence. *FEMS Microbiol Rev* **39**: 331–349.
- Ostrolenk, M., Kramer, N., and Cleverdon, R.C. (1947) Comparative studies of enterococci and *Escherichia coli* as indices of pollution. *J Bacteriol* **53**: 197–203.
- Owttrim, G.W. (2013) RNA helicases. *RNA Biology* **10**: 96–110.
- Pagliai, F.A., Murdoch, C.C., Brown, S.M., Gonzalez, C.F., and Lorca, G.L. (2014) A dual role of the transcriptional regulator TstR provides insights into cyanide detoxification in *Lactobacillus brevis*. *Mol Microbiol* **92**: 853–871.
- Pandiani, F., Chamot, S., Brillard, J., Carlin, F., Nguyen-the, C., and Broussolle, V. (2011) Role of the five RNA helicases in the adaptive response of *Bacillus cereus* ATCC 14579 cells to temperature, pH, and oxidative stresses. *Appl Environ Microbiol* **77**: 5604–5609.
- Paulsen, I.T., Banerjee, L., Myers, G.S.A., Nelson, K.E., Seshadri, R., Read, T.D., *et al.* (2003) Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* **299**: 2071–2074.



- Postma, P.W., Lengeler, J.W., and Jacobson, G.R. (1993) Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol Rev* **57**: 543–594.
- Ran, S., Chu, M., Gu, S., Wang, J., and Liang, J. (2018) *Enterococcus faecalis* induces apoptosis and pyroptosis of human osteoblastic MG63 cells via the NLRP3 inflammasome. *Int Endod J*.
- Ren, G.-X., Guo, X.-P., and Sun, Y.-C. (2017) Regulatory 3' untranslated regions of bacterial mRNAs. *Frontiers in Microbiology* **8**  
<http://journal.frontiersin.org/article/10.3389/fmicb.2017.01276/full>. Accessed September 24, 2018.
- Resch, A., Večerek, B., Palavra, K., and Bläsi, U. (2010) Requirement of the CsdA DEAD-box helicase for low temperature riboregulation of rpoS mRNA. *RNA Biology* **7**: 796–802.
- Rincé, A., Le Breton, Y., Verneuil, N., Giard, J.-C., Hartke, A., and Auffray, Y. (2003) Physiological and molecular aspects of bile salt response in *Enterococcus faecalis*. *International Journal of Food Microbiology* **88**: 207–213.
- Roux, C.M., DeMuth, J.P., and Dunman, P.M. (2011) Characterization of components of the *Staphylococcus aureus* mRNA degradosome holoenzyme-like complex. *J Bacteriol* **193**: 5520–5526.
- Ruiz de los Mozos, I., Vergara-Irigaray, M., Segura, V., Villanueva, M., Bitarte, N., Saramago, M., *et al.* (2013) Base pairing interaction between 5'- and 3'-UTRs controls icaR mRNA translation in *Staphylococcus aureus*. *PLoS Genet* **9**: e1004001.
- Santiago-Frangos, A., and Woodson, S.A. (2018) Hfq chaperone brings speed dating to bacterial sRNA. *Wiley Interdiscip Rev RNA* e1475.
- Schleifer, K.H., and 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.
- Shams, F., Oldfield, N.J., Lai, S.K., Tunio, S.A., Wooldridge, K.G., and Turner, D.P.J. (2016) Fructose-1,6-bisphosphate aldolase of *Neisseria meningitidis* binds human plasminogen via its C-terminal lysine residue. *Microbiologyopen* **5**: 340–350.
- Shankar, N., Baghdayan, A.S., and Gilmore, M.S. (2002) Modulation of virulence within a pathogenicity island in vancomycin-resistant *Enterococcus faecalis*. *Nature* **417**: 746–750.
- Sherman, J.M. (1937) The streptococci. *Bacteriol Rev* **1**: 3–97.
- Shimizu, T., Yaguchi, H., Ohtani, K., Banu, S., and Hayashi, H. (2002) Clostridial VirR/VirS regulon involves a regulatory RNA molecule for expression of toxins. *Mol Microbiol* **43**: 257–265.
- 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**: e23948.

- Shiroda, M., Pratt, Z.L., Döpfer, D., Wong, A.C.L., and Kaspar, C.W. (2014) RpoS impacts the lag phase of *Salmonella enterica* during osmotic stress. *FEMS Microbiol Lett* **357**: 195–200.
- Sillanpaa, J., Xu, Y., Nallapareddy, S.R., Murray, B.E., and Höök, M. (2004) A family of putative MSCRAMMs from *Enterococcus faecalis*. *Microbiology* **150**: 2069–2078.
- Solheim, M., Aakra, A., Vebø, H., Snipen, L., and Nes, I.F. (2007) Transcriptional responses of *Enterococcus faecalis* V583 to bovine bile and sodium dodecyl sulfate. *Appl Environ Microbiol* **73**: 5767–5774.
- Stork, M., Di Lorenzo, M., Welch, T.J., and Crosa, J.H. (2007) Transcription termination within the iron transport-biosynthesis operon of *Vibrio anguillarum* requires an antisense RNA. *J Bacteriol* **189**: 3479–3488.
- Storz, G., Opdyke, J.A., and Zhang, A. (2004) Controlling mRNA stability and translation with small, noncoding RNAs. *Current Opinion in Microbiology* **7**: 140–144.
- Suárez, C.A., Blancato, V.S., Poncet, S., Deutscher, J., and Magni, C. (2011) CcpA represses the expression of the divergent cit operons of *Enterococcus faecalis* through multiple cre sites. *BMC Microbiol* **11**: 227.
- Sun, X., Zhulin, I., and Wartell, R.M. (2002) Predicted structure and phyletic distribution of the RNA-binding protein Hfq. *Nucleic Acids Res* **30**: 3662–3671.
- Suzuki, K., Shimizu, M., Sasaki, N., Ogawa, C., Minami, H., Sugimoto, H., and Watanabe, T. (2016) Regulation of the chitin degradation and utilization system by the ChiX small RNA in *Serratia marcescens* 2170. *Bioscience, Biotechnology, and Biochemistry* **80**: 376–385.
- Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D.L., *et al.* (2018) Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect Dis* **18**: 318–327.
- Thiercelin, M.E. (1899) Sur un diplocoque saprophyte de l'intestin susceptible de devenir pathogene. *Comptes rendus des séances de la Société de biologie et de ses filiales* **51**: 269–271.
- Thurlow, L.R., Thomas, V.C., and Hancock, L.E. (2009) Capsular polysaccharide production in *Enterococcus faecalis* and contribution of CpsF to capsule serospecificity. *J Bacteriol* **191**: 6203–6210.
- Tomita, H., and Clewell, D.B. (2000) A pAD1-encoded small RNA molecule, mD, negatively regulates *Enterococcus faecalis* pheromone response by enhancing transcription termination. *J Bacteriol* **182**: 1062–1073.
- Tramonti, A., De Canio, M., and De Biase, D. (2008) GadX/GadW-dependent regulation of the *Escherichia coli* acid fitness island: transcriptional control at the gadY-gadW divergent promoters and identification of four novel 42 bp GadX/GadW-specific binding sites. *Mol Microbiol* **70**: 965–982.
- Troxell, B., and Hassan, H.M. (2013) Transcriptional regulation by Ferric Uptake Regulator (Fur) in pathogenic bacteria. *Front Cell Infect Microbiol* **3**: 59.

- Tsigrelis, C., Singh, K.V., Coutinho, T.D., Murray, B.E., and Baddour, L.M. (2006) Vancomycin-resistant *Enterococcus faecalis* endocarditis: linezolid failure and strain characterization of virulence factors. *Journal of Clinical Microbiology* **45**: 631–635.
- Uhlmann, J., Siemens, N., Kai-Larsen, Y., Fiedler, T., Bergman, P., Johansson, L., and Norrby-Teglund, A. (2016) Phosphoglycerate kinase-A novel streptococcal factor involved in neutrophil activation and degranulation. *J Infect Dis* **214**: 1876–1883.
- Van Tyne, D., and Gilmore, M.S. (2014) Friend turned foe: evolution of enterococcal virulence and antibiotic resistance. *Annu Rev Microbiol* **68**: 337–356.
- Vebø, H.C., Snipen, L., Nes, I.F., and Brede, D.A. (2009) The transcriptome of the nosocomial pathogen *Enterococcus faecalis* V583 reveals adaptive responses to growth in blood. *PLOS ONE* **4**: e7660.
- Vebø, H.C., Solheim, M., Snipen, L., Nes, I.F., and Brede, D.A. (2010) Comparative genomic analysis of pathogenic and probiotic *Enterococcus faecalis* isolates, and their transcriptional responses to growth in human urine. *PLoS ONE* **5**: e12489.
- Waters, L.S., and Storz, G. (2009) Regulatory RNAs in bacteria. *Cell* **136**: 615–628.
- Weaver, K.E. (2007) Emerging plasmid-encoded antisense RNA regulated systems. *Curr Opin Microbiol* **10**: 110–116.
- Weaver, K.E. (2015) The Type I toxin-antitoxin par locus from *Enterococcus faecalis* plasmid pAD1: RNA regulation by both cis- and trans-acting elements. *Plasmid* **78**: 65–70.
- 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. *Infection Control & Hospital Epidemiology* **37**: 1288–1301.
- Winstedt, L., Frankenberg, L., Hederstedt, L., and Wachenfeldt, C. von (2000) *Enterococcus faecalis* V583 contains a cytochrome bd-type respiratory oxidase. *J Bacteriol* **182**: 3863–3866.
- Yan, X., Budin-Verneuil, A., Verneuil, N., Gilmore, M.S., Artigaud, S., Auffray, Y., and Pichereau, V. (2015) Transcriptomic response of *Enterococcus faecalis* V583 to low hydrogen peroxide levels. *Curr Microbiol* **70**: 156–168.
- Yuen, G.J., and Ausubel, F.M. (2014) Enterococcus infection biology: lessons from invertebrate host models. *J Microbiol* **52**: 200–210.

## Annexe 1 : Résultats complémentaires de la publication 1 « Identification of the general stress stimulon related to colonization in *Enterococcus faecalis* »

**Table S1** : List of *E. faecalis* V19 genes considered for expression analysis. The name of the ORF or genes in the V583 strain, as well as the number of genes in each operon are indicated. The names of the proteins encoded by each gene considered are found in the “annotation” column and the proteins are classified according to their role in the cell in the column “functional category”. The FC represents the fold-change of the RNA expression between stress condition (Bile salts (BS), urine, serum, pH 5.8 (acid pH), or containing 1 mM H<sub>2</sub>O<sub>2</sub> (oxidative stress)) compared to V19 strain grown in BHI. The poorly or not expressed genes were filtered out. The genes differentially expressed in at least one stress condition, with a cut-off fold  $\geq 5$ ,  $p \leq 0.05$  and  $\log\text{CPM} > 1$  are indicated.

ORF	genes/ operon	Gene Name	Annotation	Functional category	BS	urine	Serum	acid	oxidative
EF0007	3	<i>rpsF</i>	ribosomal protein S6	Protein synthesis		-37.2			
EF0008	1	<i>ssb-1</i>	single-strand binding protein	DNA metabolism	-8.6		-8.8	-6.4	-5.6
EF0009	1	<i>rpsR</i>	ribosomal protein S18	Protein synthesis	-13.8	-19.5	-103	-6.1	
EF0017	1		ABC transporter, ATP-binding protein	Transport			-9.7	-6.5	
EF0019- EF0022	4	<i>mptBACD</i>	PTS system IIBC components mannose-specific	Transport				-10.4	6.7
EF0025	1		membrane protein, putative	cell envelope		6.9			
EF0026	1			Hypothetical or unknown protein			-65.2		
EF0051	1	<i>ispE</i>	4-diphosphocytidyl-2C-methyl-D-erythritol kinase	Fatty acids and phospholipids metabolism		-5.8	-52.8	-5.5	-6.5
EF0059	1	<i>glmU</i>	UDP-N-acetylglucosamine pyrophosphorylase	Purines, Pyrimidines, nucleosides, nucleotides		-18.3	-73.6		
EF0071	1		lipoprotein	cell envelope				11.6	
EF0073	1		transcriptional regulator, Cro/Ci family	Regulatory functions		-10.5			
EF0076	1		short chain dehydrogenase	Other metabolism				6.0	
EF0080	1	<i>gls24</i>	<i>gls24</i> protein	cellular process		27.2	7.7	19.1	
EF0081	1	<i>glsB</i>	membrane protein	cell envelope	-25.8				-37.2
EF0082	1		major facilitator family transporter	Transport	-46.3	-8.6	-6.1		-11.6
EF0095	1		lipoprotein	cell envelope		10.1	9.0		
EF0097	1		regulatory protein pfoR	Regulatory functions		-6.6		-5.2	
EF0097- EF0099	3	<i>sdhB-1-A-1</i>	L-serine dehydratase, seryl-tRNA synthetase	Amino-acid biosynthesis		-5.9			
EF0100	1	<i>serS-1</i>	seryl-tRNA synthetase	Protein synthesis		-7.6			
EF0102	1		transcriptional regulator, ArgR family	Regulatory functions		6.0			
EF0104- EF0106	3	<i>arcABC</i>	arginine deiminase, ornithine carbamoyltransferase, carbamate kinase	Amino-acid biosynthesis		6.0	6.7	8.3	
EF0107	1		Crp/Fnr family transcriptional regulator	Regulatory functions		6.0	6.7		
EF0108	1		C4-dicarboxylate transporter, putative	Transport		6.0	6.7		
EF0110	1		transcriptional regulator, ArsR family	Regulatory functions	6.1				
EF0111	1		oxidoreductase, zinc-binding	central and intermediary metabolism		62.8	9.8	10.1	
EF0117	1		transcriptional regulator, GntR family	Regulatory functions		6.5	19.3		
EF0120	1			Hypothetical or unknown protein	-31.9	-11.5	-31.1	-12.5	-160.4
EF0121	1			Hypothetical or unknown protein					-6.4
EF0125	1		IS256, transposase	Mobile elements function		-31.5	-28.6	19.5	
EF0143	1		transcriptional regulator, Cro/Ci family	Regulatory functions				8.8	
EF0145	1			Hypothetical or unknown protein	-16.5				-15.2
EF0149	1		aggregation substance	cellular process	-24.6	19.1	10.1		-99.4
EF0153	1		cell wall surface anchor family protein	cell envelope			9.8	7.3	
EF0159	1		membrane protein, putative	cell envelope					-7.9
EF0164	1		lipoprotein	cell envelope	-25.6	-5.9	-6.7	-14.1	-35.2
EF0167	1	<i>guaA</i>	GMP synthase	Purines, Pyrimidines, nucleosides, nucleotides			-25.0		
EF0173	1	<i>pyn</i>	pyrimidine-nucleoside phosphorylase	Purines, Pyrimidines, nucleosides, nucleotides		11.2			
EF0174	1	<i>deoC</i>	deoxyribose-phosphate aldolase	central and intermediary metabolism			7.9		
EF0175	1	<i>cdd</i>	cytidine deaminase	Purines, Pyrimidines, nucleosides, nucleotides		6.7			
EF0176- EF0177	2		ABC-type sugar transport system	Transport			8.8		

