

UNIVERSITÉ FRANÇOIS – RABELAIS DE TOURS

ÉCOLE DOCTORALE SSBCV

UMR1282 - Infectiologie et Santé Publique

Equipe 5 : Bactéries et Risque Materno-Fœtal

THÈSE présentée par :

Mohamed Amine ZORGANI

soutenue le : **07 décembre 2016**

pour obtenir le grade de : **Docteur de l'université François – Rabelais de Tours**

Discipline/ Spécialité : Biologie / Microbiologie

Caractérisation des ARN régulateurs chez *Streptococcus agalactiae*

THÈSE dirigée par :

Pr LARTIGUE Marie-Frédérique
Pr QUENTIN Roland

PU-PH, université François – Rabelais de Tours
PU-PH, université François – Rabelais de Tours

RAPPORTEURS :

Dr DRAMSI Shaynoor
Pr SOUTOURINA Olga

HDR, institut Pasteur, Paris
PU, Paris – Saclay

JURY :

Dr DRAMSI Shaynoor
Pr FELDEN Brice
Pr LARTIGUE Marie-Frédérique
Pr QUENTIN Roland
Pr SOUTOURINA Olga
Pr VANDENESCH François

HDR, institut Pasteur
PU, université de Rennes 1
PU-PH, université François – Rabelais de Tours
PU-PH, université François – Rabelais de Tours
PU, Paris – Saclay
PU-PH, université de Lyon

UNIVERSITÉ FRANÇOIS – RABELAIS DE TOURS

ÉCOLE DOCTORALE SSBCV

UMR1282 - Infectiologie et Santé Publique

Equipe 5 : Bactéries et Risque Materno-Fœtal

THÈSE présentée par :

Mohamed Amine ZORGANI

soutenue le : **07 décembre 2016**

pour obtenir le grade de : **Docteur de l'université François – Rabelais de Tours**

Discipline/ Spécialité : Biologie / Microbiologie

Caractérisation des ARN régulateurs chez *Streptococcus agalactiae*

THÈSE dirigée par :

Pr LARTIGUE Marie-Frédérique
Pr QUENTIN Roland

PU-PH, université François – Rabelais de Tours
PU-PH, université François – Rabelais de Tours

RAPPORTEURS :

Dr DRAMSI Shaynoor
Pr SOUTOURINA Olga

HDR, institut Pasteur – Paris
PU, Paris – Saclay

JURY :

Dr DRAMSI Shaynoor
Pr FELDEN Brice
Pr LARTIGUE Marie-Frédérique
Pr QUENTIN Roland
Pr SOUTOURINA Olga
Pr VANDENESCH François

MCU-HDR, institut Pasteur
PU, université de Rennes 1
PU-PH, université François – Rabelais de Tours
PU-PH, université François – Rabelais de Tours
PU, Paris – Saclay
PU-PH, université de Lyon

Remerciements

Tables des matières

INTRODUCTION GENERALE.....	1
REVUE DE LA LITTERATURE.....	3
I. <i>Streptococcus agalactiae</i>.....	6
A. Caractéristiques phénotypiques et génétiques	6
1. Classification du genre <i>Streptococcus</i>	6
2. Taxonomie	6
3. Organisation du génome	8
B. Caractéristiques morphologiques et structurales	10
1. Caractères morphologiques, culturels et biochimiques	10
2. La paroi bactérienne	12
3. Les acides lipotéichoïques	12
a- Biosynthèse des acides lipotéichoïques.....	14
b- La D-alanylation des acides lipotéichoïques	14
c- Rôle et fonction des acides lipotéichoïques.....	16
C. Résistance aux agents antimicrobiens.....	16
1. Classification et mécanismes d'action des antibiotiques	16
2. Mécanismes de résistance aux antibiotiques	18
3. Résistance aux PAM.....	19
a- Mode d'action des peptides antimicrobiens cationiques	19
b- Principaux mécanismes de résistance aux PAM chez les <i>Streptococcus</i>	21
II. Les infections à <i>Streptococcus agalactiae</i>.....	25
A. <i>Streptococcus agalactiae</i> : un pathogène opportuniste	25
1. Infection chez le bovin	25
2. Infection chez l'Homme	27
a- Infection chez le nouveau-né	27
b- Infection chez la femme enceinte.....	28
c- Infection chez l'adulte, en dehors de la femme enceinte.....	28
B. Facteurs de virulence de <i>Streptococcus agalactiae</i>	30
1. Facteurs d'adhérence et de colonisation.....	30
2. Facteurs d'invasion cellulaire.....	30
3. Echappement au système immunitaire	32
III. Les systèmes de régulation chez <i>Streptococcus agalactiae</i>.....	35
A. Systèmes de régulation à deux composants.....	37
1. Le système CovS/CovR	37
2. Le système CiaR/CiaH	38

3.	Le système DltS/DltR	39
4.	Le système RgfC/RgfA	39
B.	Régulateurs transcriptionnels	40
1.	Le régulateur MtaR	40
2.	Le régulateur RogB.....	40
3.	Le régulateur RovS	41
C.	Les facteurs sigma	41
D.	Les systèmes de signalisation de type eucaryote	42
IV.	Les ARN régulateurs bactériens.....	45
A.	Découverte et caractéristiques des ARN régulateurs.....	45
B.	Mécanismes et modes d'action	48
1.	Les ARN régulateurs agissant en <i>cis</i>	48
a-	Les <i>riboswitch</i>	50
b-	Les thermosenseurs	52
2.	Les ARN régulateurs agissant en trans ciblant les ARNm	52
a-	Les ARN codés en <i>cis</i> ou antisens.....	54
b-	Les ARN codés en <i>trans</i>	57
3.	Les ARN régulateurs ciblant les protéines	61
C.	Exemple de certains ARN régulateurs	63
1.	ARN régulateurs impliqués dans la virulence chez <i>Staphylococcus aureus</i>	63
1.	ARN régulateurs chez <i>Clostridium</i>	65
2.	ARN régulateurs chez <i>Enterococcus faecalis</i>	65
3.	ARN régulateurs impliqués dans la virulence de <i>Streptococcus</i>	65
a-	ARN régulateurs chez <i>Streptococcus pyogenes</i>	66
b-	ARN régulateurs chez <i>Streptococcus pneumoniae</i>	67
c-	ARN régulateurs chez <i>Streptococcus agalactiae</i>	67
	Objectifs de travail.....	69
A.	Introduction	71
B.	Données préliminaires	71
1.	Identification des ARNrég dans la souche NEM316.....	71
2.	Identification des cibles de CetR par microarray	75
C.	Objectifs du travail de thèse	76
	Résultats.....	77
	Chapitre 1 : Caractérisation de l'ARN régulateur CetR : rôle dans la résistance aux PAM et dans la virulence de <i>S. agalactiae</i>	79

A.	Introduction	79
B.	Caractéristiques de l'ARNrég CetR	79
C.	Rôle de CetR dans l'intégrité cellulaire de <i>S. agalactiae</i>	83
D.	Publication N° 1	84
E.	Résultats complémentaires	114
1.	CetR, code-t-il différents transcrits fonctionnels ?	114
2.	Caractérisation des différents transcrits de CetR	114
3.	DltR, le régulateur du régulateur ?	116
4.	L'expression de CetR est inhibée par le stress acide et activée par le stress salin	116
5.	CetR favorise l'adhérence aux protéines de la matrice extracellulaire	118
6.	Rôle de CetR dans la croissance en milieu chimiquement défini	120

Chapitre 2 : CetR, un ARNrég ayant un rôle pléiotropique chez *S. agalactiae* 121

A.	Introduction	121
B.	Identification des nouvelles cibles directes ou indirectes de l'ARNrég CetR	121
1.	Identification des cibles directes de CetR par approche <i>in silico</i>	121
2.	Identification des cibles de CetR par approche <i>in vivo</i>	123
a-	Identification de cibles potentielles de CetR par protéomique comparative	123
b-	Identification des cibles par microarray par qRT-PCR	136
3.	Synthèse de l'analyse <i>in silico</i> et <i>in vivo</i>	138

Chapitre 3 : Prévalence des ARNrég chez *S. agalactiae* 144

A.	Introduction	142
B.	Etude de la prévalence des ARNrég par PCR	142
C.	Etude de la prévalence des ARNrég par analyse <i>in silico</i>	144
D.	Synthèse de l'étude de la prévalence	144

Chapitre 4 : Etude d'autres ARNrég de *S. agalactiae* 150

A.	Introduction	148
B.	Le duplex Srn017-Srn018 est-il un système Toxine-Antitoxine ?	148
C.	Les ARNrég Srn015, Srn024 et Srn070 ont-ils le même mode d'action ?	150

Discussion et perspectives..... 153

A.	Distribution des ARN régulateurs chez <i>S. agalactiae</i>	152
B.	CetR, un long ARN régulateur qui a un rôle pléiotrope	154

C.	CetR, un régulateur global de virulence ?	155
D.	CetR, régulateur des protéines moonlighting chez <i>S. agalactiae</i> ?	157
E.	CetR, interagit-il avec des protéines ?	161
F.	CetR et la résistance aux antibiotiques : nouvelle cible thérapeutique ?	162
	Conclusion générale.....	166
	Références bibliographiques.....	168
	Annexes.....	187

Liste des Tableaux

Tableau 1 : Classification des principales familles d'antibiotiques en fonction de leur mécanisme d'action	15
Tableau 2 : Facteurs de virulence chez <i>Streptococcus agalactiae</i> d'après (Rajagopal, 2009)	29
Tableau 3 : Liste non-exhaustive d'ARNrég impliqués dans le contrôle de la virulence chez plusieurs bactéries pathogènes pour l'Homme, d'après (Reignier, 2010)	62
Tableau 4 : Les ARNrég étudiés chez les streptocoques d'après (Le Rhun and Charpentier, 2012)	64
Tableau 5 : Gènes différenciellement exprimés dans le mutant <i>cetR</i> induit (I) par rapport au mutant <i>cetR</i> non-induit (NI), évalués par microarray	74
Tableau 6 : Liste des 10 meilleurs ARNm cibles formant potentiellement un duplex avec CetR <i>in silico</i>	122
Tableau 7 : Nombre total de protéines identifiées par LC-MS/MS retrouvées dans les différents compartiments cellulaires des deux souches WT et $\Delta cetR$	124
Tableau 8 : Protéines exprimées différenciellement entre le WT et $\Delta cetR$ identifiées dans le cytoplasme et la membrane	126
Tableau 9 : Protéines exprimées différenciellement entre le WT et $\Delta cetR$ identifiées dans le cytoplasme et le surnageant	129
Tableau 10 : Protéines exprimées différenciellement entre le WT et $\Delta cetR$ identifiées dans la membrane et le surnageant	129
Tableau 11 : Protéines exprimées différenciellement entre le WT et $\Delta cetR$ identifiées uniquement dans le cytoplasme	131
Tableau 12 : Protéines exprimées différenciellement entre la WT et $\Delta cetR$ identifiées uniquement dans la fraction membranaire	133
Tableau 12 : Suite	135
Tableau 13 : Protéines exprimées différenciellement entre le WT et $\Delta cetR$ identifiées uniquement dans le surnageant	135
Tableau 14 : Gènes différenciellement exprimés dans le mutant <i>cetR</i> induit (I) par rapport au mutant <i>cetR</i> non-induit (NI), évalués par microarray et confirmés par qRT-PCR	137
Tableau 15 : Cibles d'ARNm prédites par bioinformatique et identifiées par LC-MS/MS	139
Tableau 16 : Cibles identifiées par LC/MS-MS et par microarray	139
Tableau 17 : Répartition des souches de <i>S. agalactiae</i> étudiées par PCR en fonction de leur origine, de leur ST et du sérotypage capsulaire	141
Tableau 18 : Prévalence obtenue par analyse <i>in silico</i> des sept ARNrég de <i>S. agalactiae</i>	143
Tableau 19 : Prévalence des six ARNrég dans une population de 374 souches de <i>S. agalactiae</i> étudiées par PCR et analyse <i>in silico</i>	145
Tableau 20 : ARNm cibles de Srn015 prédites <i>in silico</i>	149

Liste des figures

Figure 1 : Arbre phylogénétique du genre <i>Streptococcus</i> basé sur les séquences d'ARNr 16S.	5
Figure 2 : La résolution taxonomique relative de diverses techniques moléculaires.	7
Figure 3 : Caractéristiques morphologiques et culturelles de <i>Streptococcus agalactiae</i>	9
Figure 4 : Représentation schématique de la paroi des bactéries à Gram positif.	11
Figure 5 : Représentation schématique de la structure des acides lipotéichoïques (ALT) et leur ancrage à la membrane des bactéries à Gram positif.	13
Figure 6 : Modèles de mécanismes de D-alanylation des acides lipotéichoïques (ALT).	13
Figure 7 : Rôles et fonctions des acides lipotéichoïques.....	15
Figure 8 : Illustration des divers mécanismes de résistance aux antibiotiques utilisés par les bactéries.....	17
Figure 9 : Principaux mécanismes de résistance aux PAM chez les <i>Streptococcus</i>	20
Figure 10 : <i>Streptococcus agalactiae</i> , un pathogène du nouveau-né.....	24
Figure 11 : Schéma représentatif des différentes entités d'infections materno-fœtale et leur transmission.	26
Figure 12 : Mécanismes d'adhérence et d'invasion cellulaire chez <i>Streptococcus agalactiae</i>	31
Figure 13 : Mécanismes d'échappement au système immunitaire chez <i>Streptococcus agalactiae</i>	31
Figure 14 : Principaux régulateurs de facteurs de virulence chez <i>Streptococcus agalactiae</i>	34
Figure 15 : Structure des systèmes de régulation à deux composants (TCS).	36
Figure 16 : Chronologie des études marquantes et les innovations technologiques dans l'étude des ARN régulateurs bactériens.	44
Figure 17 : Critères de prédiction des ARNrég par bio-informatique.	46
Figure 18 : Vue d'ensemble des méthodes à haut débit pour l'identification et la caractérisation des ARNrég bactériens.....	46
Figure 19 : Mécanisme d'action des <i>riboswitch</i>	49
Figure 20 : ARNrég agissant en <i>trans</i>	51
Figure 21 : Mécanisme d'action des ARN codés en <i>cis</i> ou antisens.....	53
Figure 22 : Mécanisme d'action d'ARN codés en <i>trans</i>	56
Figure 23 : Activation de la traduction par les ARN codés en <i>trans</i>	58
Figure 24 : Mode d'action de la protéine Hfq.....	58
Figure 25 : Représentation schématique des voies de dégradation des ARN chez les bactéries à Gram positif.	60
Figure 26 : Propriétés générales des sRNA séquestrant des protéines.	60

Figure 27 : Profil d'expression northern blot de sept ARNrég de <i>S. agalactiae</i> NEM316 en fonction de la phase de croissance.....	70
Figure 28 : Profil d'expression northern blot de sept ARNrég de <i>S. agalactiae</i> NEM316 en fonction des conditions de stress.	72
Figure 29 : Mécanisme d'action du vecteur d'expression pG+off portant un promoteur inducible par la tétracycline.	74
Figure 30 : Identification et validation de l'expression de l'ARNrég CetR.	78
Figure 31 : Structure secondaire potentielle de l'ARNrég CetR.	80
Figure 32 : Diagramme représentant l'environnement génétique de <i>cetR</i>	82
Figure 33 : Rôle de CetR dans la physiologie de <i>S. agalactiae</i>	82
Figure 34 : Profil d'expression de CetR.	113
Figure 35 : Caractérisation de la région d'interaction CetR- <i>dltD</i>	113
Figure 36 : Région homologue entre les promoteurs de <i>dltR</i> et <i>cetR</i>	115
Figure 37 : Effet de stress acide et salin sur l'ARNrég CetR.	117
Figure 38 : CetR favorise l'adhérence aux protéines de la matrice extracellulaire.....	117
Figure 39 : Rôle de CetR dans la croissance en milieu chimiquement défini.....	119
Figure 40 : Relations phylogénétiques entre les 125 souches de <i>S. agalactiae</i> typées par <i>Multi-Locus Sequence Typing</i> (MLST) et prévalence des six ARNrég étudiée par PCR.	141
Figure 41 : Diagramme représentant les caractéristiques du duplex Srn017-Srn018.	147
Figure 42 : Structures secondaires des csRNA de la souche <i>S. agalactiae</i> NEM316.....	149
Figure 43 : La voie de glycolyse et l'activité <i>moonlighting</i> des protéines glycolytiques bactériennes.	159

Liste des annexes

A.	Annexe 1 : Méthodes supplémentaires	189
1.	Cloning of Transcriptional Gene Reporter to cetR Putative Promoters.....	189
2.	β -Galactosidase Assays	189
3.	Determination of 5' ends of CetR by RACE PCR	189
4.	Northern Blot Analysis	190
5.	Coating of microtiter plates with various ECM proteins.....	190
6.	Adhesion test to ECM proteins	190
B.	Annexe 2 : Rôle de CetR dans la croissance en MCD.....	192
C.	Annexe 3 : Méthodes de l'approche protéomique	195
a-	Bacterial strains and growth conditions	195
b-	Protein preparation.....	195
c-	Protein identification by nano-LC-MS/MS	196
D.	Annexe 4: Etude de la prévalence des ARNrég chez <i>S. agalactiae</i>	199
1.	Amplification génique par PCR	199
2.	Etude de la prévalence par analyse <i>in silico</i>	199
E.	Annexe 5 : ARNm cibles de Srn024 prédites <i>in silico</i>.....	200
F.	Annexe 6 : ARNm cibles de Srn070 prédites <i>in silico</i>.....	201
G.	Annexe 7 : Motifs conservés entre les séquences codant les ARNrég Srn015, Srn024 et Srn070	202
H.	Annexe 8 : Publication 3.....	203
I.	Annexe 9 : Publication 4.....	213
J.	Annexe 10 : Publication 5.....	217
K.	Annexe 11 : Publication 6.....	273
L.	Annexe 12 : Publication 7.....	288

Liste des acronymes

3' ou 5'UTR : 3' ou 5' untranslated region

AA : Acide aminé

ADN : Acide désoxyribonucléique

ADNg : ADN génomique

ALT : Acide lipotéichoïque

ARN : Acide ribonucléique

ARNm : ARN messager

ARNr : ARN ribosomique

AENrég : ARN régulateur

ARNt : ARN de transfert

BHI : « Brain Heart Infusion »

CC : Complexe Clonal

CetR : « cell envelop targeting RNA »

DMEM : Dulbecco's modified Eagle Medium

IGR : région intergénique

LB : « Luria Bertani broth »

LPS : Lipopolysaccharide

MCD : Milieu Chimiquement Défini

miARN: micro-ARN

MLST : « MultiLocus Sequence Typing »

MOI : Multiplicity of Infection

nt : nucléotide

OMP : outer membrane protein

ORF : Open Reading Frame

pb : paire de bases

PCR : Polymerase chain reaction

RBS : « Ribosome Binding Site »

RT-PCR : « Reverse Transcriptase - Polymerase Chain Reaction »

SD : séquence Shine Dalgarno

SGB : Streptocoque du groupe B

sRNA: small RNA

SVF : Sérum de Veau Fœtal

TH : « Todd Hewitt »

TIR : Translation Initiation Region

TSS : Transcription start site

UFC : Unité Formant Colonie

INTRODUCTION GENERALE

Streptococcus agalactiae est une bactérie présente physiologiquement dans le tube digestif et le tractus génital de l'Homme. La plupart du temps, la présence de cette bactérie est sans danger, cependant dans certains cas, *S. agalactiae* peut également être responsable de graves infections chez l'Homme (méningite, pneumonie et septicémie), principalement chez le nouveau-né (c'est la première cause d'infections néo-natales) et chez les personnes âgées ou immunodéprimées.

Ce travail s'inscrit dans l'une des thématiques de recherche de l'équipe B.R.M.F (Bactéries et Risque Materno-Fœtal) qui vise à étudier (i) le rôle des ARN régulateurs (ARNrég) dans la physiologie, la virulence et l'adaptation de *S. agalactiae* à différents hôtes et niches écologiques, et (ii) leur mécanisme d'action.

Dans ce mémoire de thèse, nous présenterons en première partie une revue de la littérature concernant : (i) *S. agalactiae*, notre modèle d'étude ; (ii) les différents systèmes de régulation chez *S. agalactiae* et (iii) les mécanismes d'actions et le rôle des ARNrég chez les bactéries.

Dans une seconde partie, nous exposerons les résultats de nos travaux expérimentaux portant sur (i) la caractérisation de l'ARNrég CetR et son rôle dans la résistance aux peptides antimicrobiens et dans la virulence (publication N°1), (ii) rôle de l'ARNrég CetR dans la croissance en milieu chimiquement défini et l'impact de sa délétion sur le protéome de *S. agalactiae* (publication N°2, en préparation) et (iii) enfin, nous présenterons les résultats non publiés portant sur la caractérisation des ARNrég Srn008, Srn015, Srn017, Srn024, Srn046 et Srn070.