EF0185- EF0187	3	<i>deoBD-1-D-2</i>	phosphopentomutase, purine nucleoside phosphorylase,	Purines, Pyrimidines, nucleosides, nucleotides		19.1	9.6		
EF0191- EF0193	3		ferrichrome ABC transporter	Transport	-6.3				
EF0195	1	<i>gpm</i>	phosphoglycerate mutase 1	central and intermediary metabolism			-9.7		
EF0198	2	<i>rpsL</i>	ribosomal protein S12	Protein synthesis					43.7
EF0199	1	<i>rpsG</i>	ribosomal protein S7	Protein synthesis			-5.7		12.5
EF0200	1	<i>fusA</i>	translation elongation factor G	Protein synthesis			-7.1		
EF0201	1	<i>tuf</i>	translation elongation factor Tu	Protein synthesis		-14.0	-13.4	-9.6	
EF0202	1		phosphomethylpyrimidine kinase	Purines, Pyrimidines, nucleosides, nucleotides		17.5	5.1		
EF0205	1	<i>rpsI</i>	ribosomal protein S10	Protein synthesis					5.1
EF0206	1	<i>rplC</i>	ribosomal protein L3	Protein synthesis		-5.0		-6.3	5.4
EF0207	1	<i>rplD</i>	ribosomal protein L4	Protein synthesis		-5.7		-7.7	
EF0209	1	<i>rplB</i>	ribosomal protein L2	Protein synthesis					7.3
EF0210	1	<i>rpsS</i>	ribosomal protein S19	Protein synthesis		-12.4	-18.1	-6.2	-30.0
EF0211	1	<i>rplV</i>	ribosomal protein L22	Protein synthesis					-11.0
EF0212	1	<i>rpsC</i>	ribosomal protein S3	Protein synthesis		-11.3	-11.3	-26.1	
EF0213	1	<i>rplP</i>	ribosomal protein L16	Protein synthesis					5.6
EF0214	1	<i>rpmC</i>	ribosomal protein L29	Protein synthesis					6.4
EF0215	1	<i>rpsQ</i>	ribosomal protein S17	Protein synthesis					-5.0
EF0216	1	<i>rplN</i>	ribosomal protein L14	Protein synthesis		-8.3	-7.5	-9.6	
EF0217	1	<i>rplX</i>	ribosomal protein L24	Protein synthesis		-12.5			
EF0218	1	<i>rplE</i>	ribosomal protein L5	Protein synthesis					8.1
EF0220	1	<i>rpsH</i>	ribosomal protein S8	Protein synthesis			-6.5	-5.1	
EF0221	1	<i>rplF</i>	ribosomal protein L6	Protein synthesis		-12.7	-11.5	-8.6	
EF0223	1	<i>rplR</i>	ribosomal protein L18	Protein synthesis		-5.6		-6.5	12.8
EF0224	1	<i>rpsE</i>	ribosomal protein S5	Protein synthesis		-7.7	-69.6		
EF0225	1	<i>rpmD</i>	ribosomal protein L30	Protein synthesis					6.7
EF0226	1	<i>rplO</i>	ribosomal protein L15	Protein synthesis				-5.7	
EF0227	1	<i>secY</i>	preprotein translocase, SecY subunit	Protein fate		-8.7	-23.7	-17.2	-12.3
EF0229	1	<i>infA</i>	translation initiation factor IF-1	Protein synthesis		-13.6	-6.4		-5.1
EF0230	1	<i>rpmJ</i>	ribosomal protein L36	Protein synthesis					5.2
EF0231	1	<i>rpsM</i>	ribosomal protein S13	Protein synthesis		-5.1	-84.5	-6.5	16.8
EF0232	1	<i>rpsK</i>	ribosomal protein S11	Protein synthesis		-5.6	-12.3	-8.7	-9.7
EF0234	1	<i>rplQ</i>	ribosomal protein L17	Protein synthesis		-5.8	-15.4	-18.9	-6.0
EF0235	1		membrane protein, putative	cell envelope		-7.3			-9.3
EF0245	1		decarboxylase family protein	central and intermediary metabolism		5.1			16.5
EF0246	2		amino acid ABC transporter, ATP-binding protein	Transport			-11.3		-6.3
EF0251	1		transcriptional regulator, putative	Regulatory functions		-5.9	-13.6		-6.1
EF0262	1		S1 RNA-binding domain protein	Protein synthesis			-9.0	-12.7	
EF0271	1	<i>arb</i>	glycosyl hydrolase, family 1	central and intermediary metabolism		44.8			
EF0272	1		glycosyl hydrolase, family 1	central and intermediary metabolism		18.5	19.3	8.8	
EF0277	1	<i>ogt</i>	methylated-DNA--protein-cysteine S-methyltransferase	DNA metabolism					14.5
EF0279	1		HD domain protein	Signal transduction					16.3
EF0282	1	<i>fabI</i>	enoyl-(acyl-carrier-protein) reductase	Fatty acids and phospholipids metabolism			-7.4	-17.7	
EF0283- EF0284	2	<i>fabF-1-fabZ-1</i>	3-oxoacyl-(acyl-carrier-protein) synthase II, (3R)-hydroxymyristoyl-(acyl-carrier-protein) dehydratase	Fatty acids and phospholipids metabolism			-9.1	-14.1	-10.0
EF0285	1	<i>pyrD-1</i>	dihydroorotate dehydrogenase	Purines, Pyrimidines, nucleosides, nucleotides		-12.7			
EF0289- EF0290	2	<i>metCcsK</i>	cystathionine gamma-lyase-cysteine synthase	Amino-acid biosynthesis		-15.5			-27.1
EF0291	1	<i>celA</i>	glycosyl hydrolase, family 1	central and intermediary metabolism		-10.7		9.8	-10.7
EF0297	1	<i>copY</i>	transcriptional repressor CopY	Regulatory functions			34.2		11.6
EF0298- EF0299	2		copper-translocating P-type ATPase, copper transport protein CopZ	Transport			34.2		11.6
EF0300	1		membrane protein, putative	cell envelope			12.9		
EF0301	1		transcriptional regulator, GntR family	Regulatory functions		17.6			
EF0303	1		phage integrase (pp1)	Mobile and extrachromosomal elements function					-6.1
EF0313	1		hypothetical protein (pp1)	Mobile and extrachromosomal elements function				9.8	

EF0319	1		hypothetical protein (pp1)	Mobile and extrachromosomal elements function		7.8			
EF0324	1		hypothetical protein (pp1)	Mobile and extrachromosomal elements function	-8.7				-12.4
EF0332	1		hypothetical protein (pp1)	Mobile and extrachromosomal elements function	7.4		9.8	6.8	
EF0334	1		portal protein (pp1)	Mobile and extrachromosomal elements function		5.6			
EF0344	1		hypothetical protein (pp1)	Mobile and extrachromosomal elements function			19.3		
EF0349	1		tail protein, putative (pp1)	Mobile and extrachromosomal elements function	8.1	18.3		27.6	
EF0351	1		structural protein, putative (pp1)	Mobile and extrachromosomal elements function				5.3	
EF0358	1		glyoxalase family protein	cellular process		11.6		9.0	
EF0363	1		ISEf1, transposase	Mobile elements function			-18.3	-6.5	
EF0366	1			Hypothetical or unknown protein					-5.2
EF0367	1			Hypothetical or unknown protein			-7.8		
EF0373- EF0372	2	<i>ehk13-err13</i>	two-component system	Signal transduction		6.0		-12.2	
EF0374	1		lipoprotein	cell envelope		9.1	19.3		
EF0381	1			Hypothetical or unknown protein		24.0			
EF0383	1		protein FdrA	Protein synthesis		49.2	35.1		
EF0387	1		sodium/dicarboxylate symporter family protein	Transport		11.6	9.8		
EF0389	1		membrane protein	cell envelope				10.2	
EF0390	1		N-acyl-D-amino-acid deacylase family protein	Amino-acid biosynthesis			19.3		
EF0394	1	<i>salB</i>	secreted antigen	cellular process	16.6	-40.7	-84.9	-13.0	-15.9
EF0396	1			Hypothetical or unknown protein		-6.4		13.2	
EF0402	1	<i>nhaC-1</i>	Na <sup>+</sup> /H <sup>+</sup> antiporter	Transport				-7.2	
EF0403	1	<i>llrb</i>	transcriptional regulator, MarR family	Regulatory functions	5.8				
EF0405	1		hydrolase, haloacid dehalogenase-like family	central and intermediary metabolism		97.0	18.6	10.8	-5.8
EF0411- EF0412	2	<i>EF0411-mltF</i>	PTS system, IIBCA components mannitol specific	Transport			35.1		
EF0413	1	<i>mtlD</i>	mannitol-1-phosphate 5-dehydrogenase	central and intermediary metabolism			35.1		
EF0414	1		oxidoreductase, DadA family	central and intermediary metabolism	-11.3			5.3	
EF0418	1		ABC transporter, ATP-binding protein	Transport					-5.5
EF0420	1		drug resistance transporter, EmrB/QacA family protein	Transport	15.6				
EF0421	1		transcriptional regulator, MerR family	Regulatory functions	13.6				
EF0432	1		transcriptional regulator, AraC family	Regulatory functions				-8.1	-21.2
EF0440	1		di-/tripeptide transporter	Transport		-15.7			
EF0443	1		LysM domain protein	cellular process	-6.1	-15.8	-162.4	-15.1	
EF0445- EF0450	6	<i>menB- EF0446- EF0447- menD- EF0449</i>	short chain dehydrogenase, O-succinylbenzoic acid--CoA ligase, menaquinone-specific isochorismate synthase, 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase, 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase, mandelate racemase	central and intermediary metabolism		-11.9			
EF0453	1	<i>ohr gsp65</i>	OsmC/Ohr family protein	cellular process				17.8	6.3
EF0465	1		transcriptional regulator	Regulatory functions			-5.3		
EF0466	1	<i>nagB</i>	glucosamine-6-phosphate isomerase	central and intermediary metabolism		20.3			
EF0467	1			Hypothetical or unknown protein		12.5			
EF0470- EF0473	1	<i>nrdHIEF</i>	ribonucleoside-diphosphate reductase	Purines, Pyrimidines, nucleosides, nucleotides	-12.7	-5.1			-8.5
EF0475- EF0476	2	<i>feoAB</i>	ferrous iron transport proteins AB	Transport		9.9	5.4		-7.9
EF0485	1		aggregation substance	cellular process	8.3	7.0			
EF0492	1			Hypothetical or unknown protein				6.7	
EF0500	1			Hypothetical or unknown protein		12.6			
EF0504	1			Hypothetical or unknown protein					10.3
EF0514	1		IS1216, transposase	Mobile elements function				-13.8	
EF0520	1		IS1216, transposase, degenerate	Mobile elements function		-5.5			
EF0522	1		IS1216, transposase, authentic point mutation	Mobile elements function				-15.6	

EF0525- EF0527	3	<i>cyll-Lcyll- ScylM</i>	cytolysin	cellular process					-10.2
EF0529	1		IS256, transposase	Mobile elements function		-40.8	-54.9		
EF0547	1		ammonium transporter family protein	Transport		18.5	10.6	8.8	
EF0551	1		glycosyl hydrolase, family 31	central and intermediary metabolism		12.2			
EF0552- EF0555	4		PTS system IICBDA components	Transport		58.5			
EF0556- EF0557	2	<i>xylAB</i>	xylose isomerase - kinase	central and intermediary metabolism		13.7		5.8	
EF0563	1			Hypothetical or unknown protein		10.8			
EF0566	1	<i>kdpF</i>	potassium transporter stabilization peptide KdpF	Transport					-11.0
EF0567- EF0570	4	<i>kdpABCD</i>	potassium-transporting ATPase	Transport	6.9			9.6	17.5
EF0575- EF0576	2		cationic ABC transporter	Transport		5.3			
EF0577	1		adhesion lipoprotein	cell envelope	5.5	12.4			
EF0581- EF0584	1		ABC transporter	Transport					-6.1
EF0595	1		ISEf1, transposase	Mobile elements function			-90.6		-6.4
EF0600	1		transcriptional regulator, TetR family	Regulatory functions		7.2			-5.6
EF0602	1		transposase, IS3 family	Mobile elements function		-6.9			-6.4
EF0604	1		gls24 protein	cellular process				-10.6	-8.1
EF0606	1		Dps family protein	cellular process			5.8		
EF0607	1		ParB-like nuclease domain protein	DNA metabolism		13.7	7.2		
EF0608	1			Hypothetical or unknown protein		59.5	7.2	5.2	
EF0609	1			Hypothetical or unknown protein				18.8	
EF0622	1			Hypothetical or unknown protein				14.5	
EF0629	1		oxidoreductase, aldo/keto reductase family	central and intermediary metabolism	8.5		19.3		16.4
EF0633	1	<i>tryS-1</i>	tyrosyl-tRNA synthetase	Protein synthesis		6.8			5.0
EF0634	1		decarboxylase, putative	central and intermediary metabolism	7.8			40.6	12.9
EF0635	1		amino acid permease family protein	Transport		-6.9	-5.0	-12.6	-7.9
EF0636	1	<i>nhaC-2</i>	Na <sup>+</sup> /H <sup>+</sup> antiporter	Transport		5.7			-5.4
EF0641	1	<i>ldh-2</i>	L-lactate dehydrogenase	central and intermediary metabolism		-6.0		9.8	-7.3
EF0655	1		nitroreductase family protein	central and intermediary metabolism					11.3
EF0657	1		transcriptional regulator, DeoR family	Regulatory functions			9.8		
EF0660	1		MATE efflux family protein	Transport	21.1			14.2	20.8
EF0666	1		glyoxalase family protein	cellular process	14.9			16.1	
EF0669	1		polysaccharide biosynthesis family protein	central and intermediary metabolism		-8.6	-10.2		
EF0671	1		xaa-his dipeptidase	Protein fate					5.8
EF0673	1		membrane protein, putative	cell envelope	6.2				
EF0674- EF0675	2		glycine betaine/carnitine/choline ABC transporter	Transport	9.4			10.5	8.4
EF0676	1	<i>argR</i>	arginine repressor	Regulatory functions		-7.3			
EF0677	1		phosphoglucomutase/phosphomannomutase family protein	central and intermediary metabolism		6.3	7.7		
EF0688	1		ABC transporter, ATP-binding protein	Transport		-22.5	6.0		
EF0693- EF0696	2	<i>fruK-1- EF0694- EF0695- lacD-1</i>	1-phosphofruktokinase, tagatose 1,6-diphosphate aldolase	central and intermediary metabolism			17.8	37.3	10.5
EF0694- EF0695	2		PTS system, IIBCA components	Transport			17.8	37.3	10.5
EF0696	1			Hypothetical or unknown protein		-9.6			-6.4
EF0698	1		acetyltransferase, GNAT family	central and intermediary metabolism	6.6				
EF0700	1		hemolysin	cellular process		-20.3	-132.3	-6.5	
EF0701	1	<i>prfC</i>	peptide chain release factor 3	Protein synthesis		-28.1	-7.8		
EF0704	1		lipoprotein, putative	cell envelope			8.9		7.2
EF0706	1	<i>clpE</i>	ATP-dependent Clp protease, ATP-binding subunit ClpE	Protein fate		16.5			
EF0709- EF0710	2	<i>ptsH-ptsI</i>	phosphocarrier protein HPr-PEP phosphotransferase EI	Signal transduction			-12.8		43.3
EF0717	1		PTS system, fructose-specific IIBC components	Transport			-8.8	-11.1	-108.1



EF0718	1	<i>fruK-2</i>	1-phosphofruktokinase	central and intermediary metabolism						-69.6
EF0719	1		transcriptional regulator, DeoR family	Regulatory functions	5.1					-32.1
EF0727	1		diacylglycerol kinase catalytic domain protein	Fatty acids and phospholipids metabolism	-12.4	-12.3	-80.4			-9.0
EF0731	1		transcriptional regulator, luxR family	Regulatory functions		5.8				
EF0740	1		deoxynucleoside kinase	Purines, Pyrimidines, nucleosides, nucleotides		7.4	35.1			
EF0741	1			Hypothetical or unknown protein	-7.6					-7.9
EF0744	1		sodium/dicarboxylate symporter family protein	Transport		-13.6				
EF0746	1	<i>pbp5</i>	penicillin-binding protein 5	cell envelope	11.6	5.6			11.3	
EF0747	1			Hypothetical or unknown protein					41.3	
EF0756	1			Hypothetical or unknown protein					60.2	
EF0758	1		cadmium-translocating P-type ATPase	Transport	8.5				11.6	
EF0761- EF0760	2		amino acid ABC transporter	Transport	-5.5	-38.7	-10.7	-63.5		-5.0
EF0764	1			Hypothetical or unknown protein	-7.4					
EF0765	1			Hypothetical or unknown protein	-9.4					-7.5
EF0768	1			Hypothetical or unknown protein		9.4	17.8			8.5
EF0769	1			Hypothetical or unknown protein					5.8	
EF0771	1	<i>clpP</i>	ATP-dependent Clp protease, proteolytic subunit ClpP	Protein fate		12.5	6.2			-10.3
EF0773	1			Hypothetical or unknown protein					23.9	
EF0781	1		cold shock domain family protein	cellular process		-16.7				
EF0789- EF0790	2		ABC transporter	Transport	-5.3				-12.4	
EF0794	1			Hypothetical or unknown protein					-6.3	
EF0797	1			Hypothetical or unknown protein		5.2				
EF0806- EF0804	3		amino acid ABC transporter	Transport					8.3	
EF0813	1		glycosyl hydrolase, family 35	central and intermediary metabolism	7.9	27.2				9.3
EF0815- EF0816	3		PTS system, IIABCD components	Transport		13.7				
EF0817	1			Hypothetical or unknown protein		6.0				
EF0820	1	<i>rply</i>	ribosomal protein L25	Protein synthesis		6.6			7.8	
EF0821	1		ribosomal small subunit pseudouridine synthase A, putative	Protein synthesis					9.0	
EF0825	1	<i>udk</i>	uridine kinase	Purines, Pyrimidines, nucleosides, nucleotides	-18.6	-14.2				-10.1
EF0827	1		oxidoreductase, Gfo/Idh/MocA family	central and intermediary metabolism	-5.4				5.8	
EF0845	1	<i>murF</i>	UDP-N-acetylmuramoylalanyl-D-glutamyl-diaminopimelate-D-alanyl-ligase	cell envelope					-5.1	
EF0846	1		ATP-dependent RNA helicase, DEAD/DEAH box family	Regulatory functions				-6.6		
EF0848- EF0849	2	<i>acpS-<i>alr</i></i>	holo-ACP-synthase, alanine racemase	cell envelope				-65.5		
EF0850	1		transcriptional regulator, PemK family	Regulatory functions		-12.4				
EF0865- EF0862	4		glycine betaine/carnitine/choline ABC transporter	Transport				-8.9		
EF0875	1		copper-translocating P-type ATPase	Transport					9.6	
EF0882	1		replication initiation and membrane attachment protein DnaB, putative	DNA metabolism		5.3			7.1	6.4
EF0888	1			Hypothetical or unknown protein						-8.6
EF0891	1		aspartate aminotransferase, putative	Amino-acid biosynthesis	5.3					5.3
EF0897	1		preprotein translocase, YajC subunit	Protein fate	-10.8	-9.3			-6.6	
EF0900	1	<i>adhE</i>	aldehyde-alcohol dehydrogenase	central and intermediary metabolism		23.7	8.3		-5.1	
EF0901- EF0904	4		isopentenyl diphosphate delta isomerase	cell envelope		-5.8	-15.1		-7.0	
EF0909- EF0912	4		peptide ABC transporter	Transport		-12.9				
EF0916	1	<i>rplT</i>	ribosomal protein L20	Protein synthesis		-9.3	-8.4		-16.0	
EF0919	1		acetyltransferase, GNAT family	central and intermediary metabolism		7.8			5.1	
EF0926- EF0927	2	<i>err09-<i>ehk09</i></i>	two-component system	Signal transduction		14.3	5.1			
EF0928	1		glucose uptake protein	Transport	-8.2	-10.7	-15.0		-14.5	
EF0929	1		amino acid permease family protein	Transport	-5.8	-17.6			-8.0	



EF0931	1			Hypothetical or unknown protein		7.0			
EF0934	1		hydrolase, TatD family	Purines, Pyrimidines, nucleosides, nucleotides		14.8			13.6
EF0938	1		ABC transporter, ATP-binding/TOBE domain protein	Transport		83.5	6.8		
EF0939	1	<i>mgsA</i>	methylglyoxal synthase	central and intermediary metabolism	9.4	11.6		13.5	7.9
EF0944	1		extracellular protein	cellular process		-7.4	-103.8	-5.1	
EF0948	1	<i>ung</i>	uracil-DNA glycosylase	DNA metabolism				5.6	9.2
EF0954	1	<i>malR</i>	MalR regulator	Regulatory functions		-5.8			-6.6
EF0957- EF0955	3	<i>malPpgcMm alM</i>	maltose phosphorylase, beta-phosphoglucomutase, mutarotase	central and intermediary metabolism	16.2	-5.8			-6.6
EF0958- EF0960	2		PTS system, IIABC components	Transport	17.6		50.9	48.8	
EF0961	1		pyrroline-5-carboxylate reductase	Amino-acid biosynthesis	7.9			50.6	8.3
EF0964	1			Hypothetical or unknown protein				14.1	
EF0965	1			Hypothetical or unknown protein		9.2			
EF0970	1	<i>rpmA</i>	ribosomal protein L27	Protein synthesis	-39.7	-104.8	-70.2	-44.1	-44.4
EF0972	1		DNA repair exonuclease family protein	DNA metabolism	8.7				
EF0977	1	<i>nusB</i>	N utilization substance protein B	Regulatory functions	-5.6	-7.7	-7.4		
EF0982	1	<i>tlyA</i>	hemolysin A	cellular process	-23.9	-13.7	-10.1	-18.0	-58.3
EF0987	1		lipoprotein	cell envelope	-5.6	-24.4		-7.5	
EF0988	1			Hypothetical or unknown protein		-15.0	-172.1		
EF0992- EF0994	3	<i>mraY-murD- murG</i>	phospho-N-acetylmuramoyl-pentapeptide-transferase, UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase, UDPdiphospho-muramoylpentapeptide beta-N- acetylglucosaminyltransferase	cell envelope		-63.2	-11.0	-5.4	
EF0996- EF0997	2	<i>ftsAZ</i>	cell division proteins FtsAZ	cellular process		-19.9			
EF0999	1			Hypothetical or unknown protein		-5.6			
EF1003	1	<i>ileS</i>	isoleucyl-tRNA synthetase	Protein synthesis		-6.7			
EF1008	1		oxidoreductase, Gfo/ldh/MocA family	central and intermediary metabolism		15.2		7.6	5.3
EF1017- EF1019	3		PTS system, IIABC components	Transport	7.3				8.3
EF1020	1			Hypothetical or unknown protein		18.6			
EF1021	1			Hypothetical or unknown protein				6.8	
EF1022	1			Hypothetical or unknown protein		-7.0			
EF1024	1			Hypothetical or unknown protein				99.7	
EF1035	1		lipoprotein	cell envelope			19.3		
EF1036	1		nucleoside diphosphate kinase	Purines, Pyrimidines, nucleosides, nucleotides		18.5			
EF1038	1		lipoprotein	cell envelope				31.6	
EF1041	1		xanthine/uracil permeases family protein	Transport		8.6	12.6		
EF1045	1	<i>pfk</i>	6-phosphofructokinase	central and intermediary metabolism					7.0
EF1046	1	<i>pyk</i>	pyruvate kinase	central and intermediary metabolism		-6.4	-13.7		
EF1048	1	<i>rpmF-3</i>	ribosomal protein L32	Protein synthesis	-8.3	-9.8			
EF1050- EF1051	2	<i>etaRS</i>	two-component system	Signal transduction	-6.7	6.3	-66.0	20.8	-28.2
EF1057	1		Mn2+/Fe2+ transporter, NRAMP family	Transport		28.1			9.6
EF1058	1		universal stress protein family	cellular process		9.6			-7.1
EF1064	1			Hypothetical or unknown protein	12.2				
EF1065	1			Hypothetical or unknown protein	8.0	18.2	42.7	110.0	
EF1068	1	<i>galM</i>	aldose 1-epimerase	central and intermediary metabolism		18.5			
EF1069- EF1071	3	<i>galK-galE-1- galT</i>	galactokinase, UDP-glucose 4-epimerase, galactose-1-phosphate uridylyltransferase, galactose-1-phosphate uridylyltransferase	central and intermediary metabolism			8.9		
EF1072	1	<i>galR</i>	galactose operon repressor galR	Regulatory functions					-5.1
EF1078	1	<i>emeA</i>	multidrug resistance protein	Transport			-6.5	8.8	
EF1086	1		spermine/spermidine acetyltransferase	Amino-acid biosynthesis	-9.2				
EF1089	1			Hypothetical or unknown protein	-7.3				
EF1096	1			Hypothetical or unknown protein				-5.4	5.1
EF1097	1			Hypothetical or unknown protein	-5.4				
EF1098	1			Hypothetical or unknown protein	-38.0	-8.3		-8.7	-35.3
EF1104	1			Hypothetical or unknown protein	19.4			25.4	16.8

EF1108- EF1110	4	<i>lutABC</i>	oxidoreductase, iron sulfur cluster binding protein, YkgG family protein	central and intermediary metabolism					8.1	
EF1112	2	<i>rexB</i>	exonuclease RexB	DNA metabolism						6.5
EF1120- EF1117	4	<i>glnQHMP</i>	amino acid ABC transporter	Transport	-19.7					6.0
EF1132	1		Cystathionine beta-synthase domain protein	Signal transduction					-11.4	
EF1143	1		HD domain protein	Signal transduction					8.2	
EF1147	1	<i>pyrG</i>	CTP synthase	Purines, Pyrimidines, nucleosides, nucleotides				-37.8	-5.8	
EF1151	1		cell division protein DivIVA	cellular process						5.4
EF1155	1	<i>nth</i>	endonuclease III	DNA metabolism					-9.0	
EF1162	1		helicase, putative	DNA metabolism	-7.0				-5.1	
EF1166	1		YitT family protein	cell envelope					9.1	
EF1167	1	<i>fba</i>	fructose-bisphosphate aldolase class-II	central and intermediary metabolism					-10.1	
EF1169	1	<i>murAB</i>	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 2	cell envelope				-9.5		
EF1171	1	<i>rpmE</i>	ribosomal protein L31	Protein synthesis	-15.6	-30.5	-64.0	-9.1		-6.3
EF1179	1	<i>cscK</i>	fructokinase	central and intermediary metabolism	15.0	7.7		58.7		15.4
EF1181	1		nitroreductase family protein	Other metabolism					8.6	
EF1184- EF1185	2	<i>dapA-rnjB</i>	4-hydroxy-tetrahydrodipicolinate synthase, metallo-beta-lactamase superfamily protein	Purines, Pyrimidines, nucleosides, nucleotides				-28.2	-33.5	-7.6
EF1191	1		DegV family protein	Fatty acids and phospholipids metabolism					-9.4	
EF1193- EF1194	2	<i>vicRK (err07-ehk07)</i>	two-component system	Signal transduction				-5.3	-14.1	8.3
EF1211	1	<i>npr</i>	NADH peroxidase	cellular process						10.2
EF1213	1	<i>alsS</i>	acetolactate synthase	central and intermediary metabolism				7.4		
EF1214	1	<i>budA</i>	alpha-acetolactate decarboxylase	central and intermediary metabolism	-6.8					-44.6
EF1224	1		transcriptional regulator, Cro/Ci family	Regulatory functions				40.5		
EF1226	1		oxidoreductase, putative	central and intermediary metabolism				6.2		
EF1227	1			Hypothetical or unknown protein				49.2		10.2
EF1228	1			Hypothetical or unknown protein	-6.7					
EF1232- EF1234	3		ABC transporter	Transport				16.1		
EF1240	1		sugar-binding transcriptional regulator, LacI family	Regulatory functions	-22.3	-17.3			-19.4	-21.1
EF1242	1		transcriptional regulator, GntR family	Regulatory functions	-11.5				-9.5	
EF1246	1			Hypothetical or unknown protein						8.3
EF1260- EF1261	2	<i>err06-ehk06</i>	two-component system	Signal transduction	11.0					6.8
EF1263	1			Hypothetical or unknown protein	38.8	22.7		16.8		61.3
EF1266	1			Hypothetical or unknown protein						10.9
EF1273	1		ribosomal protein L7A family	Protein synthesis						-6.0
EF1274	1	<i>infB</i>	translation initiation factor IF-2	Protein synthesis					-6.2	
EF1277	1		transcriptional regulator, Cro/Ci family (pp2)	Mobile and extrachromosomal elements function					6.1	
EF1288	1		hypothetical protein (pp2)	Mobile and extrachromosomal elements function				7.4		
EF1300- EF1301	2		cell division proteins, FtsW/RodA/SpovE family	cellular process				66.7	19.3	
EF1302	1		transcriptional regulator, putative	Regulatory functions				5.7		
EF1303	1		transcriptional regulator, LysR family	Regulatory functions	6.1	22.7	9.8			
EF1304	1		magnesium-translocating P-type ATPase	Transport	7.2		7.2			
EF1306- EF1310	4	<i>hrcAgrpEdna KdnaJ</i>	heat shock proteins	Protein fate	10.0	12.3	5.9	8.7		5.6
EF1311	1			Hypothetical or unknown protein				-12.2		-14.9
EF1317	1	<i>nagA-1</i>	N-acetylglucosamine-6-phosphate deacetylase	central and intermediary metabolism	5.3				-5.5	
EF1326	1		transcriptional regulator, TetR family	Regulatory functions				-5.3	-21.0	-7.6
EF1328	1		transcriptional regulator, GntR family	Regulatory functions				5.3		
EF1335- EF1336	2	<i>err16-ehk16</i>	two-component system	Signal transduction	-43.9				-5.2	-327.2
EF1338	1			Hypothetical or unknown protein						-14.0
EF1342	1		transcriptional regulator, Mar family	Regulatory functions						-7.1

EF1345- EF1343	3	<i>mdxEFG</i>	sugar ABC transporter	Transport	16.2	10.1			
EF1347- EF1349	3	<i>mmdHgmDH mmgT</i>	maltodextrin hydrolase, maltooligosaccharide glycosyl transferase	central and intermediary metabolism	22.1	9.3	19.3		
EF1352	1		magnesium-translocating P-type ATPase	Transport			5.8		
EF1358- EF1361	4	<i>gldAdhaMKL</i>	dihydroxyacetone kinase	central and intermediary metabolism	5.3	53.2	26.6		6.4
EF1362	1			Hypothetical or unknown protein		7.9		7.0	
EF1368	1			Hypothetical or unknown protein					8.5
EF1371	1			Hypothetical or unknown protein	7.1				
EF1379	1	<i>alaS</i>	alanyl-tRNA synthetase	Protein synthesis		-40.0			
EF1397- EF1399	3		molybdenum ABC transporter	Transport				-11.7	
EF1400	1		cadmium-translocating P-type ATPase	Transport		12.6	21.1		
EF1412	1			Hypothetical or unknown protein					-5.8
EF1419	1		hypothetical protein (pp3)	Mobile and extrachromosomal elements function			-83.4		
EF1420	1		hypothetical protein (pp3)	Mobile and extrachromosomal elements function		5.9			
EF1428	1		hypothetical protein (pp3)	Mobile and extrachromosomal elements function				24.2	
EF1449	1		hypothetical protein (pp3)	Mobile and extrachromosomal elements function	-17.5	-24.5	-21.4	-20.9	-114.7
EF1456	1		hypothetical protein (pp3)	Mobile and extrachromosomal elements function			19.3		
EF1457	1		minor head protein (pp3)	Mobile and extrachromosomal elements function		-9.2			-71.9
EF1458	1		hypothetical protein (pp3)	Mobile and extrachromosomal elements function	24.5			8.8	
EF1479	1		hypothetical protein (pp3)	Mobile and extrachromosomal elements function					-8.6
EF1480	1		hypothetical protein (pp3)	Mobile and extrachromosomal elements function			22.6	8.1	
EF1492- EF1500	9	<i>ntpFIKECGA BD</i>	V-type ATPase	Transport	7.6	-11.6		-5.1	6.0
EF1504	1	<i>lysA</i>	diaminopimelate decarboxylase	Amino-acid biosynthesis	-9.0	-5.5		-62.7	-10.3
EF1510	1			Hypothetical or unknown protein				13.1	
EF1512	1			Hypothetical or unknown protein		-11.6			-8.2
EF1519	1		cation-transporting ATPase, E1-E2 family	Transport			-6.4		
EF1521	1	<i>dnaG</i>	DNA primase	DNA metabolism			-9.2		
EF1522	1	<i>sigA</i>	RNA polymerase sigma-43 factor	Regulatory functions				-15.2	
EF1526	1	<i>gap-1</i>	glyceraldehyde 3-phosphate dehydrogenase	central and intermediary metabolism		5.5		6.9	
EF1527	1	<i>obgE</i>	GTP-binding protein	Purines, Pyrimidines, nucleosides, nucleotides		-6.6			
EF1534	1		peptidyl-prolyl cis-trans isomerase, cyclophilin-type	Protein synthesis		23.9		-5.4	
EF1535	1			Hypothetical or unknown protein					-5.2
EF1543	1	<i>fer</i>	ferredoxin	cellular process	8.7			34.8	
EF1545	1	<i>recQ-1</i>	ATP-dependent DNA helicase RecQ	DNA metabolism	6.1			57.8	
EF1547	1	<i>cmk</i>	cytidylate kinase	Amino-acid biosynthesis					7.9
EF1548	1	<i>rpsA</i>	ribosomal protein S1	Protein synthesis			-7.5		
EF1550	1	<i>hup</i>	DNA-binding protein HU	DNA metabolism		8.4	-69.0	8.5	
EF1553	1		TPR domain protein	Protein synthesis		8.0			5.6
EF1555	1		YitT family protein	cell envelope	9.2			8.1	5.3
EF1558	1	<i>papS</i>	poly A polymerase	Purines, Pyrimidines, nucleosides, nucleotides	6.5	9.2	8.7	67.6	17.3
EF1561	7	<i>aroE- EF1562- aroBC- EF1565- aroAK- EF1568</i>	shikimate 5-dehydrogenase, 3-deoxy-7- phosphoheptulonate synthase, 3- dehydroquininate synthase, chorismate synthase, chorismate synthase, 3- phosphoshikimate 1- carboxyvinyltransferase, shikimate kinase, prephenate dehydratase	Amino-acid biosynthesis					6.1
EF1570	1		DegV family protein	Fatty acids and phospholipids metabolism	-7.3	-9.4		-5.9	-11.5
EF1589	1		acetyltransferase, GNAT family	central and intermediary metabolism		-26.9			-5.2
EF1598	1	<i>phrB</i>	deoxyribodipyrimidine photolyase	Purines, Pyrimidines, nucleosides, nucleotides				7.1	
EF1601	1		PTS system, IIABC components	Transport		9.4	9.8		