REVUE DE LA LITTERATURE

I. *STREPTOCOCCUS AGALACTIAE*

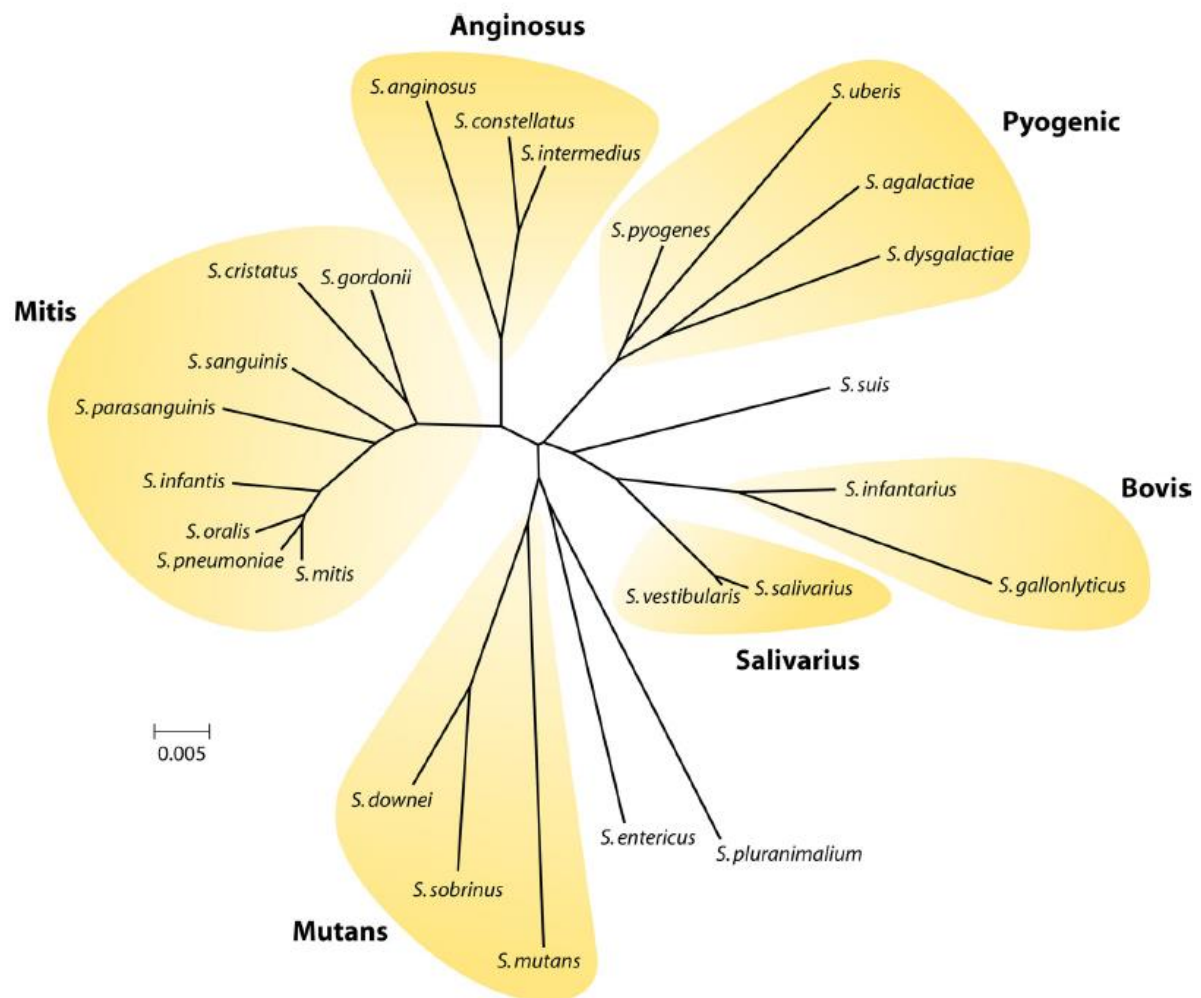


Figure 1 : Arbre phylogénétique du genre *Streptococcus* basé sur les séquences d'ARNr 16S.
D'après (Nobbs *et al.*, 2009)

I. *Streptococcus agalactiae*

A. Caractéristiques phénotypiques et génétiques

1. Classification du genre *Streptococcus*

Le genre *Streptococcus* a été découvert pour la première fois en 1879 par Louis Pasteur lors de l'examen d'une infection du postpartum. Il observa chez les femmes malades des granules arrondis disposés sous forme de chaînes de perles. A cette époque, ce « microorganisme », *Streptococcus pyogenes*, causait des taux de mortalité élevés dans les maternités. Quant à *Streptococcus agalactiae*, il a été initialement décrit en 1887, comme responsable de mammites chez les bovins (Nocard and Mollereau, 1887). Des années après cette découverte, Rebecca Lancefield proposa la première classification des streptocoques basée sur la composition en carbone des antigènes de la paroi (Lancefield, 1933). Ainsi, on distingue vingt sérogroupes (groupes : A à H, K à P, R à V et non groupable). Les antigènes permettant cette classification sont soit des polysaccharides (ex. groupes A, B, C, E, F et G), des acides téichoïques (ex. groupe D et N) ou des acides lipotéichoïques (LTA) (ex. groupe H). Il existe également une autre classification basée sur la capacité des streptocoques à lyser les globules rouges. On distingue, les alpha-hémolytiques, les bêta-hémolytiques et les gamma-non-hémolytiques. L'espèce *Streptococcus agalactiae* appartient aux « streptocoques du groupe B » bêta-hémolytique, appelé également « SGB » (Lancefield and Freimer, 1966). Son implication en pathologie humaine n'a été établie qu'à partir des années 1960 (Eickhoff *et al.*, 1964).

2. Taxonomie

Streptococcus agalactiae appartient au phylum des *Firmicutes*, à l'ordre des *Lactobacillales*, et plus précisément à la famille des *Streptococcaceae*. Cette famille regroupe les genres *Streptococcus*, *Lactococcus* et *Lactovum*. A partir de la fin du XXème siècle, l'arrivée des techniques de biologie moléculaire a contribué à l'amélioration de nos connaissances concernant la phylogénie bactérienne. Le typage moléculaire basé sur l'ARNr 16S a permis d'organiser le genre streptocoque en six clusters (Figure 1). Chaque cluster inclut plusieurs espèces, dont le cluster Pyogène auquel appartient *S. agalactiae*. Au moment de la rédaction de ce manuscrit, le genre *Streptococcus* comportait 119 espèces.

Famille	Genre	Espèce	Sous-espèce	Souche
Séquençage génomique				
	Séquençage de l'ARN 16S			
	Pourcentage en GC			
	Hybridation ADN-ADN			
		Typage de séquences multilocus		
		Profilage des protéines totales		
		Prise d'empreinte génomique		

Figure 2 : La résolution taxonomique relative de diverses techniques moléculaires.

D'après (Prescott *et al.*, 2010)

D'autres études proposent l'utilisation de gènes de ménage « *housekeeping* », l'analyse de la présence/absence de gènes de virulence afin d'identifier de nouvelles espèces et de discriminer les espèces qui sont très proches en termes de séquences d'ARNr 16S (Kilian *et al.*, 2008; Bishop *et al.*, 2009). Ainsi, des techniques de biologie moléculaire, tel que la prise d'empreinte génomique, profilage de protéines totales et séquençage massif ont apporté une aide précieuse dans la caractérisation et la différenciation des souches bactérienne (Figure 2) (Prescott *et al.*, 2010).

L'emploi des séquences d'ADN pour déterminer l'identité d'une espèce ou d'une souche exige l'analyse de gènes qui évoluent plus rapidement que ceux qui codent l'ARNr. Souvent, on séquence et on compare de cinq à sept gènes de ménage, une technique appelée typage de séquence multilocus (MLST pour « *multilocus sequence typing* »). A l'origine, le MLST fut mis au point pour distinguer les souches pathogènes, mais son application s'est élargie à la taxonomie bactérienne (Maiden *et al.*, 1998). Utilisée pour la première fois chez *S. agalactiae* en 2003, cette technique se base sur l'étude de la composition nucléotidique de sept gènes de ménage : *adhP* (*Alcohol dehydrogenase, propanol-preferring*), *pheS* (*Phenylalanine--tRNA synthetase*), *atr* (*amino-acid transporter protein*), *glnA* (*Glutamine synthetase, type I*), *sdhA* (*L-serine dehydratase*), *glcK* (*Glucose kinase*) et *tkt* (*Transketolase*) (Jones *et al.*, 2003). En fonction des variations nucléotidiques, un numéro d'allèle est donné pour chacun des sept gènes étudiés, permettant ainsi d'attribuer une valeur de ST « *Sequence Type* » et une valeur de CC « *Clonal Complex* » à la souche étudiée. Un CC est défini comme un groupe MLST dont le profil allélique des souches est identique pour au moins six gènes sur les sept étudiés, et il doit être composé d'au moins deux ST différents, sinon on appelle ceci un singleton.

3. Organisation du génome

L'espèce *S. agalactiae* appartient à la famille des *Streptococcaceae* qui regroupe des espèces bactériennes ayant un faible taux en GC (guanine-cytosine) représentant environ 35,5 % de leur génome. C'est au début de XXI^{ème} siècle que le premier génome de la souche *S. agalactiae* NEM316 a été totalement séquencé et assemblé (Glaser *et al.*, 2002). Cette étude a permis de montrer que le génome de cette souche est composé de 2118 gènes, dont environ 55% représentent le « *core genome* », d'un génome dispensable constitué de gènes

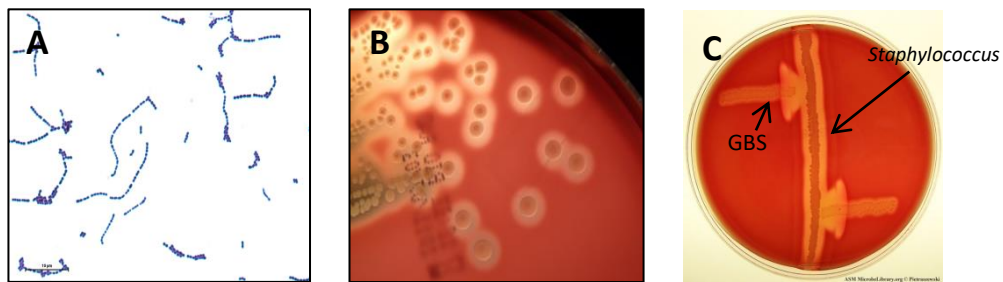


Figure 3 : Caractéristiques morphologiques et culturelles de *Streptococcus agalactiae*.

(A) Image en microscopie photonique d'une coloration du Gram de *S. agalactiae*. (B) Image de colonies de streptocoques sur une gélose au sang de cheval. (C) Image du CAMP test.

(<http://lib.jiangnan.edu.cn/asm/>)

partiellement communs ou spécifiques à la souche et environ 50% de gènes appartenant à 14 îlots génomiques. Ces derniers, sont composés de gènes codant des protéines de surface, des gènes associés à des éléments génétiques mobiles, ou encore des gènes de virulence. Cette composition de gènes suggère que parmi ces îlots génomiques, certains pourraient être des îlots de pathogénicité (Glaser *et al.*, 2002). Actuellement, on dénombre 31 génomes de l'espèce *S. agalactiae* totalement séquencés et assemblés, disponibles dans la base de données NCBI (<http://www.ncbi.nlm.nih.gov/>).

B. Caractéristiques morphologiques et structurales

1. Caractères morphologiques, culturels et biochimiques

Les streptocoques partagent les caractéristiques communes des coques à Gram positif, immobiles, mésophiles et non sporulants (Prescott *et al.*, 2010). L'espèce *S. agalactiae* se présente sous forme sphérique ou ovoïde, de diamètre mesurant 0,6 à 1,2 µm et groupée en diplocoques ou en chaînettes de longueur variable (Figure 3A).