EF1602	1		glycosyl hydrolase, family 13	central and intermediary metabolism		12.3			
EF1603- EF1604	2	<i>scrB-1- scrR-1</i>	sucrose-6-phosphate dehydrogenase	central and intermediary metabolism		9.4		7.6	
EF1612- EF1613	2	<i>pflAB</i>	pyruvate formate-lyase activating enzyme	central and intermediary metabolism	-5.2			-5.9	
EF1641- EF1639	3		iron compound ABC transporter	Transport		10.8	17.3		
EF1647	1	<i>hslV</i>	heat shock protein HslV	cellular process					5.5
EF1649	1	<i>gid</i>	glucose-inhibited division protein	Signal transduction	-6.3	-25.4	-7.1		-6.1
EF1655	1		2-dehydropantoate 2-reductase	central and intermediary metabolism	-64.3				-69.8
EF1656	1		transcriptional regulator, LysR family	Regulatory functions		132.4	50.9		
EF1658- EF1663	5	<i>ptbbukbkD ABC</i>	branched-chain alpha-keto acid, E2, acetyltransferase	Amino-acid biosynthesis	-20.9		19.3	8.8	-7.2
EF1662	1			Hypothetical or unknown protein			9.8		
EF1671	1		oxidoreductase, zinc-binding	central and intermediary metabolism		9.7		5.7	
EF1672	1		permease protein, putative	Transport		22.7	28.0		
EF1678	1		signal peptidase I	Protein fate	-6.2				-11.8
EF1680	1			Hypothetical or unknown protein			9.8		
EF1685	1	<i>hlyIII</i>	hemolysin III	cellular process	-7.1	-34.7	-96.1	-8.2	-16.8
EF1689	1			Hypothetical or unknown protein					7.2
EF1693	1		KH domain protein	Signal transduction	-9.3			-5.1	
EF1694	1	<i>rpsP</i>	ribosomal protein S16	Protein synthesis	-7.8	-7.9	-71.8	-13.8	5.3
EF1707	1		glycosyl hydrolase, family 38	central and intermediary metabolism		36.1			
EF1710	1		transcriptional regulator, LysR family	Regulatory functions		-11.8			-5.9
EF1719- EF1712	8	<i>pyrBCpyrAc arbBpyrDIID -2FE</i>	dihydroorotase-carbamoyl phosphate synthase, dihydroorotate dehydrogenase, OMP decarboxylase; OMPDCase; OMPdecase, orotate phosphoribosyltransferase	Purines, Pyrimidines, nucleosides, nucleotides	-6.2	-10.4		-15.5	
EF1720	1		uracil permease	Transport	-6.2	-10.4		-15.5	
EF1721	1	<i>pyrR</i>	pyrimidine regulatory protein PyrR	Regulatory functions	-6.2	-10.4		-7.1	
EF1724	1		Cystathionine beta synthase domain protein	Signal transduction				7.6	12.1
EF1733- EF1732	1		ABC transporter	Transport					-10.2
EF1736	1	<i>nfo</i>	endonuclease IV	DNA metabolism		-10.7		5.6	
EF1739	1	<i>tryS-2</i>	tyrosyl-tRNA synthetase	Protein synthesis	-5.4		-13.1		
EF1740	1	<i>pbp1B</i>	penicillin-binding protein 1B	cell envelope		-5.7			
EF1744	1		general stress protein	cellular process	6.6			5.5	
EF1751	1		membrane protein, putative	cell envelope	-6.2	60.7	21.1	-6.0	-11.1
EF1754	1		PhoU family protein	Signal transduction				-5.1	
EF1760	1		cell division ABC transporter	Transport		-6.2		-34.8	
EF1764	1	<i>yfiA</i>	ribosomal subunit interface protein	Protein synthesis		6.5			
EF1769	1		PTS system, IIB component	Transport	6.6			9.1	-5.7
EF1770	1			Hypothetical or unknown protein		-6.1			-7.6
EF1774	1			Hypothetical or unknown protein				11.6	
EF1787- EF1777	11	<i>purE-K1- CSQLFMNH D</i>	5-(carboxyamino)imidazole ribonucleotide mutase and synthase, phosphoribosylaminoimidazole-succinocarboxamide synthase, phosphoribosylformylglycinamide synthase, phosphoribosylformylglycinamide synthetase I, phosphoribosylformylglycinamide synthase II, amidophosphoribosyltransferase, phosphoribosylaminoimidazole synthetase, phosphoribosylglycinamide formyltransferase, phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase, phosphoribosylamine-glycine ligase	Purines, Pyrimidines, nucleosides, nucleotides			82.6		
EF1789	1		SPFH domain/Band 7 family protein	cell envelope	6.3			24.2	
EF1804- EF1802	4		PTS system, IIBCD components	Transport		22.7			



EF2036	1		hypothetical protein (pp4)	Mobile and extrachromosomal elements function	-8.7				
EF2040	1		transcriptional regulator, Cro/Ci family (pp4)	Mobile and extrachromosomal elements function					13.4
EF2041			hypothetical protein (pp4)	Mobile and extrachromosomal elements function				10.5	
EF2051	1		transcriptional regulator, GntR family	Regulatory functions	-6.4	-14.6	-9.1		-25.7
EF2055	1		oxidoreductase, pyridine nucleotide-disulfide family	Purines, Pyrimidines, nucleosides, nucleotides		-8.9			
EF2056	1		1,4-dihydroxy-2-naphthoate octaprenyltransferase	Other metabolism		-9.0		7.8	
EF2057	1		heptaprenyl diphosphate synthase	Fatty acids and phospholipids metabolism				19.4	
EF2061-EF2060	2	<i>cydAB</i>	cytochrome d ubiquinol oxidase	central and intermediary metabolism	-5.2		-90.0	-5.2	-7.9
EF2063	1	<i>soxS</i>	transcriptional regulator, AraC family	Regulatory functions				-8.4	-9.4
EF2067	1			Hypothetical or unknown protein		12.6		8.0	
EF2074-EF2075	2		ABC transporter	Transport			-8.1	-10.1	
EF2076	1	<i>efaA</i>	endocarditis specific antigen	cellular process	5.5		5.8	14.3	
EF2078	1			Hypothetical or unknown protein	-5.7	-19.2	-42.4		-228.4
EF2085	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function					-5.5
EF2086	1		endolysin (pp5)	Mobile and extrachromosomal elements function				5.2	
EF2087	1		holin (pp5)	Mobile and extrachromosomal elements function			35.1		
EF2090	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function		10.3			
EF2093	1		endolysin domain protein (pp5)	Mobile and extrachromosomal elements function		6.1			
EF2094	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function			9.8		
EF2095	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function		11.6		12.8	
EF2096	1		tail protein (pp5)	Mobile and extrachromosomal elements function		7.2	9.7	5.7	
EF2097	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function		6.2			
EF2098	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function		9.4			
EF2099	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function		11.5			
EF2103	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function		10.8			
EF2104	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function			13.0		
EF2105	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function		7.8			
EF2109	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function		22.9	19.3		
EF2112	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function		9.3			
EF2113	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function			9.8		
EF2115	1		transcriptional regulator, ArpU family (pp5)	Mobile and extrachromosomal elements function	-33.1	-22.0	-25.7	-6.1	-35.8
EF2118	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function		22.9	35.1		
EF2120	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function		18.5	19.3		
EF2128	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function		18.5			
EF2131	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function		31.6			
EF2132	1		recT protein, putative (pp5)	Mobile and extrachromosomal elements function		13.7			
EF2138	1		transcriptional regulator, Cro/Ci family (pp5)	Mobile and extrachromosomal elements function		31.6	19.4		
EF2140	1		hypothetical protein (pp5)	Mobile and extrachromosomal elements function		12.3		5.1	



EF2142	1		transcriptional regulator, Cro/Ci family (pp5)	Mobile and extrachromosomal elements function					5.8
EF2150	1		FemAB family protein	Protein synthesis				32.9	
EF2151	1	<i>glmS</i>	glucosamine--fructose-6-phosphate aminotransferase	Purines, Pyrimidines, nucleosides, nucleotides		-11.5			
EF2154	1			Hypothetical or unknown protein	-11.0	-13.8	-107.4	-11.3	
EF2159	1	<i>glnA</i>	glutamine synthetase, type I	Amino-acid biosynthesis		-8.7		-5.6	
EF2165	1		NAD-dependent epimerase/dehydratase family protein	central and intermediary metabolism		-13.4			
EF2167	1		glycosyl transferase, group 2 family protein	central and intermediary metabolism		-12.7			
EF2173	1		ISEf1, transposase	Mobile elements function			-86.2		
EF2175	1		licD-related protein	Signal transduction	-6.7	-16.0		-5.1	-7.3
EF2177	1		bacterial sugar transferase	central and intermediary metabolism	-9.7	-12.4	-8.8	-11.8	
EF2180	2		glycosyl transferase, group 2 family protein	central and intermediary metabolism		-6.9			
EF2181	1		glycosyl transferase, group 2 family protein	central and intermediary metabolism		-5.7			
EF2183- EF2182	2		ABC transporter	Transport	-27.6	-21.1	-111.0	-42.0	-60.3
EF2185	1		ISEf1, transposase	Mobile elements function			-5.7		
EF2187	1		IS256, transposase	Mobile elements function		-35.0	-60.2		
EF2194- EF2191	4	<i>rfbACB- EF2191</i>	glucose-1-phosphate thymidyltransferase, dTDP-4-dehydrorhamnose 3,5-epimerase, dTDP-glucose 4,6-dehydratase, dTDP-4-dehydrorhamnose reductase	Purines, Pyrimidines, nucleosides, nucleotides		-12.2	-5.5		
EF2205	1			Hypothetical or unknown protein			-6.0		
EF2209	1			Hypothetical or unknown protein				6.5	
EF2219- EF2218	2	<i>ehk01-err01</i>	two-component system	Signal transduction		7.0		5.8	-7.2
EF2223- EF2221	3		ABC transporter	Transport		10.3	17.8	7.1	
EF2224	1		cell wall surface anchor family protein	cell envelope	5.1	5.6		11.0	6.8
EF2234	1		sugar ABC transporter	Transport		7.8			
EF2238	1		sugar-binding transcriptional regulator, LacI family	Regulatory functions	-16.4	11.0			-187.2
EF2247	1		transcriptional regulator	Regulatory functions				11.3	
EF2255	1		site-specific recombinase, phage integrase family	DNA metabolism				17.4	
EF2271- EF2269	3		PTS system, IIBCD components	Transport				14.5	
EF2280	1			Hypothetical or unknown protein		-6.4	-15.0		
EF2291	1		transcriptional regulator, Cro/Ci family	Regulatory functions		-5.2	-35.7		-7.6
EF2297- EF2293	5	<i>vanYBvanW vanBvanX</i>	D-alanyl-D-alanine carboxypeptidase, vancomycin B-type resistance protein VanW, alpha-keto acid dehydrogenase, D-Ala-D-lactate ligase, D-alanyl-D-alanine dipeptidase	cellular process				17.3	
EF2304	1		transcriptional regulator, Cro/Ci family	Regulatory functions		-6.6	-19.2		
EF2305	1		toprim domain protein	DNA metabolism	17.2			5.8	11.4
EF2307	1			Hypothetical or unknown protein	31.6				9.8
EF2308	1			Hypothetical or unknown protein			-14.8		
EF2320	1			Hypothetical or unknown protein				9.8	
EF2322	1			Hypothetical or unknown protein		12.3		8.1	
EF2342	1			Hypothetical or unknown protein		-6.9		-6.5	-10.1
EF2352	1	<i>lepA</i>	GTP-binding protein LepA	Protein synthesis		-11.0			-7.0
EF2355	1	<i>clpB</i>	ATP-dependent Clp protease, ATP-binding subunit ClpB	Protein fate		77.7	30.8		
EF2359	1		transposase, IS3 family	Mobile elements function				12.9	-8.6
EF2362- EF2361	2	<i>purK-1purB</i>	phosphoribosylaminoimidazole carboxylase, ATPase subunit	Transport					6.1
EF2364	1		xanthine permease	Transport	-14.8				-9.5
EF2365	1	<i>xpt</i>	xanthine phosphoribosyltransferase	Amino-acid biosynthesis		-6.1	-7.2		8.5
EF2370	1		oxidoreductase, Gfo/Idh/MocA family	central and intermediary metabolism	5.9			11.5	10.7
EF2377	1		amino acid permease family protein	Transport		-24.3			
EF2378	1	<i>polC</i>	DNA polymerase III, alpha subunit	DNA metabolism				6.4	
EF2379	1	<i>proS</i>	prolyl-tRNA synthetase	Protein synthesis		-9.3	-6.8		
EF2385	1			Hypothetical or unknown protein		11.6	9.8		

EF2390- EF2393	4	<i>sufDSUB</i>	FeS cluster assembly proteins	central and intermediary metabolism			5.2		5.1
EF2394	1	<i>sufC</i>	ABC transporter	Transport			5.2		5.1
EF2395	1	<i>frr</i>	ribosome recycling factor	Protein synthesis		-16.5		-5.8	
EF2397	1	<i>tsf</i>	translation elongation factor Ts	Protein synthesis	-8.7	-18.9	-33.7		
EF2398	1	<i>rpsB</i>	ribosomal protein S2	Protein synthesis		-8.6	-99.4		
EF2399	1		acetyltransferase, GNAT family	central and intermediary metabolism				66.9	
EF2400	1		RNA methyltransferase, TrmH family	Purines, Pyrimidines, nucleosides, nucleotides				-5.4	
EF2401	1	<i>acyP</i>	acylphosphatase	central and intermediary metabolism	15.0				
EF2404	1			Hypothetical or unknown protein				-5.4	-5.3
EF2410	1	<i>era</i>	GTP-binding protein Era	Fatty acids and phospholipids metabolism		-10.7			
EF2411	1	<i>dgkA</i>	diacylglycerol kinase	Fatty acids and phospholipids metabolism	-7.3				-12.0
EF2417	1		transcriptional regulator, Fur family	Regulatory functions		15.2			
EF2422	3	<i>hom-thrCB</i>	homoserine dehydrogenase, threonine synthase, homoserine kinase	Amino-acid biosynthesis				-6.2	
EF2423	1		transcriptional regulator, ArsR family	Regulatory functions					15.8
EF2424	1		pyrroline-5-carboxylate reductase	Amino-acid biosynthesis	-5.8	-5.5		-11.6	-6.2
EF2429	1	<i>guaC</i>	GMP reductase	Purines, Pyrimidines, nucleosides, nucleotides				-5.4	
EF2430	1		xanthine/uracil permease family protein	Transport			35.1		
EF2431	1		chlorohydrolase family protein	Other metabolism			19.3		11.3
EF2443	1	<i>rpsT</i>	ribosomal protein S20	Protein synthesis	-16.2	-40.6	-46.8	-10.0	
EF2444	1		acyl-CoA thioester hydrolase	Fatty acids and phospholipids metabolism	-56.2	-106.6	-162.1	-26.7	-33.7
EF2451	1	<i>coaD</i>	pantetheine-phosphate adenylyltransferase	Other metabolism		8.1			
EF2456	1	<i>pycA</i>	pyruvate carboxylase	Fatty acids and phospholipids metabolism			-5.2		
EF2460	1		GTP-binding protein TypA	Other metabolism		-6.2	-8.0		
EF2461	1		inositol monophosphatase protein family	Signal transduction	-7.1				6.9
EF2462	1			Hypothetical or unknown protein				7.6	
EF2464	1			Hypothetical or unknown protein				8.8	
EF2470	1		HD domain protein	Signal transduction			-17.2		
EF2473	1		ribosomal-protein-alanine acetyltransferase	Protein synthesis				7.6	9.0
EF2474	1		ribosomal-protein-alanine acetyltransferase	Protein synthesis				5.3	9.0
EF2486- EF2485	2		ABC transporter	Transport	-11.2			-19.3	
EF2489	1	<i>cpsG</i>	MurB family protein	cell envelope			-6.5		
EF2491	1	<i>cpsE</i>	glycosyl transferase, group 2 family protein	central and intermediary metabolism		-9.2	-152.7	-6.7	
EF2494	1	<i>cdsA</i>	phosphatidate cytidylyltransferase	Fatty acids and phospholipids metabolism		-5.2			
EF2495	1	<i>uppS</i>	undecaprenyl diphosphate synthase	Fatty acids and phospholipids metabolism		-22.2	-6.1		
EF2498- EF2497	2		ABC transporter	Transport	-5.1				
EF2500	1		GcvH family protein	Amino-acid biosynthesis				-12.2	
EF2502	1		cell division protein, FtsW/RodA/SpoVE family	cellular process		-29.6		8.8	
EF2512	1		lipoprotein	cell envelope		7.9			
EF2517	1		conjugal transfer protein	Signal transduction				21.6	
EF2521	1			Hypothetical or unknown protein			9.8		
EF2529	1			Hypothetical or unknown protein				5.7	
EF2535	1		nucleotidyltransferase domain protein	Protein fate					6.9
EF2549	1	<i>upp</i>	uracil phosphoribosyltransferase	Purines, Pyrimidines, nucleosides, nucleotides	-5.5		-7.4		
EF2555	1	<i>tdK</i>	thymidine kinase	Purines, Pyrimidines, nucleosides, nucleotides		5.2			
EF2556	1		fumarate reductase flavoprotein subunit precursor	central and intermediary metabolism		-35.9	-10.0	-12.8	
EF2572	1		molybdenum transport domain protein	Transport	7.4			5.6	
EF2586	1		cobyric acid synthase	Other metabolism		-22.2			
EF2587	1		inosine-uridine preferring nucleoside hydrolase	Purines, Pyrimidines, nucleosides, nucleotides		-7.0		-6.6	



EF2593- EF2592	2		ABC transporter	Transport	8.9	6.9	7.4		
EF2594	1		transcriptional regulator, TetR family	Regulatory functions	10.1		11.9		
EF2595	1	<i>gmk-1</i>	guanylate kinase	Purines, Pyrimidines, nucleosides, nucleotides				11.7	
EF2599	1		transcriptional antiterminator, bglG family	Regulatory functions				17.4	
EF2603	1			Hypothetical or unknown protein	-24.8		-9.6	-23.8	-55.6
EF2607- EF2612	6	<i>atpFHAGDC</i>	ATP synthase F1	Transport	-10.0	-10.1	-9.7	12.2	
EF2613	1			Hypothetical or unknown protein					-11.9
EF2616	1	<i>smpB</i>	SsrA-binding protein	Protein synthesis				9.8	6.9
EF2620	1	<i>secG</i>	preprotein translocase, SecG subunit	Protein fate		-10.9			
EF2632	1		IS256, transposase	Mobile elements function		-39.4	-55.1		
EF2634- EF2633	2	<i>groELgroES</i>	chaperonin	Protein fate	15.9	41.3	6.7	6.6	
EF2641- EF2642	2		glycine betaine/L-proline ABC transporter	Transport		-7.6		-7.7	
EF2652- EF2649	4		spermidine/putrescine ABC transporter, ATP-binding subunit	Transport	-25.1	-16.8	-67.7	-10.4	-17.2
EF2657	4		membrane protein, putative	cell envelope		-12.7			18.8
EF2661	1		diacylglycerol kinase catalytic domain protein	Fatty acids and phospholipids metabolism		-8.5			-10.2
EF2668	1	<i>mgtE</i>	magnesium transporter	Transport		-7.3	-7.0		
EF2669	1		ribosomal large subunit pseudouridine synthase, RluD subfamily	Protein synthesis				8.0	
EF2681	1		hydrolase, haloacid dehalogenase-like family	central and intermediary metabolism	-16.0			-7.3	-5.5
EF2687	1			Hypothetical or unknown protein					6.1
EF2694	1			Hypothetical or unknown protein	14.3			8.4	
EF2695	1			Hypothetical or unknown protein		19.4			
EF2698	1		tellurite resistance protein	cellular process		8.2			
EF2700	1		MutT/nudix family protein	DNA metabolism		5.6		13.2	
EF2701	1		acetyltransferase, GNAT family	central and intermediary metabolism					-16.2
EF2703	1		transcriptional regulator	Regulatory functions		15.2			
EF2706	1		RNA methyltransferase, TrmA family	Purines, Pyrimidines, nucleosides, nucleotides		-11.1			
EF2708	1		membran protein, putative	cell envelope	-5.5	-48.3	-26.4	-8.7	-15.9
EF2713	1		cell wall surface anchor family protein	cell envelope		22.9	13.0		
EF2715	1	<i>rplL</i>	ribosomal protein L7/L12	Protein synthesis		-19.3	-246.9	-10.4	
EF2716	1	<i>rplJ</i>	ribosomal protein L10	Protein synthesis			-6.5		
EF2718	1	<i>rplA</i>	ribosomal protein L1	Protein synthesis		-7.0	-6.0	-9.9	
EF2719	1	<i>rplK</i>	ribosomal protein L11	Protein synthesis		-6.2	-7.4		
EF2720	1		ABC transporter, ATP-binding protein	Transport	-7.4				-5.7
EF2729	1	<i>nusG</i>	transcription antitermination protein NusG	Regulatory functions		-11.5			
EF2730	1	<i>secE</i>	preprotein translocase, SecE subunit	Protein fate	-5.5				
EF2731	1	<i>rpmG-2</i>	ribosomal protein L33	Protein synthesis					-6.8
EF2733	1	<i>murB</i>	UDP-N-acetylenolpyruvoylglucosamine reductase	cell envelope		-6.2		-9.4	
EF2739- EF2738	2	<i>ahpCF</i>	alkyl hydroperoxide reductase, thioredoxin reductase/glutathione-like protein	cellular process		-5.8	-24.9	-6.2	8.2
EF2750- EF2746	5	<i>dltXABCD</i>	D-alanine--poly(phosphoribitol) ligase, membrane protein DtlB, dltD protein	cell envelope	-10.2	-14.5	-7.2	-9.8	-15.7
EF2759	1		tetrapyrrole methylase family protein	central and intermediary metabolism					5.3
EF2766	1	<i>recR</i>	recombination protein RecR	DNA metabolism	-5.5	-13.7			
EF2774	1			Hypothetical or unknown protein	-5.8				
EF2780	1			Hypothetical or unknown protein	-8.0	-30.1		-25.7	-16.8
EF2784	1			Hypothetical or unknown protein					
EF2785	1			Hypothetical or unknown protein	6.1			5.3	
EF2787	1		rhodanese family protein	cellular process	-5.4	-10.9			
EF2790	1			Hypothetical or unknown protein	-5.3	-14.0	-16.7		
EF2794	1		membrane protein, putative	cell envelope	5.3				
EF2795	1		LysM domain lipoprotein	cell envelope	6.7				
EF2797	1	<i>gsp64</i>	general stress protein	cellular process	6.2			5.6	5.3
EF2800	1		ISEf1, transposase (pp6)	Mobile and extrachromosomal elements function			-16.8	-7.9	
EF2811	1		minor structural protein (pp6)	Mobile and extrachromosomal elements function				7.6	
EF2828	1		transcriptional regulator, ArpU family (pp6)	Mobile and extrachromosomal elements function				26.2	-143.9

EF2829	1		hypothetical protein (pp6)	Mobile and extrachromosomal elements function	19.5			28.8	24.5
EF2833	1		hypothetical protein (pp6)	Mobile and extrachromosomal elements function	-10.2		-10.9		-6.7
EF2834	1		hypothetical protein (pp6)	Mobile and extrachromosomal elements function	-32.8			-12.5	-20.8
EF2839	1		DnaD domain protein (pp6)	Mobile and extrachromosomal elements function	41.4			98.1	
EF2847	1		hypothetical protein (pp6)	Mobile and extrachromosomal elements function	-11.3	-11.8			
EF2848	1		hypothetical protein (pp6)	Mobile and extrachromosomal elements function		5.4			
EF2850	1		hypothetical protein (pp6)	Mobile and extrachromosomal elements function				8.4	
EF2852	1		transcriptional regulator, Cro/Ci family (pp6)	Mobile and extrachromosomal elements function					7.9
EF2853	1		hypothetical protein (pp6)	Mobile and extrachromosomal elements function		16.7			
EF2856	1	<i>rpmG-3</i>	ribosomal protein L33	Protein synthesis		-12.3	-141.7		
EF2859	1		4-oxalocrotonate tautomerase, putative	Other metabolism		7.8			
EF2860	1		ErfK/YbiS/YcfS/YnhG family protein, putative	cell envelope	9.8	-24.9			
EF2863	1		endo-beta-N-acetylglucosaminidase	central and intermediary metabolism		-7.9		5.3	
EF2868	1			Hypothetical or unknown protein					6.0
EF2870	1		HD domain protein	Signal transduction		8.3			9.6
EF2882- EF2875	8	<i>fabDGF-2- accBfabZacc CDA</i>	ACP S-malonyltransferase, 3-ketoacyl-ACP reductase, 3-oxoacyl-ACP synthase, 3-oxoacyl-ACP synthase, (3R)-hydroxymyristoyl-(acyl-carrier-protein) dehydratase	Fatty acids and phospholipids metabolism		-9.3		5.4	5.4
EF2883	1	<i>fabK</i>	enoyl-(acyl-carrier-protein) reductase II	Fatty acids and phospholipids metabolism		-9.8	-76.2		
EF2885	1	<i>fabH</i>	3-oxoacyl-(acyl-carrier-protein) synthase III	Fatty acids and phospholipids metabolism			-129.8		
EF2886	1		transcriptional regulator, MarR family	Regulatory functions				5.2	
EF2898	1		peptidyl-prolyl cis-trans isomerase, cyclophilin-type	Protein synthesis		-12.6	-82.3	-7.3	
EF2908	1		glycosyl transferase, group 2 family protein	central and intermediary metabolism		-9.8	-112.7		
EF2910	1		potassium uptake protein	Transport		-9.7	-7.1	10.9	
EF2920- EF2919	2		ABC transporter	Transport		-6.5	5.1	7.0	
EF2923- EF2924	2	<i>rnjA</i>	metallo-beta-lactamase superfamily protein	Purines, Pyrimidines, nucleosides, nucleotides		-10.1	-141.7		
EF2929	1		membrane protein, putative	cell envelope	6.6				
EF2933	1		DNA-binding protein	Regulatory functions		8.1			7.7
EF2934	1	<i>thil</i>	thiazole biosynthesis protein Thil	Purines, Pyrimidines, nucleosides, nucleotides	-5.3				
EF2935	1		xanthine/uracil permeases family protein	Purines, Pyrimidines, nucleosides, nucleotides			-7.2	-5.0	
EF2938	1		hypothetical protein (pp7)	Mobile and extrachromosomal elements function				6.8	
EF2939	1		cold-shock domain family protein (pp7)	Mobile and extrachromosomal elements function	-7.8				-13.3
EF2940	1		hypothetical protein (pp7)	Mobile and extrachromosomal elements function		14.8	12.9	5.8	
EF2942	1		hypothetical protein (pp7)	Mobile and extrachromosomal elements function		42.1	17.3	21.8	
EF2943	1		hypothetical protein (pp7)	Mobile and extrachromosomal elements function	8.5	22.9		43.0	
EF2944	1		hypothetical protein (pp7)	Mobile and extrachromosomal elements function				5.6	
EF2948	1		DNA primase domain protein (pp7)	Mobile and extrachromosomal elements function			7.4		
EF2950	1		hypothetical protein (pp7)	Mobile and extrachromosomal elements function	7.8	-20.5	-49.4	26.7	
EF2951	1		hypothetical protein (pp7)	Mobile and extrachromosomal elements function		6.9			
EF2953	1		hypothetical protein (pp7)	Mobile and extrachromosomal elements function	7.4	11.6			
EF2958	1	<i>hypR</i>	transcriptional regulator, LysR family	Regulatory functions					7.3

EF2959	1		ribose uptake protein, putative	Transport	33.9				
EF2960	1		ribose transporter protein RbsD	Transport	7.8				
EF2961	1	<i>rbsK</i>	ribokinase	central and intermediary metabolism	79.2		5.7		
EF2964	1		putative transport protein, SgaT family	Transport	14.6				
EF2966	1		transcriptional antiterminator, bglG family	Regulatory functions	20.0				5.6
EF2986- EF2685	2		ABC transporter	Transport		9.8	15.7		
EF2994	1		aminotransferase, class V	Protein synthesis	11.6		5.6		
EF3008	1			Hypothetical or unknown protein					-14.1
EF3009	1			Hypothetical or unknown protein				-6.1	
EF3022	1		sodium:dicarboxylate symporter family protein	Transport	-25.4	-6.3			-5.6
EF3030	1			Hypothetical or unknown protein		27.3			
EF3033- EF3029	4		PTS system, IIA/BCD components	Transport		22.9			
EF3034	1		transcriptional regulator, GntR family	Regulatory functions	-5.4	5.2			
EF3035	1		universal stress protein family	cellular process			5.8		-6.4
EF3041	1		pheromone binding protein	Signal transduction	-6.2		-7.8		
EF3044	1	<i>nagA-2</i>	N-acetylglucosamine-6-phosphate deacetylase	central and intermediary metabolism					-14.1
EF3045- EF3046	2		PTS system, IIB components	Transport	-19.7		7.3		-12.6
EF3053	1			Hypothetical or unknown protein			17.8		
EF3054	1			Hypothetical or unknown protein					-5.7
EF3056	1	<i>srtA</i>	sortase family protein	Protein fate	30.6	-5.6	5.1	9.5	
EF3060	1	<i>salA</i>	secreted lipase	cellular process		-11.9	-36.7	-8.6	
EF3062- EF3061	2	<i>mreC/mreD</i>	rod shape-determining proteins	cellular process		-11.1			-7.8
EF3063	1		immunity protein, putative	cellular process	5.8	-40.1		12.9	
EF3064	1	<i>pnpA</i>	polyribonucleotide nucleotidyltransferase	Purines, Pyrimidines, nucleosides, nucleotides		-6.1	-9.4		
EF3067	1		hydrolase, haloacid dehalogenase-like family	central and intermediary metabolism				-6.4	
EF3069	1		formate/nitrite transporter family protein	Transport		-14.1			
EF3070	1	<i>rpsD</i>	ribosomal protein S4	Protein synthesis		-12.0	-183.1	-6.4	
EF3072	1		BioY family protein	Transport		7.0		5.2	
EF3076	1		cell wall surface anchor family protein	cell envelope			9.8		
EF3079	1		acetyltransferase, GNAT family	central and intermediary metabolism				-5.1	
EF3080	1	<i>pepT-2</i>	peptidase T	Protein fate	-6.7				
EF3085	1	<i>fatB</i>	iron transporter	Transport	5.9				
EF3087	1			Hypothetical or unknown protein	14.1				
EF3090	1		isochorismatase family protein	Other metabolism	7.9			-6.1	
EF3094	1	<i>ftsY</i>	signal recognition particle-docking protein FtsY	Transport		-7.7	-7.4		
EF3097	1	<i>rnc</i>	ribonuclease III	Purines, Pyrimidines, nucleosides, nucleotides			-6.2		
EF3100	1		IS256, transposase	Mobile elements function	5.2	-39.6	-55.2		-7.6
EF3103	1		membrane protein, putative	cell envelope					-9.4
EF3110- EF3106	5		peptide ABC transporter	Transport	5.4		9.0		5.6
EF3111	1	<i>acpP</i>	acyl carrier protein	Other metabolism		-7.0			-9.0
EF3114	1		DAK2 domain protein	Other metabolism	-6.6		-10.1	-6.1	
EF3116	1	<i>rpmB</i>	ribosomal protein L28	Protein synthesis	-16.6	-24.9	-60.0	-8.0	
EF3121- EF3120	2	<i>ireK (prkC)</i>	Protein phosphatase 2C-serine/threonine protein kinase	Signal transduction			-6.7		
EF3124	1	<i>def-2</i>	polypeptide deformylase	Protein synthesis				5.3	7.9
EF3129	1		D-alanyl-D-alanine carboxypeptidase	cell envelope		-9.8			
EF3130	1			Hypothetical or unknown protein					6.7
EF3131	1			Hypothetical or unknown protein	7.4				18.4
EF3139- EF3136	4		PTS system, IICDBA components	Transport			19.3		
EF3140	1		alcohol dehydrogenase, iron-containing	central and intermediary metabolism	-5.7	18.5			-5.7
EF3148	1	<i>pgsA</i>	CDP-diacylglycerol-glycerol-3-phosphate 3-phosphatidyltransferase	Fatty acids and phospholipids metabolism	-14.2	-9.2	-105.8	-22.1	
EF3165	1	<i>maf</i>	maf protein	Regulatory functions	-24.6	-56.0	-85.1	-62.7	-25.6
EF3171	1			Hypothetical or unknown protein	13.4	18.3	9.8	10.5	11.1
EF3173	1			Hypothetical or unknown protein		-13.4	-6.1	-5.8	-12.7
EF3187	1		cell wall surface anchor family protein	cell envelope	10.5				

EF3188	1			Hypothetical or unknown protein	12.2			6.1	
EF3202	1	<i>rpsN-3</i>	ribosomal protein S14	Protein synthesis					8.3
EF3207	1		dihydrouridine synthase family protein	Purines, Pyrimidines, nucleosides, nucleotides		-11.6			
EF3209- EF3208	2		ABC transporter	Transport	-7.1	-11.4			5.6
EF3210- EF3213	4		PTS system, IIA/BCD components	Transport	11.1	176.2	161.6		
EF3215	1		IS256, transposase	Mobile elements function	5.0	-33.8	-59.2		
EF3230	1	<i>rpsI</i>	ribosomal protein S9	Protein synthesis		-8.8	-199.7	-6.5	
EF3231	1	<i>rplM</i>	ribosomal protein L13	Protein synthesis		-9.1	-5.4		
EF3235	1		gluconate kinase	central and intermediary metabolism				12.4	
EF3238	1			Hypothetical or unknown protein					-5.7
EF3240	1	<i>birA</i>	BirA bifunctional protein	Protein synthesis	12.4	-8.7	-67.7	-7.5	
EF3245	1		cell-envelope associated acid phosphatase	cell envelope	8.8	-10.1	-13.2	7.9	
EF3247	1			Hypothetical or unknown protein				15.3	
EF3254	1	<i>ubiA</i>	1,4-dihydroxy-2-naphthoate octaprenyltransferase	Other metabolism		-5.0		-6.7	
EF3257	1		oxidoreductase, pyridine nucleotide-disulfide family	Protein fate	-15.7	-22.9	-42.9	-19.5	-5.3
EF3267	1	<i>folE</i>	GTP cyclohydrolase I	Other metabolism			19.3		
EF3271	1			Hypothetical or unknown protein	11.1	40.4		30.2	
EF3273	1			Hypothetical or unknown protein	367.4		19.3	222.9	66.1
EF3275	1		hydantoinase/ oxoprolinase	Amino-acid biosynthesis	5.4		9.8	12.8	
EF3277	1		cytosine permease, putative	Transport			25.8		
EF3283	1	<i>ctsR</i>	transcriptional regulator CtsR	Regulatory functions	6.6				
EF3285	1		PTS system, IIC component	Transport			19.3		
EF3286	1		phosphosugar-binding transcriptional regulator, putative	Regulatory functions	129.4	5.2		336.1	
EF3289- EF3290	2	<i>croRS</i>	two-component system	Signal transduction			7.0		
EF3293	1	<i>guaB</i>	inosine-5'-monophosphate dehydrogenase	Signal transduction			-9.5		-5.0
EF3295	1			Hypothetical or unknown protein	12.4			5.1	
EF3297	1			Hypothetical or unknown protein		12.2	10.2		6.6
EF3301	1			Hypothetical or unknown protein	23.7	5.4		14.3	28.3
EF3302	1			Hypothetical or unknown protein				14.9	
EF3307	1			Hypothetical or unknown protein			-23.1		-63.5
EF3314	1		cell wall surface anchor family protein	cell envelope		7.0			
EF3322- EF3317	6	<i>citCDEFXM</i>	citrate lyase ligase, 2-(5'-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase, oxaloacetate decarboxylase	central and intermediary metabolism	5.1	30.8	19.3	10.5	
EF3327	1	<i>citH</i>	citrate transporter	Transport		136.2		13.5	
EF3328	1	<i>citO</i>	citrate dependant regulator	Regulatory functions		136.2		13.5	
EF3333	1	<i>rpmH</i>	ribosomal protein L34	Protein synthesis		-23.3	-10.2		7.9

**Table S2 :** Comparison of detected sRNA and their flanking genes RNA expression. The names of the genes and sRNA are indicated. In the "Name" column are found names of sRNA are that have been previously defined by Fouquier d'Hérouel *et al.* (Fouquier d'Hérouel *et al.*, 2011), Innocenti *et al.* (Innocenti *et al.*, 2015), or Shioya *et al.* (Shioya *et al.*, 2011). The product of each gene is presented in the "Annotation" column. For each stress (Bile salt (BS, 0.06%), urine, serum, acid pH (pH 5.8) and oxidative (H<sub>2</sub>O<sub>2</sub> 1mM)), the FC (Fold-change of the RNA expression comparing stress condition and V19 strain grown in BHI) are indicated with a cut-off  $\geq 2$ ,  $p < 0.05$  and  $\log\text{CPM} > 1$  for both sRNA and genes. sRNA that had their coordinates confirmed on the V583 genome are indicated (\*).