Streptococcus agalactiae est décrite comme une bactérie peu exigeante, capable de croître dans des milieux de culture commercialisés, tels que le milieu Todd Hewitt (TH), ou cœur cerveau (« *Brain Heart Infusion* », BHI). Les streptocoques nécessitent pour leur multiplication de nombreux facteurs de croissance présents dans ces milieux de culture complexes, tels la phénylalanine, le glutamate, la tyrosine, le tryptophane, l'arginine, la valine, la leucine, la lysine, la méthionine, l'isoleucine, la cystéine et l'histidine. Après une période d'incubation de 24 heures à 37°C sur un milieu gélosé additionné de sang de cheval ou de mouton, *S. agalactiae* se présente sous forme de colonies rondes, de petite taille d'environ 3 à 5 mm de diamètre, de couleur blanchâtre, plate et lisse (Figure 3B). De plus, les colonies de *S. agalactiae* masquent une étroite zone d'hémolyse totale, appelée β-hémolyse. Ces colonies peuvent produire un pigment jaune à rouge brique (Figure 3B).

S. agalactiae est une bactérie aérobie-anaérobie facultative, chimio-hétérotrophe, ne possédant pas de catalase et incapable de réduire les nitrates. Les principaux caractères phénotypiques d'identification de cette espèce sont les tests de résistance à la bacitracine, au triméthoprim-sulfaméthoxazole et l'hydrolyse de l'hippurate. En microbiologie clinique,

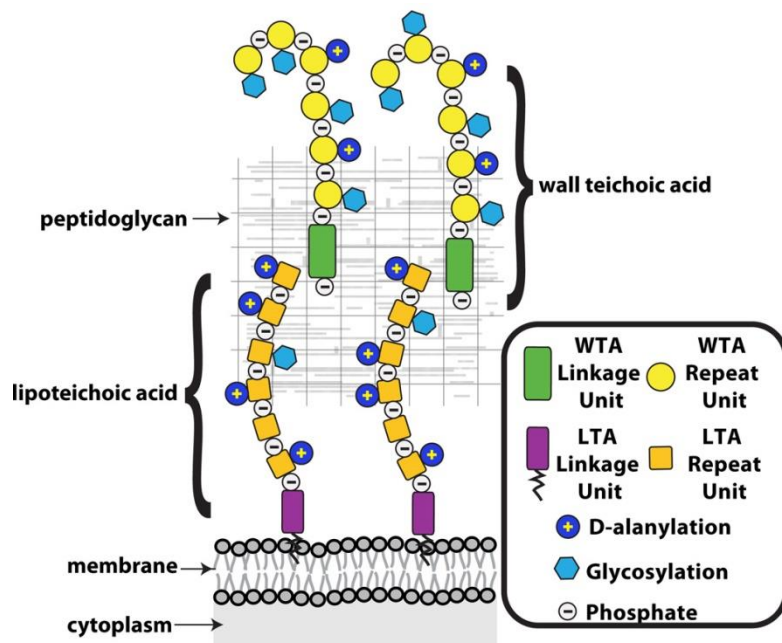


Figure 4 : Représentation schématique de la paroi des bactéries à Gram positif.

LTA : Lipoteichoic acid, WTA : Wall teichoic acid.

D'après (Brown *et al.*, 2013)

cette identification est rendue possible grâce à l'analyse des profils protéomiques des souches par la spectrométrie de masse. Le « CAMP test » (Christie-Atkins-Munch-Peterson) (Christie *et al.*, 1944) est également utilisé pour cette identification. Il met en évidence une hémolyse synergique sur gélose au sang de mouton en présence de la β -toxine de *Staphylococcus aureus* (Figure 3C). De plus, l'ensemble des bactéries appartenant à la famille des *Streptococcaceae* sont des bactéries lactiques. Leur métabolisme énergétique aboutit à la production d'acides lactique, acétique, formique, d'éthanol et de dioxyde de carbone par fermentation du glucose (Brown and Collins, 1977).

2. La paroi bactérienne

A l'inverse des bactéries à Gram négatif, *S. agalactiae* fait partie des bactéries à Gram positif dont la paroi cellulaire est constituée d'une seule bicouche phospholipidique (Figure 4). La protection de la membrane est assurée par une couche épaisse de peptidoglycane, qui agit également comme une structure d'ancrage pour les autres composants de la surface cellulaire, tels que les polymères d'acides téichoïques. D'autres composants fixés aux lipides membranaires, les acides lipotéichoïques (ALT), préservent l'intégrité de la paroi cellulaire et fournissent la majeure partie de la charge négative globale de la surface cellulaire. Ces polymères peuvent représenter plus de 60% de la masse de la paroi cellulaire des Gram positif (Silhavy *et al.*, 2010). Chez *S. agalactiae*, les protéines exposées à la surface cellulaire sont attachées à la paroi et peuvent être classées en deux familles : les protéines à motifs LPXTG, exposées à la surface, et les lipoprotéines qui ne sont pas accessibles en surface puisqu'elles sont masquées par d'autres composants de l'enveloppe cellulaire (Lindahl *et al.*, 2005). Vingt et une protéines à motifs LPXTG et 36 lipoprotéines ont été identifiées chez la souche *S. agalactiae* NEM316 (Glaser *et al.*, 2002). Vingt-quatre protéines à motifs LPXTG et 51 lipoprotéines ont été identifiées chez la souche 2603V/R (Tettelin *et al.*, 2005). Les fonctions de ces protéines sont diverses (adhésines, nucléases, peptidase, protéases, enzymes, système de transport) ou encore inconnues. Enfin, les polysaccharides capsulaires forment la couche la plus externe de l'enveloppe cellulaire et jouent également le rôle des antigènes.

3. Les acides lipotéichoïques

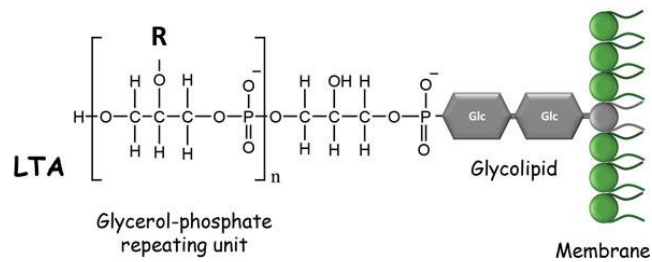


Figure 5 : Représentation schématique de la structure des acides lipotéichoïques (ALT) et leur ancrage à la membrane des bactéries à Gram positif.

Le R indique la substitution avec le D-Ala.

D'après (Chapot-Chartier and Kulakauskas, 2014)

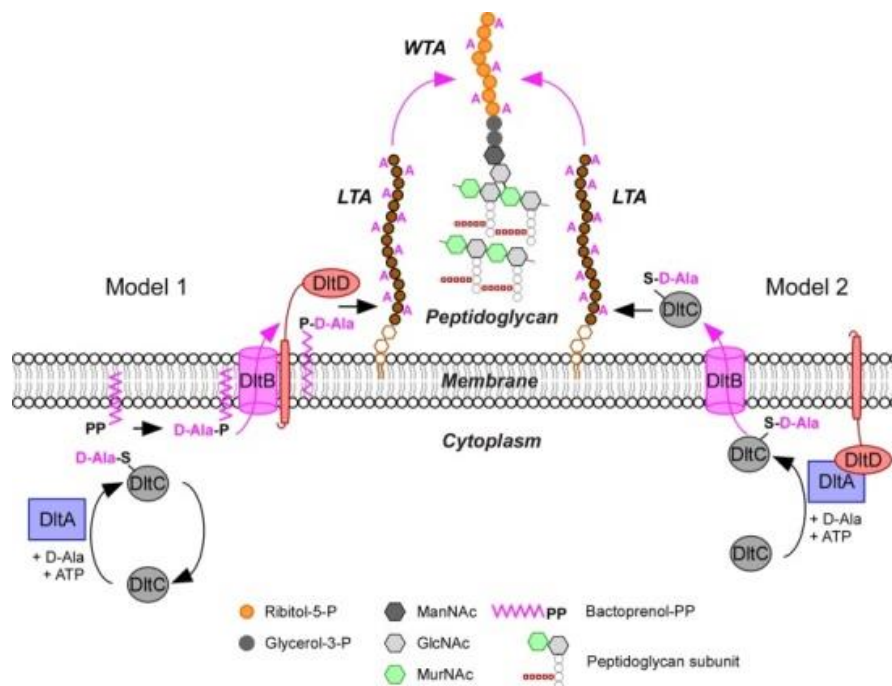


Figure 6 : Modèles de mécanismes de D-alanylation des acides lipotéichoïques (ALT).

Le modèle 1 propose que l'activité combinée de DltA et DltC, dans le cytoplasme, permet le transfert de D-Ala sur le « *lipid carrier bactoprenol pyrophosphate* » (PP) (Perego *et al.*, 1995). Ensuite, DltB transloque le D-Ala-bactoprenol-PP à travers la membrane. Il est difficile de savoir si DltD améliore l'action de la flippase (DltB) ou bien permet le transfert de D-Ala aux ALT (Reichmann *et al.*, 2013). Dans le modèle 2, DltD se trouvant du côté du cytoplasme, facilite le transfert de D-Ala à DltC. Ensuite, le DltC-S-D-Ala est transloqué à travers la membrane par DltB et incorpore le D-Ala aux ALT (Neuhaus and Baddiley, 2003).

D'après (Schneewind and Missiakas, 2014)

Les acides lipotéichoïques (ALT) sont des polymères de glycérol ou de ribitol reliés à des groupements phosphates, liés directement aux lipides de la membrane. Ils confèrent à la paroi cellulaire sa charge négative. Malgré des études montrant le rôle des LTA dans le processus d'adhésion des cellules bactériennes aux cellules épithéliales humaines (Nealon and Mattingly, 1984; Teti *et al.*, 1987), leur fonction est controversée (Edwards and Baker, 2005). D'autre part, il a été montré que l'incorporation de D-alanine au niveau des ALT, *via* l'expression de l'opéron *dlt* permet à la bactérie de résister aux peptides cationiques (Poyart *et al.*, 2001; Poyart *et al.*, 2003). De plus, les ALT jouent le rôle d'un ligand pour les TLR (Toll like receptor), activant la production de cytokines inflammatoires en réponse à l'infection par *Staphylococcus aureus* (Gillrie *et al.*, 2010).

a- Biosynthèse des acides lipotéichoïques

Les ALT sont liés à la cellule bactérienne par un glycolipide ancré dans la couche externe de la membrane (Figure 5). Ce glycolipide est un diglucosyldiacylglycérol synthétisé à partir de diacylglycérols par l'addition successive de deux molécules de glucose (Glu) fourni par l'UDP-Glu (Reichmann and Grundling, 2011). Le diglucosyldiacylglycérol est transféré de l'intérieur vers le côté extérieur de la membrane par une transférase, LtaA. Ainsi, l'ajout des groupements Gro-1-P (glycérol-phosphate) permet la polymérisation de la partie diglucosyl de la chaîne de l'ALT par une ALT synthase LtaS. Dans la majorité des espèces, cette chaîne peut contenir jusqu'à cinquante unités de Gro-1-P. Les groupes hydroxyles libres des unités de glycérol-phosphates constituent la matrice d'ancrage de D-alanine lors de la D-alanylation des LTAs.