Small RNA	Flanking genes	Name	Annotation	BS	Urine	Serum	Acid	Oxidative
EF0096							2.2	
EF0097	EF0096		conserved hypothetical protein	-2.9				
	EF0097		regulatory protein pfoR	-4.3	-6.6		-5.2	
EF0129		Ref50					4.3	3.8
EF0130	EF0129		transcriptional regulator, Cro/CI family				4.8	
	EF0130		hypothetical protein		-4.0		-2.1	-4.3
EF0147		Ref 52		2.4				
EF0149	EF0147		hypothetical protein	-4.5				
	EF0149		aggregation substance	-24.6	19.1	10.1		-99.4
EF0156							-2.6	
EF0157	EF0156		conserved hypothetical protein					
	EF0157		conserved domain protein					3.0
EF0401				-2				2.3
EF0402	EF0401	<i>pcp</i>	pyrrolidone-carboxylate peptidase	-2.5				
	EF0402	<i>nhaC-1</i>	Na <sup>+</sup> /H <sup>+</sup> antiporter	-2.5			-7.2	
EF0408		Ref 57		4.3	4.1	3.3	4.5	4
EF0409*	EF0408		PTS system, IIA component					
	EF0409		hypothetical protein		4.9	4.6	4.7	
EF0513					2.4			
EF0514	EF0513		conserved hypothetical protein					4.3
	EF0514		IS1216, transposase	2.4	-4.7	-13.8		
EF0595		Ref 64		2.2			2.8	2.3
EF0598	EF0595		ISEf1, transposase		-3.6	-90.6	-4.2	
	EF0598		conserved hypothetical protein, degenerate					-6.4
EF0605		Ref 25C		2.3	2	2.4	2.8	2.3
EF0606*	EF0605							
	EF0606		Dps family protein		2.6	5.8		
EF0628				2.8				3.1
EF0629	EF0628		conserved hypothetical protein					
	EF0629		oxidoreductase, aldo/keto reductase family	8.5		19.3		16.4
EF0631					2.3		3.3	
EF0633	EF0631		cardiolipin synthetase, putative					
	EF0633	<i>tryS-1</i>	tyrosyl-tRNA synthetase		6.8			5.0
EF0723		Ref 66			-2.0	-2.8		
EF0724	EF0723		hypothetical protein				-3.5	-3.5
	EF0724	<i>gatC</i>	glutamyl-tRNA(Gln) amidotransferase, C subunit	3.3				2.9
EF0747					2.0			
EF0748	EF0747		Hypothetical or unknown protein	4.5	-4.1		11.3	
	EF0748		rhodanese family protein					
EF0806		Ref 67					2.5	
EF0807	EF0806		amino acid ABC transporter				8.3	
	EF0807		pheromone binding protein, putative	3.4				
EF0827		Ref 68				-3.2		
EF0828	EF0827		oxidoreductase, Gfo/Idh/MocA family	-5.4			5.8	
	EF0828		conserved hypothetical protein					
EF0845				3.0			3.2	
EF0846	EF0845	<i>murF</i>	UDP-N-acetylmuramoylalanyl-D-glutamyl-diaminopimelate-D-alanyl-ligase	-3.2	-2.2		-5.1	
	EF0846		ATP-dependent RNA helicase, DEAD/DEAH box family			-6.6		3.5
EF0865		Ref 27C- Ref69 antisense					2.0	2.2
EF0867	EF0865		glycine betaine/carnitine/choline ABC transporter	2.2		-8.9		4.1
	EF0867		glyoxalase family protein	2.4				
EF0868							4.7	
EF0869	EF0868	<i>queA</i>	S-adenosylmethionine:tRNA ribosyltransferase-isomerase					
	EF0869		transcriptional regulator, Cro/CI family					-3.6

EF0869		Ref 19C			4.1	4.8		
EF0871*	EF0869		transcriptional regulator, Cro/CI family					-3.6
	EF0871		cation-transporting ATPase, E1-E2 family				-2.5	
EF0913		Ref 70				-2.4		
EF0914	EF0913		transposase, putative				4.0	
	EF0914	<i>infC</i>	translation initiation factor IF-3				-3.6	
EF1097		SsrA		2.0	2.4	3.4	2.7	2.4
EF1098*	EF1097		Hypothetical or unknown protein	-5.4			-4.4	4.0
	EF1098		Hypothetical or unknown protein	-38.0	-8.3		-8.7	-35.3
EF1146		Ref 72		2.1				
EF1147	EF1146		DNA-directed RNA polymerase, delta subunit, putative	-2.3	-3.6	-4.7	-3.7	-3.3
	EF1147	<i>pyrG</i>	CTP synthase	-2.6	-4.8	-37.8	-5.8	
EF1151		RnpB					-2.1	
EF1152*	EF1151		cell division protein DivIVA				2.4	5.4
	EF1152		conserved hypothetical protein	2.6	-3.1		2.3	
EF1204		Ref 32C		4.4	3.1	2.4	4.3	2.4
EF1205*	EF1204		conserved hypothetical protein	-2.9	-4.3	-3.1	-2.9	
	EF1205		transcriptional regulator, putative, authentic frameshift		3.7			
EF1244		Ref 73				-2.1		
EF1246	EF1244		oxidoreductase, Gfo/Idh/MocA family					
	EF1246		Hypothetical or unknown protein	-4.5	-4.4		-3.2	
EF1303		Ref 76					2	
EF1304	EF1303		transcriptional regulator, LysR family	6.1	22.7	9.8		
	EF1304		magnesium-translocating P-type ATPase	7.2		7.2		
EF1350		Ref 77						-2.2
EF1351	EF1350		conserved hypothetical protein		4.2			
	EF1351		hypothetical protein					
EF1368		Ref 78						-2.7
EF1369*	EF1368		Hypothetical or unknown protein		4.9		4.0	8.5
	EF1369		transcriptional regulator, Cro/CI family					
EF1378				2.7			2.6	
EF1379	EF1378		transcriptional regulator, putative					
	EF1379	<i>alaS</i>	alanyl-tRNA synthetase		-5.2	-40.0		
EF1412		Ref 79		-2.4				
EF1413	EF1412		Hypothetical or unknown protein		4.1			-5.8
	EF1413		msrC protein, putative					
EF1452		Ref 81			2.7	3.1	2.6	2.1
EF1453	EF1452		adenine methyltransferase, putative					
	EF1453		conserved hypothetical protein					
EF1513				2.7	2.7	3.2	2.9	2.2
EF1515	EF1513		pheromone binding protein					
	EF1515		transcription antiterminator, bglG family	-4.6				
EF1540		Ref 82			-3.5	-2.9		
EF1541	EF1540		ribosomal large subunit pseudouridine synthase B	-2.3				
	EF1541		conserved hypothetical protein			3.3	4.5	-4.5
EF1599		Ref 83		3	2.4		3.3	
EF1601	EF1599		TPR domain transcriptional regulator, Cro/CI family					-2.5
	EF1601		PTS system, IIABC components		9.4	9.8		
EF1634				2.7			3.1	
EF1635	EF1634		propanediol utilization protein PduU					
	EF1635		propanol dehydrogenase PduQ, putative					
EF1635					2.1			-2.2
EF1637	EF1635		propanol dehydrogenase PduQ, putative					
	EF1637		ATP:cob(D)alamin adenosyltransferase, putative					
EF1935		Ref 84			2.7			
EF1936	EF1935		hypothetical protein	3.0	2.6			2.2
	EF1936		hypothetical protein					
EF1979		SsrS		-6.8	-3.8	-2.9	-4.4	
EF1980*	EF1979		ATPase, AAA family					
	EF1980		hypothetical protein	4.7			4.8	-2.2
EF1989		Ref 87- Ref 88					3.3	2.2
EF1990	EF1989	<i>hemH</i>	ferrochelatase	-2.1				
	EF1990		hypothetical protein (pp4)				5.6	
EF2021		Ref 92-93A-93B			4.6	3.9		3.9
EF2022	EF2021		hypothetical protein					
	EF2022		conserved domain protein		2.1			
EF2085		Ref 102		3.1	2.8	3.8	4.3	3.3
EF2086	EF2085		hypothetical protein (pp5)	-2.0				-5.5
	EF2086		endolysin (pp5)				5.2	
EF2205		Ffs			2.4	2.9	2.2	
EF2206*	EF2205		Hypothetical or unknown protein		4.5	-6.0		
	EF2206		cytidine/deoxycytidylate deaminase family protein					



EF2255		Ref 33B			-3.0		-3.4	
EF2257	EF2255		site-specific recombinase, phage integrase family				17.4	
	EF2257		PTS system, IIC component, putative					
EF2365							2.2	
EF2366	EF2365	<i>xpt</i>	xanthine phosphoribosyltransferase		-6.1	-7.2		8.5
	EF2366		conserved hypothetical protein	-2.1	-3.6		-2.7	
EF2463						-2.3		
EF2464	EF2463		voltage-gated chloride channel family protein					
	EF2464		Hypothetical or unknown protein				8.8	
EF2484		Ref 106			2	2.2		2.6
EF2485	EF2484		conserved hypothetical protein					
	EF2485		ABC transporter	-11.2	-4.9		-19.3	-3.4
EF2542		Ref 108 - Ref 107		2.4	2			3.2
EF2543	EF2542		hypothetical protein					
	EF2543		transcriptional regulator, putative					
EF2665								-2.2
EF2666	EF2665		RNA methyltransferase, TrmH family	-4.1				-2.9
	EF2666		ribosomal RNA large subunit methyltransferase A, putative	-4.3				2.4
EF2696				-2.2				
EF2697	EF2696		MutT/nudix family protein					
	EF2697		hypothetical protein	2.7	4.4	4.7		
EF2740		Ref 109			4.6			3.1
EF2741	EF2740		O-methyltransferase family protein					4.8
	EF2741	<i>lplA-2</i>	lipoate-protein ligase A				2.3	
EF2770		Ref 110			-2.5			-3
EF2771	EF2770		conserved hypothetical protein					
	EF2771		conserved hypothetical protein					
EF2771	EF2772		conserved hypothetical protein	2.1			2.9	
	EF2772		drug resistance transporter, putative, authentic frameshift		2.7			
EF2778		Ref 111						-4.5
EF2780	EF2778		conserved hypothetical protein					
	EF2780		Hypothetical or unknown protein	-5.8	-30.1		-25.7	-16.8
EF2858						-2.1		
EF2859	EF2858	<i>thrS</i>	threonyl-tRNA synthetase					
	EF2859		4-oxalocrotonate tautomerase, putative		4.7			
EF2863				2.3			2.8	
EF2864	EF2863		endo-beta-N-acetylglucosaminidase		-7.9		5.3	
	EF2864		conserved domain protein				2.6	
EF2944				2.1		2.4	2.2	
EF2945	EF2944		hypothetical protein (pp7)		4.3		5.6	
	EF2945		hypothetical protein					
EF3081							2.9	
EF3082	EF3081		pheromone binding protein					
	EF3082		iron compound ABC transporter, substrate-binding protein					
EF3086				2.3	-2.4			-2.3
EF3087	EF3086		conserved hypothetical protein					
	EF3087		Hypothetical or unknown protein	14.1			-4.9	
EF3087		Ref 116		2.6				
EF3088	EF3087		Hypothetical or unknown protein	14.1			-4.9	
	EF3088		hypothetical protein					
EF3158							-2.3	
EF3160	EF3158		hydrolase, haloacid dehalogenase-like family	-2.4			-3.4	-2.3
	EF3160		hypothetical protein					
EF3160					8.7	8.8	6.1	
EF3161	EF3160		hypothetical protein					
	EF3161		hypothetical protein					
EF3247		Ref 118		2.9	3.1	3.3	3.4	2.2
EF3248	EF3247		Hypothetical or unknown protein	2.6	-4.1	-4.2	15.3	
	EF3248		hypothetical protein					-3.9
EF3260		Ref 5J			5.7	3		
EF3261	EF3260		heptaprenyl diphosphate synthase, component II, putative		-3.7			
	EF3261		transcriptional regulator, AbrB family					
EF3262		Ref 46			4.1	4.5		
EF3263	EF3262		transcriptional regulator, PemK family	-2.0	-4.1		-3.1	-2.8
	EF3263		conserved hypothetical protein	4.8				
EF3314				2.5	2.2	2.6		3.8
EF3315*	EF3314		cell wall surface anchor family protein	4.8	7.0		3.9	
	EF3315		CitG family protein		3.8		4.5	

**Annexe 2 : Résultats complémentaires de la publication 2 « Characterization of a Bgl  
antiterminator in *Enterococcus faecalis* »**