b- La D-alanylation des acides lipotéichoïques

Comme mentionné ci-dessus, le mécanisme d'incorporation de D-Ala aux ALT est très bien caractérisé et implique quatre protéines encodées par l'opéron *dltABCD*. La même organisation de cet opéron est conservée chez de nombreuses bactéries à Gram positif (Neuhaus and Baddiley, 2003). La fonction de DltA et DltC dans ce processus a été bien établie. DltA est une « *D-alanine-D-alanyl carrier protein ligase* » qui permet l'activation de l'acide aminé D-Ala en D-alanyl et son transfert à DltC « *D-alanyl carrier protein* » avec la consommation d'une molécule d'ATP (Figure 6) (Fischer, 1994; Debarov *et al.*, 1996). Le rôle de DltB et DltD est moins clair. Cependant, deux modèles ont été proposés pour le

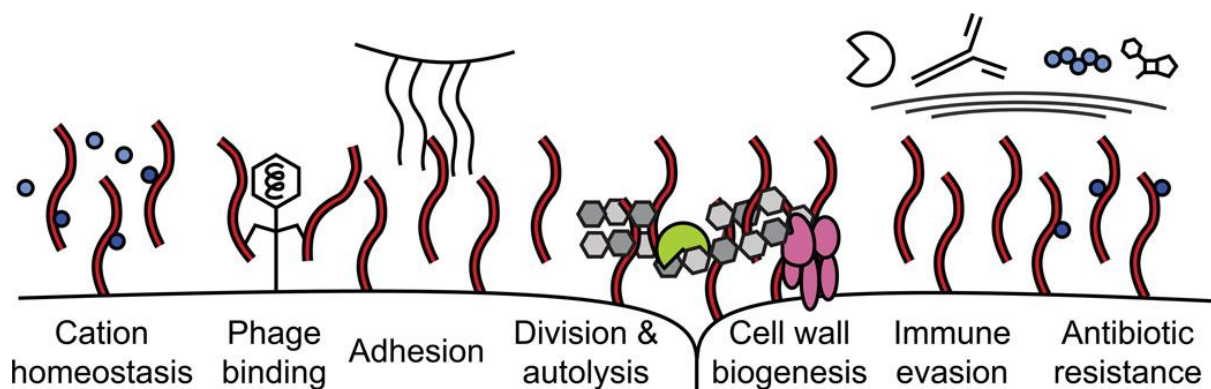


Figure 7 : Rôles et fonctions des acides lipotéichoïques.

D'après (Pasquina *et al.*, 2013)

Tableau 1 : Classification des principales familles d'antibiotiques en fonction de leur mécanisme d'action

Mécanisme d'action	Familles d'antibiotique
Inhibition de la synthèse de la paroi bactérienne	β-lactamines Pénicillines, ampicilline, céfotaxime Glycopeptides Vancomycine, teicoplanine
Inhibition de la synthèse protéique	Macrolides, lincosamides, streptogramines Erythromycine, lincomycine, pristnamycine Aminoglycosides Gentamicine Cyclines Tétracycline
Inhibition de la synthèse des acides nucléiques	Fluoroquinolones Norfloxacin, lévofloxacin Rifampicine Rifampicine
Action sur la membrane cellulaire	Peptides antimicrobiens (PAM) Colistine, polymyxine B, nisine, LL-37 Gramicidines Bacitracine Lipopeptides Daptomycine

fonctionnement de ces deux protéines (Figure 6). Dans le premier modèle, Fischer et ses collaborateurs proposèrent que la D-alanylation des ALT soit assurée par un groupement intermédiaire, un « *undecaprenyl phosphate* » (C₅₅-P), associé au lipides membranaires (Perego *et al.*, 1995). Dans ce modèle, tandis que DltB facilite le transfert de D-Ala de DltC à C₅₅-P de l'intérieur vers l'extérieur de la membrane, DltD assure le transfert de D-Ala aux ALT (Figure 6, partie gauche). Dans le deuxième modèle (Neuhaus and Baddiley, 2003), DltD agit plutôt dans le cytoplasme et permet le transfert de D-Ala à DltC (Figure 6, partie droite). Le complexe DltC-Ala est ensuite transloqué à travers la membrane par DltB, permettant ainsi l'incorporation de D-Ala aux ALT. Une étude récente consolide le premier modèle et suggère que DltD est plutôt localisé à l'extérieur de la membrane (Reichmann *et al.*, 2013). Néanmoins, cette étude n'a pas mis en évidence le rôle du groupement C₅₅-P dans le transport de D-Ala à travers la membrane.

c- Rôle et fonction des acides lipotéichoïques

Les ALT sont des cibles antibactériennes prometteuses, en raison de leur rôle crucial dans de nombreuses fonctions vitales. Il a été montré que les ALT jouent un rôle essentiel dans la viabilité bactérienne *in vitro* (Oku *et al.*, 2009), la division cellulaire (Grundling and Schneewind, 2007; Campbell *et al.*, 2011), l'échappement au système immunitaire (Atilano *et al.*, 2011; Smith *et al.*, 2012), la résistance aux peptides antimicrobiens (PAM)(Peschel *et al.*, 1999; Saar-Dover *et al.*, 2012), la résistance aux β -lactamines (Campbell *et al.*, 2011) et dans la virulence *in vivo* (Poyart *et al.*, 2003; Weidenmaier *et al.*, 2004) (Figure 7).

C. Résistance aux agents antimicrobiens

1. Classification et mécanismes d'action des antibiotiques

Les antibiotiques (ATB) ont la capacité de tuer les bactéries ou encore d'inhiber leur croissance, on parle alors respectivement de bactéricidie ou de bactériostase. Afin qu'il soit efficace et sans danger pour le patient, un ATB doit posséder une cible qui soit unique au monde bactérien et absente, ou du moins inaccessible chez le patient. Les ATB agissent, en général, à un niveau précis des structures bactériennes. Selon leur mode d'action, on distingue quatre catégories d'ATB (Tableau 1) : 1) les ATB inhibant sur la synthèse de la paroi bactérienne, une excellente cible pour les antibiotiques, car elle est absente chez les

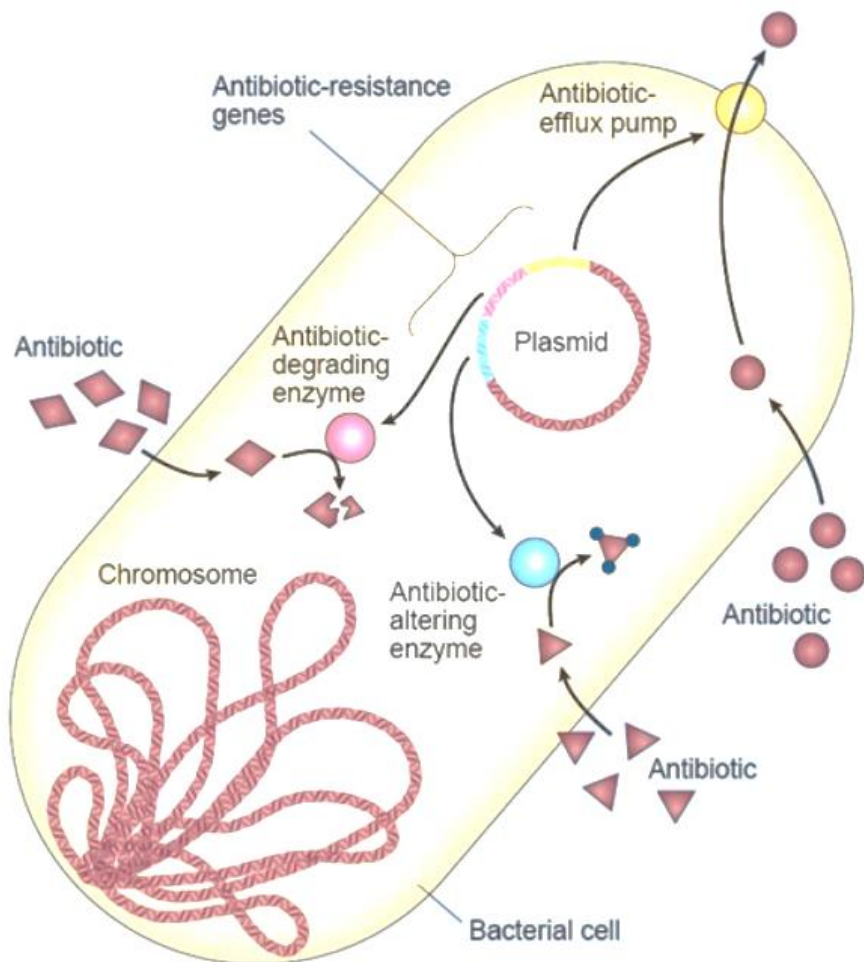


Figure 8 : Illustration des divers mécanismes de résistance aux antibiotiques utilisés par les bactéries.

D'après (Levy and Marshall, 2004)

eucaryotes ; 2) les ATB inhibant la synthèse protéique, agissent au niveau de la traduction, en ciblant soit la sous-unité 30S, soit la sous-unité 50S du ribosome bactérien, absent chez les eucaryotes ; 3) les ATB inhibant la synthèse des acides nucléiques (ARN et ADN) par l'interférence avec la fonction de la polymérase ; et 4) les ATB actifs sur la membrane, une stratégie peu employée en raison de la grande similitude existant entre les membranes bactériennes et eucaryotes. Il existe une dernière catégorie d'ATB qui inhibent certaines voies métaboliques essentielles chez la bactérie. L'acide para-aminosalicylique, analogue de l'acide para-aminobenzoïque, induit l'inhibition des purines et des pyrimidines.

2. Mécanismes de résistance aux antibiotiques

La résistance des bactéries aux antibiotiques est soit naturelle, soit acquise. La résistance naturelle d'une espèce est une caractéristique propre, appartenant à l'ensemble des souches de cette espèce. Cette résistance est chromosomique. La résistance acquise ne concerne, elle, que quelques souches d'une même espèce. Elle peut être chromosomique ou plasmidique. La résistance chromosomique peut-être secondaire à une mutation portant sur le chromosome de la bactérie. Cette résistance est transmise à la descendance de cette bactérie nouvellement résistante. Dans le cas de la résistance liée à l'acquisition d'un plasmide (ou d'un transposon) portant un gène de résistance, non seulement la bactérie transmettra ce gène à sa descendance, mais il lui sera également possible d'échanger cet élément mobile avec d'autres bactéries présentes dans son environnement par conjugaison (ou transposition). Il existe plusieurs mécanismes de résistance aux ATB (Figure 8) :

- 1- Efflux de l'antibiotique à l'extérieur de la cellule à l'aide de pompes. L'antibiotique est alors expulsé à une vitesse supérieure à sa vitesse de diffusion, de telle sorte qu'il est impossible qu'il atteigne une concentration intracellulaire suffisante pour exercer son effet.
- 2- Inactivation enzymatique de l'antibiotique. Le cas classique est l'inactivation des β -lactamines par hydrolyse de leur anneau β -lactame grâce à la production de β -lactamases.
- 3- Modification de la cible de l'antibiotique afin de lui permettre de conserver son activité physiologique, tout en la rendant invisible ou du moins protégée de l'action de l'antibiotique.