**Table S1** : List and sequences of oligonucleotides used in this study.

Name	Sequence	Use
SRC85_SP1	GAACAATCCCGTTAAACATGTT	RACE-PCR, Northern Blot
SRC85_SP2	ATTGTACACCACATTGTATTAC	RACE-PCR
SRC85_SP3	CACATTGTATTACGAAATTCTTC	RACE-PCR
ef1515_SP2	TATTAACCCGACGCCTTTAC	RACE-PCR
ef1515_SP3	TTCTCCTGTCCTTCGTCAAG	RACE-PCR
ef1515L	AGTAGCTGTCGGTAAAGGCG	qPCR
ef1515R	AGTCCTTCTGGCTCCATCAC	qPCR, RT-PCR, RACE-PCR
ef1516L	ATGTGCACGTCCTGAGAAGA	qPCR
ef1516R	GCCAGCTGTCCAATGCATAA	qPCR
GyrAL	GATGGGGAAATCAGGGATTC	qPCR
GyrAR	TCTTTTCCATTCCGCATTC	qPCR
src85_Forward	GAAGAATTCGTAATACAATGTG	Northern Blot, RT-PCR
RPV_AGEnt+1	ACATTGACTTTACCAAGTTTAG	Overexpression
FRV_AGEnt_Term	AATTACAGCACGTGTTGCTTT	Overexpression
FPI_SRC85_+1	CTAAACTGGTAAAGTCAATGTGTTAACAATGAATAGCGTTTTTC	Overexpression
RPI_SRC85_Term	AAAGCAACACGTGCTGTAATTTTCTCCTGTCCTTCGTCGAAG	Overexpression
pLT06_1_bis	GAATTCGAGCTCGGTACC	Mutant construction
pLT06_2	CTGCAGGTGCATAAACCC	Mutant construction
SRC_85_1bis	TCTCGAATTCCTACTATTACGATAAAGAA	Mutant construction
SRC_85_2bis	TATCTCTAGACGCTATTCATTTGTTAACAG	Mutant construction
SRC_85_3bis	ATCATCTAGAGTTCTATGCTTTTTTTGTTGTT	Mutant construction
SRC_85_4bis	ATGACCATGGGCTAAACCAACCGCAAATAAT	Mutant construction
SRC_85_5	AAACGAAGCATTGTAAGAA	Mutant construction
SRC_85_6	TGAAGCCTTTAATGGCTTGG	Mutant construction
ef1515_1_IVR	ACGATCCCCGGGTACCGAGCTCGAATTCAAAAGTCTATTTCCTTACGTT	Mutant construction
ef1515_2_IVRbis	TCAATGTGAATCGATGTCAAGAACAAGTACAGCATT	Mutant construction
ef1515_3_IVRbis	TGTAATTGTTCTTGACATCGATTACATTGAACGA	Mutant construction
ef1515_4_IVR	GTTTCGCTGGGTTTATCGACCTGCAGTTCAAACCTAAGAAGAAGTCTACAA	Mutant construction
ef1515_6	CCAACATCATTGGATGATTTCG	Mutant construction
ef3023_1_IVR	ATCCCCGGGTACCGAGCTCGAATTCAGGTGAAAATTAACACTAGTAACAG	Mutant construction
ef3023_2_IVRbis	CGATAACTAGAAGACATGAAAGCAACAACGACTATTA	Mutant construction
ef3023_3_IVRbis	CGTTGTTGCTTTTCATGCTTCTAGTTATCGCCAGT	Mutant construction
ef3023_4_IVR	ATGGTTGCTGGGTTTATCGACCTGCAGCTTTAATGTGTCGTTATCTGA	Mutant construction
ef3023_5	GTGGAAAGTAGCCAACTTC	Mutant construction
ef3023_6	TGGAATTGTGGGCTATAGC	Mutant construction
1516F_BamHI	ATTCGGATCCTATCCAAGCCATTAAAGGCT	Mutant construction
1516R_PstI	TTATCTGCAGAGACCGCATGTAACATATA	Mutant construction
1516_verifF	TATTATTTGCGGTTGGTTTAG	Mutant construction
1516_verifR	AACTTAAGAAGAAGTCTACAAA	Mutant construction
topo_FP2	GAATTCAGCACACTGG	pTOPO cloning
topo_RP2	CTATAGTGAGTCGATTACA	pTOPO cloning
topo65_FP1	TAATACGACTCACTATAGTATCATAATAAAGGAGGTGAAAT	pTOPO cloning
topo65_RP1	AGTGTGCTGGAATTCATTTATAATTTGTGTTATACTAAG	pTOPO cloning
topo85_FP1	TAATACGACTCACTATAGTTAACAATGAATAGCGTTTTTC	pTOPO cloning
topo85_RP1	AGTGTGCTGGAATTCATGACCCTCTTCTGCAGAA	pTOPO cloning
topo5'3023_FP1	TAATACGACTCACTATAGCGTAAACATGTATAATATTTTCATG	pTOPO cloning
topo5'3023_RP1	AGTGTGCTGGAATTCATCCGACTTTTCCAAACAC	pTOPO cloning
ef1515_pQE70_SphI	TACGCATGCGAATTAAGAGGTGCTAAATCAAAA	pQE70 cloning
ef1515_pQE70_BglII	ATGAGATCTTTGGTCATTCTTAATCGTTCA	pQE70 cloning



**Annexe 3 : Résultats complémentaires de la publication 3 « Study of RNA metabolism  
actors in *Enterococcus faecalis* »**

**Table S1** : List and sequences of oligonucleotides used in this study.

Primer name	Sequence	Use
1185-1	AAATGGATCCGCTTCGTGTACCTG	$\Delta$ rnjB construction
1185-2	ATTCGAATTCACAATCCAGTACAAAA	$\Delta$ rnjB construction
1185-3	TTGTGGAATTCGAATGGAGCAAACCTG	$\Delta$ rnjB construction
1185-4	ATATCCCGGGAAGTAACCCGCAATTT	$\Delta$ rnjB construction
1185-5	TCCAAATTGCTGAAACGCTAC	$\Delta$ rnjB construction
1185-6	AAGGTTTCAGGGATTGTGCA	$\Delta$ rnjB construction
2924-1	TTGATGGATCCATAATTGGGAAAACTGA	$\Delta$ rnjA construction
2924-2	TTCATGAATTCTACCCCATATGTGTTTT	$\Delta$ rnjA construction
2924-3	GTAGAATTCATGAAGGCCAACGTCT	$\Delta$ rnjA construction
2924-4	TGAACCCGGGTGTCGAATTTATCAAAGT	$\Delta$ rnjA construction
2924-5	TTTTAATGATTTACTTATACGATTA	$\Delta$ rnjA construction
2924-6	AAACAGTTTAAGCAACTTTCTTA	$\Delta$ rnjA construction
3097-1	TAGTGGATCCAATACCGTTGTTAAAATATT	$\Delta$ rnC construction
3097-2	CAATGAATTCGTAAGCTTGTCTAAT	$\Delta$ rnC construction
3097-3	TTTACGAATTCATTGGGTAGGCCAA	$\Delta$ rnC construction
3097-4	TGTTTCCCGGGTCTGCAATCTTCTCGA	$\Delta$ rnC construction
3097-5	TTCTGTTCGTTTCATATCTTTTG	$\Delta$ rnC construction
3097-6	TTTTTCTGCGATTGCTGTTTC	$\Delta$ rnC construction
3170-1	GGATCCTGCGTTAGTTCCTC	$\Delta$ rnY construction
3170-2	TTTTGCGAATTCCTAAAACCATATTTCCAC	$\Delta$ rnY construction
3170-3	TTTTAAGAATTCGCAAACTAGCTGGC	$\Delta$ rnY construction
3170-4	ATATCCCGGGTGTCCACGTATGGT	$\Delta$ rnY construction
3170-5	AAATGGTGGAACAGCCGCT	$\Delta$ rnY construction
3170-6	AAATTAATCACCGTAGTCCC	$\Delta$ rnY construction
0846-1	CATGGATCCTGCTCAGGAAA	$\Delta$ cshA construction
0846-2	CACGAATTCACCTAGTTTCAGGT	$\Delta$ cshA construction, pNZ273 C5 construction
0846-3	GTAGAATTCGTGGCAGCTATCGTG	$\Delta$ cshA construction, pNZ273 C5 construction
0846-4	ATACCCGGGCGCTTTTACAACCGC	$\Delta$ cshA construction,
0846-5	TTACACATGCCTGATGAAACAG	$\Delta$ cshA construction
0846-6	ATCCAATAAGGCGACACAAAAA	$\Delta$ cshA construction
1009-1	ATTGGATCCAATTGGATTGACAT	$\Delta$ cshC construction
1009-2	GAAGAATTCGCCACTGTTCTTGCCA	$\Delta$ cshC construction
1009-3	GCAGAATTCGTTGTCAAAGAACCAC	$\Delta$ cshC construction
1009-4	AGTCCCGGGTCACCTGACCAA	$\Delta$ cshC construction
1009-5	ATTACGTGACTTTTGTACGAGA	$\Delta$ cshC construction
1009-6	TAACCAAATGTTCTAAATTTGTGC	$\Delta$ cshC construction
1377-1	TTTGGATCCGAGCAATGATTTTGT	$\Delta$ cshB construction
1377-2	TTGCCCGGGTTATGCGAGTGCTTC	$\Delta$ cshB construction
1377-3	TAACCCGGGCAATTGAAGACAGCA	$\Delta$ cshB construction
1377-4	CTGCCATGGCAAGCGACAAATT	$\Delta$ cshB construction
1377-5	AACCAATTAATGGGCAATCATT	$\Delta$ cshB construction
1377-6	TTTCACTCGTTTACATACGAAA	$\Delta$ cshB construction
846TAPF1_Bam	ATCTGGATCCACTTCTATTTTCAGCAACTA	EF0846-TAP construction
846TAPR2_Bam	TTATGGATCCTTTGGCGTTTTCCGGAAT	EF0846-TAP construction
846TAPF3_Xmal	AAGTCCCGGGAATCAAAGAGCGAGTGTTCC	EF0846-TAP construction
846TAPR4_Ncol	TTCTCCATGGATTGTACATCCTTCTTTCTTC	EF0846-TAP construction
CBPF	ATTGGATCCATGGAAAAGAGAAG	EF0846-TAP construction
PROTAR	TCACCCGGGTACCAGATC	EF0846-TAP construction
P0846_FP1	TATCTTAATTTTTAAAGGGAAGTAACGG	pNZ273 C3, C4, C5 constructions
P0846_RP2	CCTTTAAAAATTAAGATAATAATACTGGTC	pNZ273 C3, C4, C5 constructions
P0846_FP2	TCATTTGTTACGTCTGTAGAAACC	pNZ273 C3 construction
P0846_RP1	AGGACGTAACAAATGATATCCTCCTAGA	pNZ273 C3 construction

P0846_FP2_C4_5	AATTCGGTTACGTCCTGTAGAAACC	pNZ273 C4, C5 constructions
P0846_RP1_C4_5	AGGACGTAACGGAATTGTTTCTGATTGAAT	pNZ273 C4, C5 constructions
topo_FP2	GAATTCACGACACTGG	pTOPO Cloning
topo_RP2	CTATAGTGAGTCGTATTACA	pTOPO Cloning
topo65_FP1	TAATACGACTCACTATAGTATCATAATAAAGGAGGTGAAAT	pTOPO Cloning
topo65_RP1	AGTGTGCTGGAATTC AATTTATAATTTGTGTTATACTAAG	pTOPO Cloning
topo0846_FP1	TAATACGACTCTATATTTAAAGGGAAGTAACG	pTOPO cloning
topo0846_RP1	AGTGTGCTGGAATTCGAATTTCTGATTGAATAG	pTOPO cloning
pNZ273_MCS_for	ATGAGATAATGCCGACTGTAC	pNZ273 cloning
pNZ273_MCS_rev	CGATCCAGACTGAATGCC	pNZ273 cloning
pQE-846-F	ATTATGGATCCTTGAAATTTAAAGAACTTGGTTTA	pQE30 cloning
pQE-846-R	ATATACTGCAGTTATTTGGCGTTTTTCGCGAAT	pQE30 cloning
pAgent0846F	TCACACCATGGAAGTAAACATAAAAAATAACGAAA	pAGEnt cloning
pAgent0846R	ACACACTGCAGTTCTCTATTAAATGAGAGCTTTA	pAGEnt cloning
pUCB0846F	TCACACTGCAGGTCACGAAGACCATATTGGT	pUCB300 cloning
pUCB0846R	ACACACCATGGATTCACCATGAACAGGCATG	pUCB300 cloning
SRC43_SP1	AATTCACCTTATGGCATTAGAA	5'RACE
SRC43_SP2	TGGTCGTTTCAGCTTCACATT	5'RACE
SRC43_SP3	TTAAACATAGAAATAGATCTATTTT	5'RACE
ef1377-SP1	ATAGTTGGCTTAACCTTTATCCA	5'RACE
ef1377-SP2	TTCGTTGATAAAAAGTTGAAATT	5'RACE
ef0846_SP2	ACTAATCCTTGAATAACGCGA	5'RACE
ef0846_SP3	TTGTTAAGCATTGGTAAGCCA	5'RACE
gusA-G	GTGTGGGTCAATAATCAGG	qPCR
gusA-D	GAAATCATGGAAGTAAGACTGC	qPCR
ef0846L2	AACCCGTGAGTTGGCTATCC	qPCR
ef0846R2	GACCGATATCAGCCCCACC	qPCR, 5'RACE (SP1), Northern blot
1009L	TGTCGTGAATCAAGCGCAATC	qPCR
1009R	ACCGAATAAAGTAGTGGGCGA	qPCR
1185L	TGGCATGTGTGGCAAGTAAT	qPCR
1185R	CGTTCAAAATCTTGGCCTGT	qPCR
1377L	GCAACCTCATGTGGTGATTG	qPCR
1377R	TCATCGACAACAAAGGCAAA	qPCR
2924L	CATTGTCGCAACAGGTGACT	qPCR
2924R	GACAAAGGACGCCTTCTTCA	qPCR
GyrAR	TCTTTTCCATTTCGGCATTTC	qPCR
GyrAL	GATGGGGAAATCAGGGATTTC	qPCR
SRC43_probe	TTTATTACTACGATTCAATTTTCTCGTCATCTTCCGGC	Northern blot

**Table S2 :** Conservation of helicases (CshA, CshB and CshC) and RNases (J1, J2, Y and III) in enterococci strains and non-enterococci strains (sheet “overview”) based on UniProt database analysis. Proteins conservation is symbolized with colored boxes. Other sheets list proteins with high homologies with CshA, CshB, CshC, RnjA, RnjB, Rny and Rnc (E-value, score and identity % and sequences are indicated).