4- Diminution de la perméabilité membranaire par réduction du nombre des porines.

Cette résistance est le fruit de la remarquable adaptabilité des bactéries, qui ont su développer différents mécanismes afin de contrer les effets nocifs des antibiotiques. L'espèce *S. agalactiae* est naturellement résistante à la bacitracine, l'acide nalidixique, aux polymyxines, au triméthoprim-sulfaméthoxazole, au métronidazole, aux aminosides à bas niveau et aux PAM tels que la colistine et la polymyxine B (Edwards and Baker, 2005; Brzychczy-Wloch *et al.*, 2010; Kasahara *et al.*, 2010; Da Cunha *et al.*, 2014). En revanche, cette bactérie est naturellement sensible aux β -lactamines telles la pénicilline G, l'amoxicilline ou le céfotaxime, aux macrolides, lincosamides et streptogramines A et B (MLS) qui inhibent la synthèse protéique bactérienne en se fixant à la sous-unité 50S du ribosome bactérien (Ungureanu, 2010).

3. Résistance aux PAM

a- Mode d'action des peptides antimicrobiens cationiques

Les PAM cationiques, appelés également CAMP « *cationic antimicrobial peptide* » en anglais, sont des peptides multifonctionnels jouant un rôle biologique fondamental, notamment dans l'élimination de micro-organismes pathogènes (Diamond *et al.*, 2009). Le caractère cationique des PAM explique leur rapidité d'action et leur affinité pour les membranes anioniques des bactéries : plus la charge est importante, plus les interactions augmentent. De plus, chaque peptide a un spectre d'activité distinct. Ainsi, la taille, la charge, la séquence en acides aminés, l'hydrophobicité, l'amphiphilie ou encore le degré de structuration sont des facteurs spécifiques à chaque peptide (Gennaro and Zanetti, 2000). Les mécanismes par lesquels les peptides perméabilisent et traversent les membranes bactériennes ne sont pas encore complètement élucidés et varient selon les peptides (Yeaman and Yount, 2003).

Chez les bactéries à Gram positif, le mécanisme d'action des peptides dépend en tout premier lieu de l'interaction du peptide avec la paroi bactérienne. Cette attraction initiale se fait par l'intermédiaire de liaisons électrostatiques entre le peptide cationique et

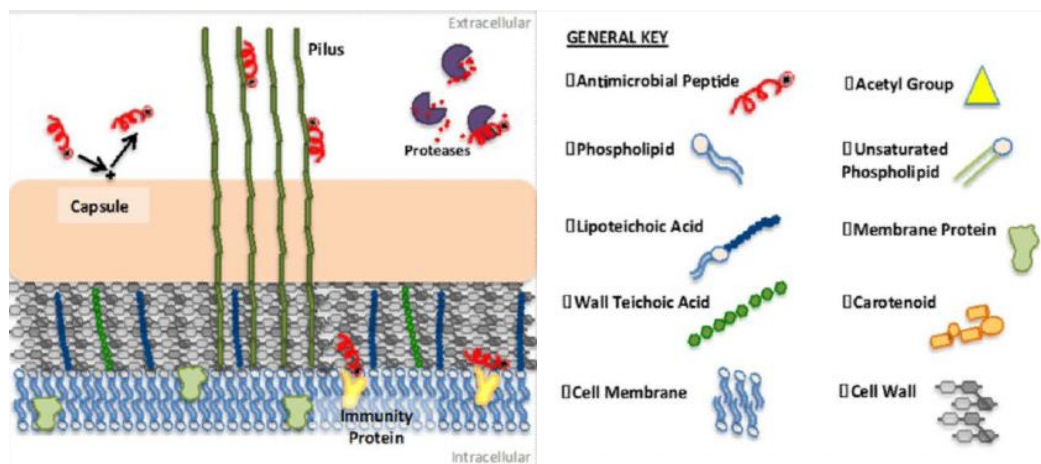


Figure 9 : Principaux mécanismes de résistance aux PAM chez les *Streptococcus*.

On distingue plusieurs mécanismes de résistance aux PAM, parmi eux : la dégradation par les protéases sécrétées, la séquestration par les protéines associées à la membrane (ex. pili), ou le blocage par les polysaccharides de la capsule.

D'après (Nawrocki *et al.*, 2014)

les composants du peptidoglycane (acides téichoïques et ALT) (Dathe *et al.*, 2001). Vient ensuite l'étape d'attachement au peptidoglycane. Cette fixation se fait grâce aux groupements phosphates qui sont situés au niveau des ALT présents à la surface de la paroi. L'importante proportion de lipides chargés négativement à la surface de la paroi des bactéries permet aux PAM de cibler les bactéries et non pas les cellules de l'hôte. Enfin, après la fixation, les peptides vont entraîner des perturbations et/ou une modification de la membrane plasmique ; cela aboutit à une modification de la perméabilité membranaire.

b- Principaux mécanismes de résistance aux PAM chez les *Streptococcus*

Depuis la découverte des premiers PAM, beaucoup d'études ont été menées afin de comprendre leur mécanisme d'action et d'essayer de les synthétiser afin d'en faire des agents antibactériens. Comme nous l'avons vu plus haut, la D-alanylation des ALT de la paroi est un des principaux mécanismes responsables de la résistance aux PAM chez *S. agalactiae* (Saar-Dover *et al.*, 2012). Cependant, d'autres mécanismes sont utilisés par cette bactérie pour contrer l'action des PAM à différents niveaux, notamment la dégradation enzymatique, la séquestration par des structures supramoléculaires et la répulsion par la capsule (Figure 9).

Dégradation enzymatique. La dégradation des PAM par des protéases est un mécanisme de résistance décrit chez plusieurs espèces à Gram positif, telles que *Enterococcus faecalis* (Schmidtchen *et al.*, 2002), *Staphylococcus aureus* (Sabat *et al.*, 2000) ou *Streptococcus pyogenes* (Hauser *et al.*, 1991). Cette résistance adaptative se fait par sécrétion de protéases et peptidases, capables de cliver les peptides. Par exemple, *S. pyogenes* sécrète une cystéine protéinase SepB qui inactive par clivage la cathélicidine humaine « LL-37 » (Schmidtchen *et al.*, 2002).

Séquestration par les protéines extracellulaires. Certaines bactéries à Gram positif produisent des protéines extracellulaires ou fixées à la surface de la paroi, qui se lient directement aux PAM et bloquent leur accès à la membrane cellulaire (Braff *et al.*, 2007; Lauth *et al.*, 2009). L'espèce *S. pyogenes* sécrète le « *Streptococcal inhibitor of complement* » (SIC) dans le milieu extracellulaire permettant ainsi la séquestration de plusieurs types de PAM. Par exemple, cette enzyme (SIC) se lie à des α -defensines, LL-37 et au lysozyme afin de

neutraliser l'action des PAM contre la paroi de *S. pyogenes* (Fernie-King *et al.*, 2002). Des études ont montré que la staphylokinase se lie à la cathélicidine de murin (mCRAMP) afin de réduire son effet bactéricide (Jin *et al.*, 2004; Braff *et al.*, 2007). Des protéines associées à l'enveloppe cellulaire peuvent aussi se lier aux PAM et empêcher leur contact avec les composants de la membrane. L'expression de la sous-unité du pili PilB de *S. agalactiae* améliore sa résistance à la cathélicidine de murin *in vitro* (Maisey *et al.*, 2008).

Inhibition de l'activité des PAM par les polysaccharides capsulaires. Les polysaccharides capsulaires protègent les bactéries contre les PAM *via* des forces d'attractions ou de répulsions électrostatiques (Campos *et al.*, 2004). L'interaction capsule-PAM est facilitée par l'interaction électrostatique entre les PAM chargés positivement et les polysaccharides capsulaires chargés négativement. Une étude chez *S. pneumoniae* a montré que l'interaction entre un extrait de la capsule et les deux PAM, la polymyxine B et l' α -défensine, empêche ces dernières d'atteindre la membrane et permet la survie bactérienne *in vitro* (Llobet *et al.*, 2008).

Enfin, les PAM sont impliqués dans de nombreuses maladies, par leur déficit mais aussi par leur surexpression. Ainsi, des maladies complexes comme la maladie de Crohn, la parodontite, l'athérosclérose, mais aussi certains cancers mettent en jeu des mécanismes jusqu'à présents mal connus faisant intervenir les PAM. Leur présence est le plus souvent indispensable, les PAM jouent un rôle protecteur contre les infections bactériennes, notamment dans la mucoviscidose, la maladie de Crohn, ou encore dans la protection du greffon de peau chez les grands brûlés (Zaiou, 2007).

II. LES INFECTIONS A *STREPTOCOCCUS AGALACTIAE*

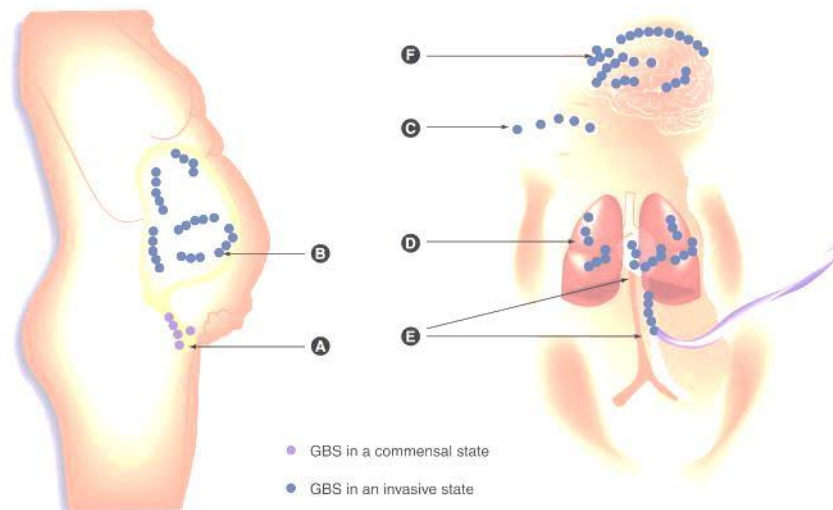


Figure 10 : *Streptococcus agalactiae*, un pathogène du nouveau-né.

(A) *S. agalactiae* réside en tant que commensal dans le tractus génital et gastro-intestinal inférieur de la femme. (B) Cette bactérie peut s'infiltrer dans le compartiment intra-utérin de la femme enceinte qui est un porteur asymptomatique. (C) Le nouveau-né inhale ou ingère *S. agalactiae* dans l'utérus ou pendant l'accouchement. (D) Ensuite, la bactérie envahit les poumons du nouveau-né causant une pneumonie. (E) A partir des poumons, *S. agalactiae* accède à la circulation sanguine du nouveau-né provoquant une septicémie, puis envahit de multiples organes néonataux, y compris le cœur. (F) Elle pénètre également la barrière hémato-encéphalique, provoquant une méningite.

D'après (Rajagopal, 2009)

II. Les infections à *Streptococcus agalactiae*

A. *Streptococcus agalactiae* : un pathogène opportuniste

Le Streptocoque du groupe B (SGB) ou *Streptococcus agalactiae* est avant tout une bactérie commensale présente au niveau du tractus gastro-intestinal et génital de l'Homme (Six *et al.*, 2014) (Figure 10). La colonisation vaginale à *S. agalactiae* est qualifiée de transitoire ou intermittente et seules 30 à 40% des femmes colonisées en début de grossesse possèdent encore la bactérie au moment de l'accouchement (Six *et al.*, 2014; Joubrel *et al.*, 2015). C'est seulement à partir des années 1970 que *S. agalactiae* est reconnu comme le principal agent responsable des infections néonatales, infections dont les conséquences peuvent être dangereuses (Baker and Edwards, 1990). *S. agalactiae* a été identifié chez diverses espèces animales telles que les poissons (Evans *et al.*, 2006; Pereira *et al.*, 2010) ou divers mammifères (bovin, chien, chat, porc, singe, dauphin....) (Elliott *et al.*, 1990; Evans *et al.*, 2006; Pereira *et al.*, 2010). Son pouvoir pathogène est peu documenté pour ces espèces. En revanche, son rôle en pathologie bovine et humaine a été établi depuis de nombreuses années.

1. Infection chez le bovin

S. agalactiae était l'un des principaux pathogènes responsables de mammites sub-cliniques. Avant l'utilisation des antibiotiques, 90 % des mammites lui étaient attribuables (inflammation modérée sans signe visible au niveau du pis de de la vache, de la mamelle ou du lait) (Keefe, 1997). *S. agalactiae* a été décrit pour la première fois par le médecin vétérinaire et microbiologiste Edmond Nocard comme le principal agent responsable de la mammite bovine (Nocard and Mollereau, 1887). Cette pathologie représente encore aujourd'hui un problème considérable pour la filière laitière de l'industrie agroalimentaire aux Etats-Unis ou le Canada (Hillerton and Berry, 2003). Ces infections, qui surviennent dans les élevages laitiers, engendrent une forte diminution de la quantité et de la qualité du lait produit par l'animal ainsi que d'importantes pertes économiques pour le secteur de l'industrie laitière (Keefe, 1997). En France, les mammites bovines sont peu fréquentes (Djabri *et al.*, 2002). La mise en place de programme d'antibioprophylaxie a conduit à une diminution globale du nombre de cas aux Etats-Unis et au Canada, dont la prévalence est parfois supérieure à 10 % (Djabri *et al.*, 2002; Hillerton and Berry, 2003).

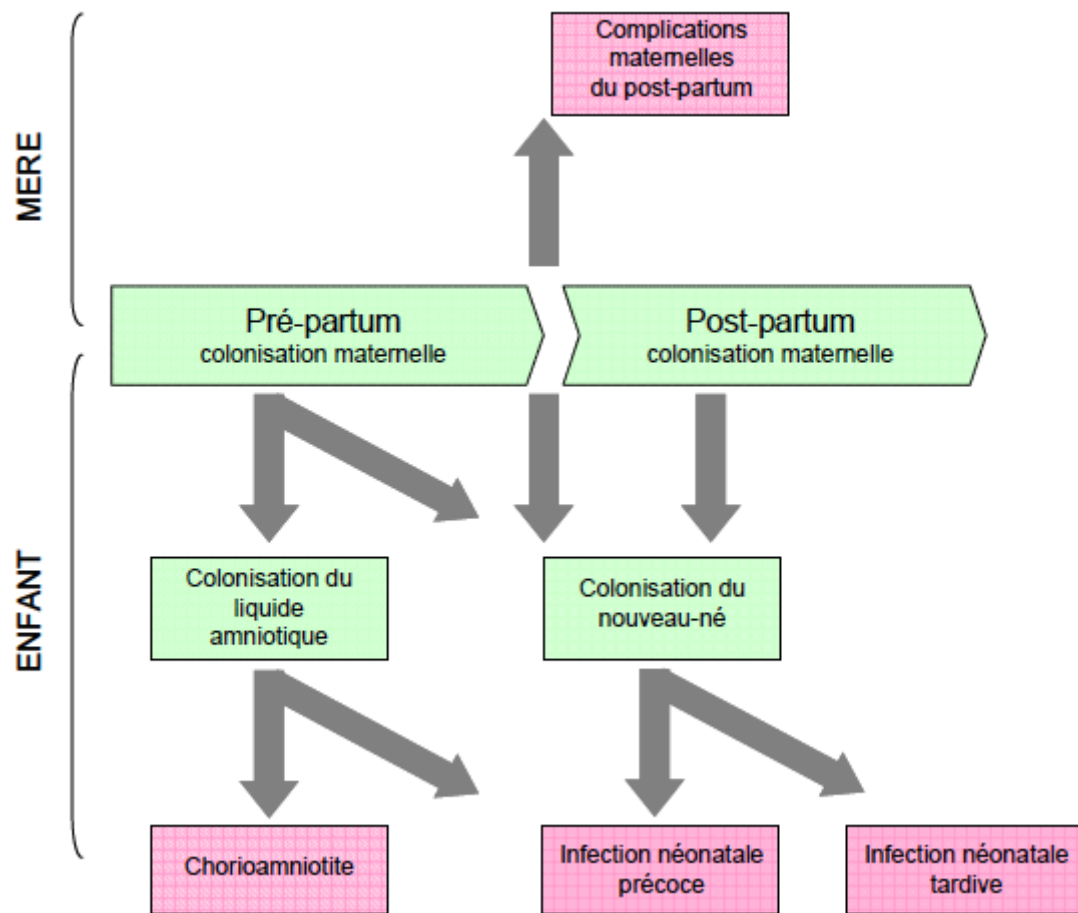


Figure 11 : Schéma représentatif des différentes entités d'infections materno-fœtale et leur transmission.

D'après (Francis, 2013)

2. Infection chez l'Homme

Les infections humaines à *S. agalactiae* représentent un réel problème de santé publique dans les pays industrialisés et suscitent un intérêt grandissant. Reconnu comme le principal agent étiologique des infections néonatales dès les années 1970 (Schuchat, 1995), *S. agalactiae* est aussi responsable d'infections graves chez la femme enceinte ou en post-partum. Elle est également à l'origine d'infections invasives chez l'adulte en dehors de la grossesse, en particulier chez les sujets âgés de plus de 65 ans et chez les sujets atteints de pathologies sous-jacentes (dénutrition, diabète, cirrhose, insuffisance rénale, cancers, etc.) (Edwards and Baker, 2005).

Parmi les 10 génotypes capsulaires décrits à ce jour (Ia, Ib, II et-IX), le sérotype III est d'une importance particulière parce qu'il est responsable d'une proportion importante d'infections précoces (« *Early Onset Disease* » - EOD, < 7 jours) et de la plupart des cas d'infections tardives (« *Late Onset Disease* » - LOD, 7-90 jours) (Simonsen *et al.*, 2014) (Figure 11). En revanche, le séquence type « hypervirulent » ST-17 est responsable en France de 60% des infections néonatales à *S. agalactiae* et de près de 80% des syndromes tardifs et des méningites, que ces dernières soient précoces ou tardives (Six *et al.*, 2014). La colonisation maternelle est une condition préalable à l'EOD.

a- Infection chez le nouveau-né

Chez le nouveau-né, l'EOD se manifeste sous forme de pneumonie qui évolue rapidement en bactériémie (Cowgill *et al.*, 2003). En revanche, la LOD se caractérise par une bactériémie, avec un risque élevé d'évolution en méningite. Les protocoles de prévention par antibioprophylaxie en per-partum largement mis en œuvre dans de nombreux pays, notamment en France, ont permis de réduire significativement l'incidence des EOD, mais l'incidence des LOD est restée inchangée (Joubrel *et al.*, 2015). La porte d'entrée de la bactérie dans l'organisme du nouveau-né est pulmonaire pour les infections précoces et à priori digestive pour les infections tardives. Dans ce dernier cas, au moment de l'accouchement, lors de la traversée des voies génitales de sa mère, l'enfant colonise la première fois son tube digestif, avec les bactéries dont sa mère est porteuse dans

son vagin et dans son rectum. Ce mode de transmission verticale, en per-partum, se fait le plus souvent par ingestion de liquide amniotique contaminé, ou lors du passage au travers des muqueuses vaginale colonisées (Tazi *et al.*, 2010) (Figure 11). Dans le cas de l'EOD, la transmission verticale entre la mère et le nouveau-né se produit dans 30 à 70% des cas, soit environ 50% des femmes avec un portage vaginal positif donnent naissance à des nouveau-nés colonisés (Quentin, 2002). En l'absence de toute intervention, on estime que 1% à 2% des nouveau-nés colonisés vont développer une infection précoce à streptocoque du groupe B (Verani *et al.*, 2010). La transmission horizontale, LOD ou post-partum, peut être maternelle (transmission par le lait, manu-portage maternel), ou beaucoup plus rarement à partir d'autres nouveau-nés par l'intermédiaire du personnel soignant (Tazi *et al.*, 2015) (Figure 11).

b- Infection chez la femme enceinte

Le portage vaginal de *S. agalactiae* est le plus souvent asymptomatique chez 6,5 à 36% des femmes enceintes. La manifestation la plus fréquente (50 % des cas) est l'infection du haut de l'appareil génital (placenta et cavité amniotique) qui conduit à la mort fœtale *in utero* (Phares *et al.*, 2008). En *per-partum*, *S. agalactiae* peut entraîner une chorioamniotite, une rupture prématurée des membranes voire un avortement (Baker and Edwards, 1990; Phares *et al.*, 2008). En *post-partum* et *post-abortion*, *S. agalactiae* est responsable de chorioamniotites et d'endométrites, accompagnées parfois de bactériémie (Krohn *et al.*, 1999).

c- Infection chez l'adulte, en dehors de la femme enceinte

Bien que l'infection à *S. agalactiae* puisse survenir chez des adultes en bonne santé, la majorité des cas survient chez les patients qui ont des pathologies sous-jacentes (dénutrition, diabète, cirrhose, insuffisance rénale, cancers) (Edwards and Baker, 2005). La prévalence du portage recto-vaginal est d'environ 25% chez les jeunes adultes en bonne santé hors grossesse (van der Mee-Marquet *et al.*, 2008). Elle ne diffère donc pas de façon significative de celle observée chez les femmes enceintes (Le Doare and Heath, 2013). Les infections nosocomiales à *S. agalactiae* représentent entre 8% et 20% des infections invasives de l'adulte (Salloum *et al.*, 2010).