**A**

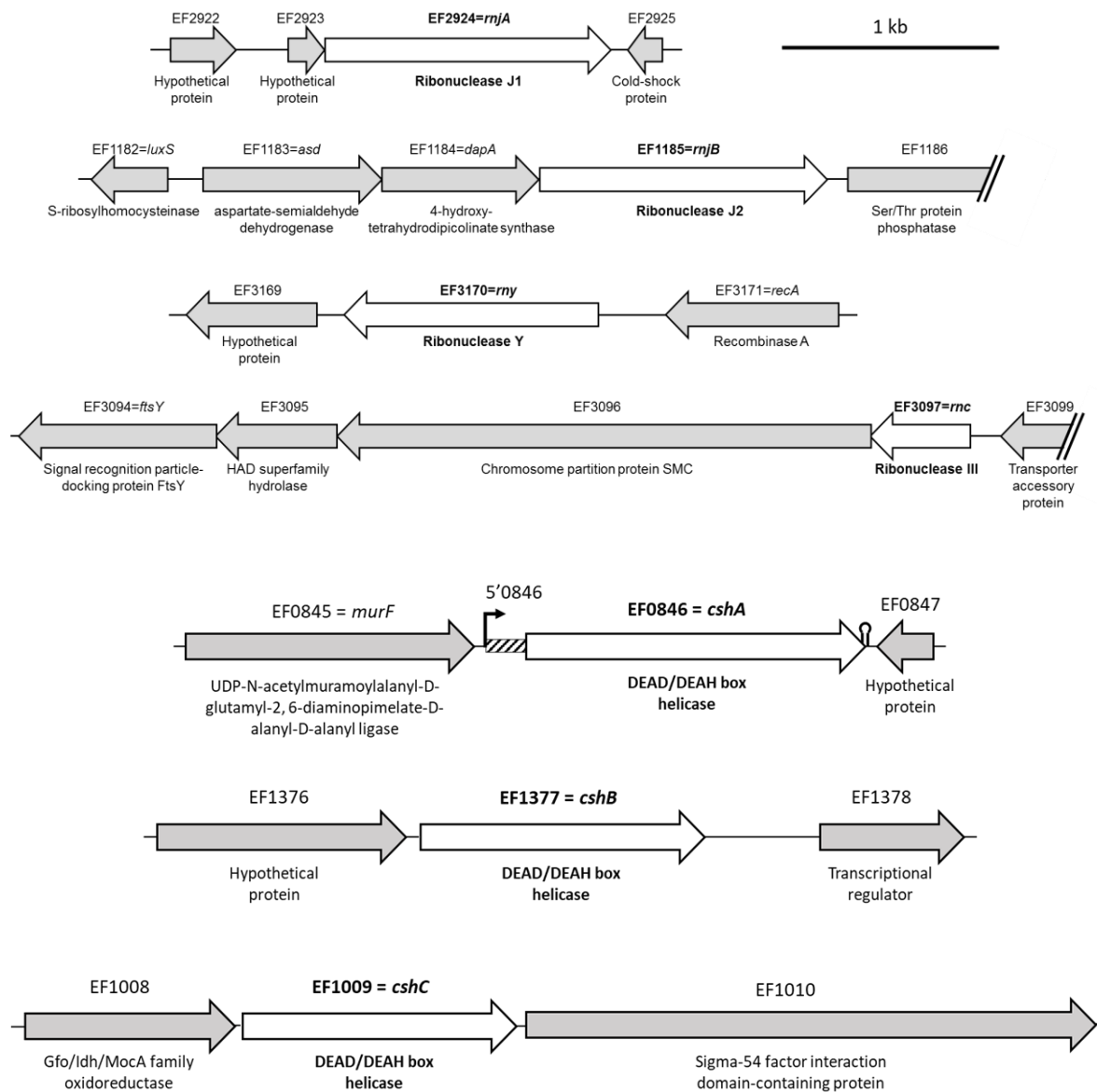
Strains of Enterococci	Helicases Csh			RNases				Strains of Enterococci	Helicases Csh			RNases			
	A	B	C	J1	J2	Y	III		A	B	C	J1	J2	Y	III
<i>E. aquimarinus</i>								<i>E. faecium</i> DO							
<i>E. asini</i> ATCC 700915								<i>E. faecium</i> E1679							
<i>E. avium</i> ( <i>Streptococcus avium</i> )								<i>E. faecium</i> E417							
<i>E. cacciae</i> ATCC BAA-1240								<i>E. faecium</i> E980							
<i>E. canintestini</i>								<i>E. faecium</i> E980							
<i>E. canis</i>								<i>E. faecium</i> EnGen0020							
<i>E. casseliflavus</i> ( <i>E. flavescens</i> )								<i>E. faecium</i> EnGen0026							
<i>E. cecorum</i>								<i>E. faecium</i> EnGen0153							
<i>E. columbae</i> DSM 7374 = ATCC 51263								<i>E. faecium</i> EnGen0191							
<i>E. devriesei</i>								<i>E. faecium</i> EnGen0263							
<i>E. dispar</i> ATCC 51266								<i>E. faecium</i> OC2A-1							
<i>E. durans</i>								<i>E. faecium</i> R496							
<i>E. faecalis</i> (strain ATCC 700802 / V583)								<i>E. faecium</i> SB2C-2							
<i>E. faecalis</i> ( <i>Streptococcus faecalis</i> )								<i>E. faecium</i> SD3B-2							
<i>E. faecalis</i> 02-MB-BW-10								<i>E. faecium</i> T110							
<i>E. faecalis</i> 06-MB-DW-09								<i>E. faecium</i> TX1337RF							
<i>E. faecalis</i> 13-SD-W-01								<i>E. flavescens</i> ATCC 49996							
<i>E. faecalis</i> 918								<i>E. gallinarum</i>							
<i>E. faecalis</i> ATCC 6055								<i>E. gilvus</i> ATCC BAA-350							
<i>E. faecalis</i> CBRD01								<i>E. haemoperoxidus</i> ATCC BAA-382							
<i>E. faecalis</i> CH188								<i>E. hermanniensis</i>							
<i>E. faecalis</i> DAPTO 512								<i>E. hirae</i>							
<i>E. faecalis</i> DAPTO 516								<i>E. italicus</i> DSM 15952							
<i>E. faecalis</i> EnGen0302								<i>E. malodoratus</i> ATCC 43197							
<i>E. faecalis</i> EnGen0311								<i>E. moraviensis</i> ATCC BAA-383							
<i>E. faecalis</i> EnGen0354								<i>E. mundtii</i>							
<i>E. faecalis</i> EnGen0359								<i>E. pallens</i> ATCC BAA-351							
<i>E. faecalis</i> EnGen0426								<i>E. phoeniculicola</i> ATCC BAA-412							
<i>E. faecalis</i> ERV103								<i>E. plantarum</i>							
<i>E. faecalis</i> ERV116								<i>E. quebecensis</i>							
<i>E. faecalis</i> ERV129								<i>E. raffinosus</i>							
<i>E. faecalis</i> ERV25								<i>E. ratti</i>							
<i>E. faecalis</i> ERV31								<i>E. rivorum</i>							
<i>E. faecalis</i> ERV37								<i>E. rotai</i>							
<i>E. faecalis</i> ERV41								<i>E. saccharolyticus</i>							
<i>E. faecalis</i> ERV62								<i>E. silesiacus</i>							
<i>E. faecalis</i> ERV63								<i>E. sp.</i> (strain 3G1_DIV0629)							
<i>E. faecalis</i> ERV65								<i>E. sp.</i> 10A9_DIV0425							
<i>E. faecalis</i> ERV68								<i>E. sp.</i> 10F3_DIV0382							
<i>E. faecalis</i> ERV72								<i>E. sp.</i> 12C11_DIV0727							
<i>E. faecalis</i> ERV73								<i>E. sp.</i> 12E11_DIV0728							
<i>E. faecalis</i> ERV81								<i>E. sp.</i> 12F9_DIV0723							
<i>E. faecalis</i> ERV85								<i>E. sp.</i> 2F9_DIV0599							
<i>E. faecalis</i> ERV93								<i>E. sp.</i> 2G9_DIV0600							
<i>E. faecalis</i> Merz96								<i>E. sp.</i> 3G6_DIV0642							
<i>E. faecalis</i> PF3								<i>E. sp.</i> 3H8_DIV0648							
<i>E. faecalis</i> R712								<i>E. sp.</i> 4E1_DIV0656							
<i>E. faecalis</i> S613								<i>E. sp.</i> 4G2_DIV0659							
<i>E. faecalis</i> T2								<i>E. sp.</i> 5B3_DIV0040							
<i>E. faecalis</i> TX0027								<i>E. sp.</i> 6C8_DIV0013							
<i>E. faecalis</i> TX0104								<i>E. sp.</i> 6D12_DIV0197							
<i>E. faecalis</i> TX0309A								<i>E. sp.</i> 7D2_DIV0200							
<i>E. faecalis</i> TX0309B								<i>E. sp.</i> 7E2_DIV0204							
<i>E. faecalis</i> TX0411								<i>E. sp.</i> 7F3_DIV0205							
<i>E. faecalis</i> TX0630								<i>E. sp.</i> 8G7_MSG3316							
<i>E. faecalis</i> TX0645								<i>E. sp.</i> 9D6_DIV0238							
<i>E. faecalis</i> TX4248								<i>E. sp.</i> 9E7_DIV0242							
<i>E. faecalis</i> VC1B-1								<i>E. sp.</i> FDAARGOS_375							
<i>E. faecalis</i> WKS-26-18-2								<i>E. sp.</i> HSI EG1							
<i>E. faecium</i> 1,141,733								<i>E. sp.</i> RIT-PI-f							
<i>E. faecium</i> 1,231,408								<i>E. sulfureus</i> ATCC 49903							
<i>E. faecium</i> 1,231,501								<i>E. termitis</i>							
<i>E. faecium</i> 10/96A								<i>E. thailandicus</i>							
<i>E. faecium</i> 13.SD.W.09								<i>E. ureasiticus</i>							
<i>E. faecium</i> 505								<i>E. ureilyticus</i>							
<i>E. faecium</i> Com15								<i>E. villorum</i>							
<i>E. faecium</i> CRL1879															

**B**

Strains (other than Enterococci)	Helicases Csh			RNases			Strains (other than Enterococci)	Helicases Csh			RNases		
	A	B	C	J1	J2	Y		A	B	C	J1	J2	Y
<i>Bacillus</i> sp.							<i>Lactobacillus uvarum</i> DSM 19971						
<i>Carnobacterium alterfunditum</i>							<i>Lactobacillus wasatchensis</i>						
<i>Carnobacterium divergens</i> ( <i>Lactobacillus divergens</i> )							<i>Lactobacillus zeae</i> DSM 20178 = KCTC 3804						
<i>Carnobacterium iners</i>							<i>Listeria aquatica</i> FSL S10-1188						
<i>Carnobacterium inhibens</i> subsp. <i>gilchinskyi</i>							<i>Listeria booriae</i>						
<i>Carnobacterium maltaromaticum</i> DSM 20342							<i>Listeria cornellensis</i> FSL F6-0969						
<i>Carnobacterium maltaromaticum</i> LMA28							<i>Listeria fleischmannii</i> 1991						
<i>Carnobacterium</i> sp.							<i>Listeria floridensis</i> FSL S10-1187						
<i>Catelicoccus marimammalium</i> M35/04/3							<i>Listeria grandensis</i> FSL F6-0971						
<i>Clostridioides difficile</i> Y384							<i>Listeria grayi</i> DSM 20601						
<i>Desemzia incerta</i>							<i>Listeria innocua</i> ATCC 33091						
<i>Floricoccus penangensis</i>							<i>Listeria ivanovii</i> (strain ATCC BAA-678 / PAM 55)						
<i>Floricoccus tropicus</i>							<i>Listeria monocytogenes</i>						
<i>Isobaculum melis</i>							<i>Listeria</i> sp.						
<i>Granulicatella balaenopterae</i>							<i>Listeriaceae</i> bacterium FSL A5-0209						
<i>Granulicatella elegans</i> ATCC 700633							<i>Marinilactibacillus piezotolerans</i>						
<i>Isobaculum melis</i>							<i>Melissococcus plutonius</i>						
<i>Klebsiella michiganensis</i>							<i>Mycobacteroides abscessus</i>						
<i>Lactobacillus agilis</i>							<i>Pilibacter termitis</i>						
<i>Lactobacillus algidus</i> DSM 15638							<i>Pisciglobus halotolerans</i>						
<i>Lactobacillus animalis</i>							<i>Planococcus donghaensis</i>						
<i>Lactobacillus apodemii</i> DSM 16634 = JCM 16172							<i>Planococcus massiliensis</i>						
<i>Lactobacillus aquaticus</i> DSM 21051							<i>Planococcus plakortidis</i>						
<i>Lactobacillus aviarius</i>							<i>Planococcus rifietoensis</i>						
<i>Lactobacillus backii</i>							<i>Planococcus</i> sp. Urea-3u-39						
<i>Lactobacillus brantae</i> DSM 23927							<i>Planomicrobium glaciei</i>						
<i>Lactobacillus bifementans</i> DSM 20003							<i>Sedimentibacter</i> sp. SX930						
<i>Lactobacillus cacaonum</i> DSM 21116							<i>Sphingobacterium faecium</i> PCAi_F2.5						
<i>Lactobacillus casei</i>							<i>Streptococcus acidominimus</i>						
<i>Lactobacillus capillatus</i> DSM 19910							<i>Streptococcus agalactiae</i>						
<i>Lactobacillus collinoides</i>							<i>Sphingobacterium faecium</i> PCAi_F2.5						
<i>Lactobacillus coryniformis</i> subsp. <i>Coryniformis</i>							<i>Streptococcus bovimastitidis</i>						
<i>Lactobacillus curvatus</i>							<i>Streptococcus cristatus</i>						
<i>Lactobacillus fuchuensis</i>							<i>Streptococcus cuniculi</i>						
<i>Lactobacillus graminis</i> DSM 20719							<i>Streptococcus dysgalactiae</i> subsp. <i>dysgalactiae</i> ATCC 27957						
<i>Lactobacillus hayakitensis</i> DSM 18933 = JCM 14209							<i>Streptococcus dysgalactiae</i> subsp. <i>Equisimilis</i>						
<i>Lactobacillus helsingborgensis</i>							<i>Streptococcus equinus</i> ( <i>Streptococcus bovis</i> )						
<i>Lactobacillus hokkaidonensis</i> JCM 18461							<i>Streptococcus henryi</i>						
<i>Lactobacillus malefermentans</i> DSM 5705 = KCTC 3548							<i>Streptococcus ictaluri</i> 707-05						
<i>Lactobacillus mali</i>							<i>Streptococcus mitis</i>						
<i>Lactobacillus murinus</i>							<i>Streptococcus oralis</i> SK100						
<i>Lactobacillus oeni</i> DSM 19972							<i>Streptococcus phocae</i>						
<i>Lactobacillus oligofermentans</i> DSM 15707 = LMG 22743							<i>Streptococcus pneumoniae</i>						
<i>Lactobacillus oryzae</i> JCM 18671							<i>Streptococcus porcinus</i> str. <i>Jelinkova</i> 176						
<i>Lactobacillus ozensis</i> DSM 23829 = JCM 17196							<i>Streptococcus pseudoporcinus</i> LQ 940-04						
<i>Lactobacillus paracasei</i>							<i>Streptococcus salivarius</i>						
<i>Lactobacillus paracollinoides</i>							<i>Streptococcus</i> sp. ACS2						
<i>Lactobacillus paraplantarum</i>							<i>Streptococcus</i> sp. FDAARGOS_256						
<i>Lactobacillus pentosus</i>							<i>Streptococcus</i> sp. M334						
<i>Lactobacillus plantarum</i>							<i>Streptococcus suis</i>						
<i>Lactobacillus rennini</i> DSM 20253							<i>Streptococcus uberis</i> (strain ATCC BAA-854 / 0140J)						
<i>Lactobacillus rhamnosus</i>							<i>Streptococcus vestibularis</i> ATCC 49124						
<i>Lactobacillus ruminis</i>							<i>Tetragenococcus halophilus</i>						
<i>Lactobacillus sakei</i>							<i>Tetragenococcus muraticus</i>						
<i>Lactobacillus salivarius</i>							<i>Trichococcus collinsii</i>						
<i>Lactobacillus saniviri</i> JCM 17471 = DSM 24301							<i>Trichococcus flocculiformis</i>						
<i>Lactobacillus satsumensis</i> DSM 16230 = JCM 12392							<i>Trichococcus ilyis</i>						
<i>Lactobacillus selangorensis</i>							<i>Trichococcus pasteurii</i>						
<i>Lactobacillus shenzhenensis</i> LY-73							<i>Trichococcus</i> sp. ES5						
<i>Lactobacillus</i> sp. HMSC056D05							<i>Vagococcus fluvialis</i> bH819						
<i>Lactobacillus</i> sp. HMSC25A02							<i>Vagococcus lutrae</i> LBD1						
<i>Lactobacillus</i> sp. UMNPBX2							<i>Vagococcus martis</i>						
<i>Lactobacillus</i> sp. UMNPBX9							<i>Vagococcus penaei</i>						
<i>Lactobacillus sucicola</i> DSM 21376 = JCM 15457							<i>Vagococcus teuberi</i>						
<i>Lactobacillus suebicus</i> DSM 5007 = KCTC 3549													

**Remarque :** Les tableaux présentés ici constituent un résumé du tableau S2. Etant donné la quantité de données qui constituent l'analyse détaillée (séquences des protéines, pourcentages d'identité,...), nous avons choisi de ne pas le présenter intégralement ici. Il sera néanmoins disponible dans l'article après publication.

**Figure S1 :** Genomic environment of genes encoding the RNases and Helicases studied in *E. faecalis* V19.



## RÉSUMÉ

### **Identification et caractérisation d'ARN régulateurs impliqués dans la réponse au stress et la virulence chez *Enterococcus faecalis*.**

*E. faecalis* est une bactérie à gram positif, responsable d'infections nosocomiales. Dans cette étude, nous avons identifié par RNA-seq 65 ARN régulateurs potentiels induits en conditions de stress et/ou d'infection chez ce pathogène opportuniste. Parmi ceux-ci, trois ARN surexprimés *in vivo*, en présence de sels biliaires et à pH acide ont fait l'objet d'une étude approfondie.

Le premier, SRC65, s'est avéré être un petit ARN (sARN) agissant probablement en *trans*. Il présenterait des fonctions redondantes avec son homologue SRC90. Différentes cibles potentielles ont été identifiées et des expériences de physiologie ont révélé un rôle de SRC65 dans le déclenchement de la phase exponentielle de croissance.

Le 2<sup>ème</sup> ARN étudié est une région 5' régulatrice non traduite (5'UTR), appelée 5'1515. Elle formerait un atténuateur et agirait de manière répressive sur le gène *ef1515*. EF1515 est un antiterminateur de la famille BglG/SacY capable de se fixer sur 5'1515 pour réguler son expression et celle du gène *ef1516* localisé en aval et codant un système de transport des sucres de type PTS. L'antiterminateur EF1515 contrôle aussi l'expression du gène *ef3023* codant une protéine impliquée dans la virulence d'*E. faecalis*.

Le dernier ARN régulateur étudié est également une 5'UTR. Celle-ci participerait à la régulation d'une hélicase à motif DEAD (CshA) codée en aval de la 5'UTR. Sa caractérisation s'inscrit dans une étude plus large concernant les éléments du métabolisme des ARN, impliquant les ribonucléases et les autres hélicases à motif DEAD d'*E. faecalis*. CshA aurait un rôle prépondérant pour la bactérie, en étant impliquée dans la réponse au stress, le fitness et la virulence. L'identification de l'interactome de CshA a notamment permis d'identifier l'énolase comme partenaire privilégié.

Mots-clés : *Enterococcus faecalis*, sARN, 5'UTR, hélicase à motif DEAD, antiterminateur, virulence, stress

## ABSTRACT

### **Identification and characterization of regulatory RNA involved in stress response and virulence in *Enterococcus faecalis*.**

*E. faecalis* is a gram-positive bacterium responsible for nosocomial infections. Using RNA-seq, we identified 65 putative regulatory RNA induced under stress and/or during infection. Of these, three RNA overexpressed *in vivo*, in the presence of bile salts and at acidic pH have were more deeply studied.

The first, SRC65, was found to be a small RNA (sRNA) probably acting in *trans*. It would present redundant functions with its homologous sRNA SRC90. Different potential targets were identified and physiology experiments revealed a role for SRC65 in triggering the exponential growth phase.

The 2nd studied RNA is a 5' untranslated region (5'UTR), called 5'1515 which would form an attenuator and act repressively on the *ef1515* gene. EF1515 is an antiterminator of the BglG/SacY family capable of binding at 5'1515 to regulate its expression and that of the downstream gene *ef1516* encoding a PTS-type sugar transporter. The EF1515 antiterminator also controls the expression of the *ef3023* gene encoding a protein involved in *E. faecalis* virulence.

The last regulatory RNA studied is also a 5'UTR. It would participate in the regulation of a DEAD-box helicase (CshA) encoded downstream of the 5'UTR. Its characterization is part of a broader study of the elements of RNA metabolism, involving ribonucleases and other DEAD-box helicases of *E. faecalis*. CshA would have a prominent role for the bacteria, being involved in stress response, fitness and virulence. The identification of the CshA interactome made it possible to identify enolase as a preferred partner.

Keywords: *Enterococcus faecalis*, sRNA, 5'UTR, DEAD-box helicase, antiterminator, virulence, stress