Tableau 2 : Facteurs de virulence chez *Streptococcus agalactiae* d'après (Rajagopal, 2009)

Facteurs de virulence	Mode d'action	Mécanisme de régulation
Toxines porogènes		
β-hémolysine/cytolysine (β-H/C, CylE)	Favorise l'invasion et déclenche la lyse de la cellule hôte Perturbe la fonction cardiaque et le foie Induit les réponses inflammatoires et l'apoptose	Inhibé par CovR/CovS Activé par Stk1 à travers CovR Activé par RovS indépendamment de CovS/CovR et Stk1
Facteur CAMP (Cfb)	Forme des pores dans la membrane de la cellule hôte Se lier aux protéines GPI	Activé par CovR/CovS Inhibé par Stk1 à travers CovR
Facteurs d'évasion immunitaire		
L'acide sialique polysaccharide capsulaire (CPS)	Empêche la reconnaissance de l'ABG par mimétisme moléculaire de la surface glycoconjuguée de la cellule hôte	Inhibé par RogB
Superoxide dismutase (SodA)	Détoxifie singulet oxygène et superoxyde	Activé par RovS
Pigment (rhamno-polyene)	Détoxifie l'oxygène singulet et le superoxyde	Inhibé par CovR/CovS
C5a peptidase (ScpB)	Empêche le recrutement des neutrophiles dû(?) au clivage du complément C5a Favorise l'adhérence en se liant à la fibronectine de l'ECM (<i>extracellular matrix</i>) et aux cellules épithéliales	Inhibé par RgfC/RgfA et CovR/CovS
Sérine protéase (CspA)	Clive le fibrinogène et chimiokines Bloque le recrutement des neutrophiles	Inconnu
Résistance aux PAM		
Alanylation des acides lipotéichoïques	Diminue la charge négative nette sur la surface cellulaire, repousse les PAM	Activated by DltR/DltS
Protéine de liaison à la pénicilline 1a (PBP1a)	Favorise la résistance aux PAM à travers un mécanisme inconnu	Inconnu
Pili	Favorise la résistance aux PAM à travers un mécanisme inconnu Favorise également l'adhésion de <i>S. agalactiae</i> aux cellules hôtes	Expression de PI-2a est activée par RogB
Adhérence et invasion des cellules hôtes		
<i>Fibrinogen-binding protein A</i> (FbsA)	Favorise l'adhérence de SGB aux cellules hôtes par liaison au fibrinogène de l'ECM	Inhibé par RovS
<i>Fibrinogen-binding protein B</i> (FbsB)	Favorise l'entrée de SGB dans la cellule hôte	Inhibé par MtaR
<i>Laminin-binding protein</i> (Lmb)	Favorise l'adhérence de SGB aux cellules hôtes par liaison à la laminine de l'ECM	Inconnu
<i>Serine-rich repeat proteins</i> (Srr)	Favorise l'adhérence de SGB à la kératine humaine et aux cellules épithéliales (HEp2)	Inconnu
<i>Immunogenic bacterial adhesin</i> (BibA)	Favorise l'adhérence de SGB aux cellules hôtes par liaison à la protéine C4bp	Inconnu
Protéine αC	Facilite l'adhérence de SGB aux cellules épithéliales	Inconnu
<i>Invasion-associated gene</i> (IagA)	Permet l'ancrage des acides lipotéichoïques à la membrane; important pour l'invasion de la barrière hématoencéphalique	Inconnu
D'autres facteurs de virulence		
<i>Hyaluronate lyase</i>	Clive le hyaluronate et favorise la propagation de SGB lors de l'infection	Inconnu

B. Facteurs de virulence de *Streptococcus agalactiae*

Le développement d'une infection à *S. agalactiae* repose sur la réussite de la colonisation bactérienne de l'épithélium vaginal, la pénétration du placenta et/ou des barrières épithéliales du nouveau-né et la résistance et l'échappement au système immunitaire. Afin de contourner ces obstacles, *S. agalactiae* exprime différents facteurs de virulence à la surface cellulaire et en sécrète d'autres dans le milieu extracellulaire (Table 2).

1. Facteurs d'adhérence et de colonisation

L'infection néonatale débute avec la colonisation du tractus génital de la femme enceinte. *S. agalactiae* possède des facteurs lui permettant d'adhérer à une variété de cellules humaines, telles que les cellules épithéliales vaginales et pulmonaires, les cellules de la barrière foeto-placentaire et celles formant l'endothélium de la barrière hémato-encéphalique. Ces facteurs d'adhérence reposent sur des protéines hydrophobes à la surface de la bactérie, en particulier les protéines FbsA, ScpB et Lra1 liant respectivement le fibrinogène (Rosenau *et al.*, 2007), la fibronectine (Beckmann *et al.*, 2002), la laminine (Lindahl *et al.*, 2005) et les protéines BibA et ssr-1 décrites comme des molécules d'adhérence aux cellules de l'épithélium pulmonaire et vaginal, respectivement (Figure 12). De plus, *S. agalactiae* possède des pili qui joueraient un rôle dans l'adhérence aux cellules de la barrière hémato-encéphalique (Konto-Ghiorgi *et al.*, 2009; Papasergi *et al.*, 2011). Une étude récente a montré qu'une protéine ancrée à la surface de la bactérie, HvgA pour « *hypervirulent GBS adhesin* », chez les ST-17 est cruciale pour l'adhérence aux cellules épithéliales intestinales, aux cellules endothéliales dans les plexus choroïdes et les cellules endothéliales micro-vasculaires qui constituent la barrière hémato-encéphalique (Tazi *et al.*, 2010).

2. Facteurs d'invasion cellulaire

La capacité de *S. agalactiae* à l'invasion cellulaire corrèle au potentiel de virulence de chaque souche. L'invasion cellulaire est médiée par les composants de cytosquelette *via* la voie de signalisation de la PI3 Kinase (Burnham *et al.*, 2007) et par différentes protéines de

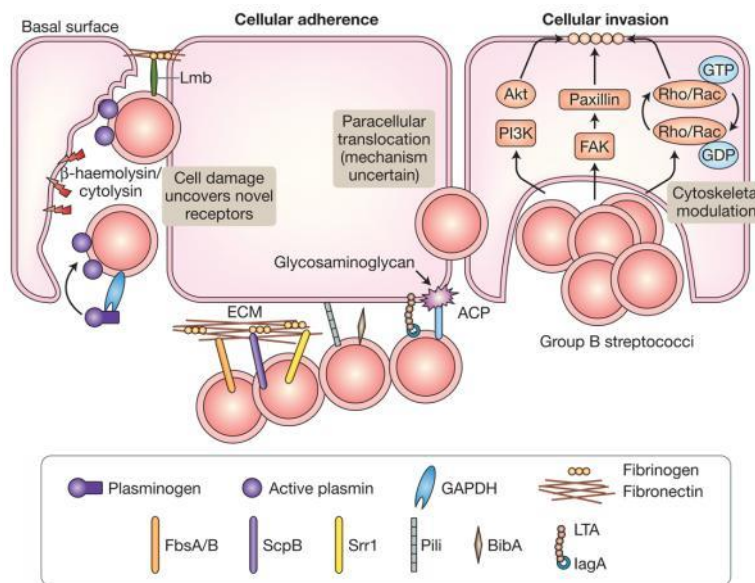


Figure 12 : Mécanismes d'adhérence et d'invasion cellulaire chez *Streptococcus agalactiae*.

ACP, *alpha C protein*; BibA, *GBS immunogenic bacterial adhesin*; ECM, *extracellular matrix*; FAK, *focal adhesion kinase*; FbsA/B, *fibrinogen-binding proteins A and B*; GAPDH, *glyceraldehyde 3-phosphate dehydrogenase*; GDP, *guanosine diphosphate*; GTP, *guanosine triphosphate*; Lmb, *laminin-binding protein*, LTA, *lipoteichoic acid*; PI3K, *phosphoinositide 3-kinase*; ScpB, *C5a peptidase*; Srr1, *serine-rich repeat domain protein 1*.

D'après (Maisey *et al.*, 2008)

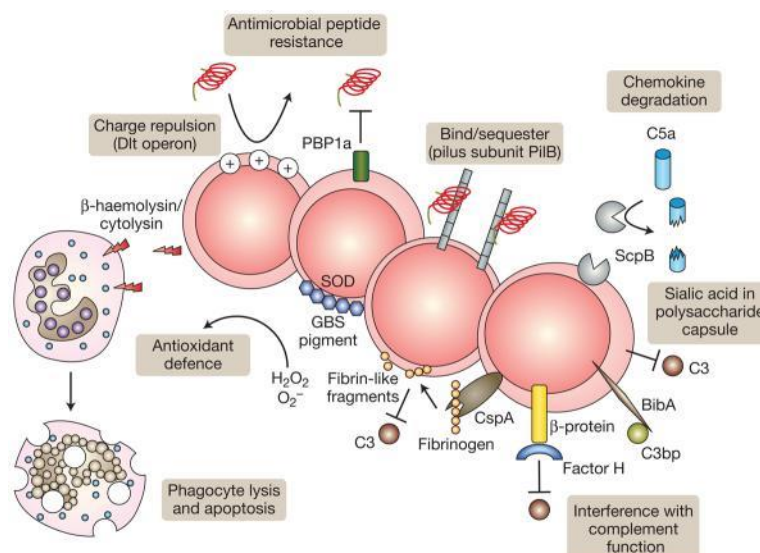


Figure 13 : Mécanismes d'échappement au système immunitaire chez *Streptococcus agalactiae*.

BibA, *GBS immunogenic bacterial adhesin*; CspA, *cell-surface protease A*; PBP1a, *penicillin-binding protein 1a*; ScpB, *C5a peptidase*; SOD, *superoxide dismutase*.

D'après (Maisey *et al.*, 2008)

surface telles que la protéine Spb1 (Adderson *et al.*, 2003), la protéine ScpB (Cheng *et al.*, 2002), l' α -protéine C (Bolduc *et al.*, 2002) ou la protéine FbsB (Gutekunst *et al.*, 2004) (Figure 12). Après pénétration dans la cellule, *S. agalactiae* entraîne la lyse cellulaire par effet cytolytique de la β -hémolysine/cytolysine et du facteur CAMP qui forme des pores dans les cellules épithéliales du poumon et les cellules endothéliales de la barrière hémato-encéphalique (Gibson *et al.*, 1999; Lang and Palmer, 2003). Le facteur HvgA est également impliqué dans la colonisation intestinale et le franchissement de la barrière hématoencéphalique (Tazi *et al.*, 2010).

3. Echappement au système immunitaire

Une fois que *S. agalactiae* a pénétré à travers les barrières cellulaires pour atteindre la circulation sanguine ou les tissus profonds, une large réponse immunitaire est activée pour éliminer l'infection, dans laquelle les cellules phagocytaires de l'hôte y compris les neutrophiles et les macrophages jouent un rôle crucial. La réponse immunitaire de l'hôte est conditionnée par les opsonines et ses cellules phagocytaires qui coopèrent à l'élimination des bactéries. Les IgG et le complément contribuent au mécanisme d'opsonisation. Une déficience combinée des défenses immunitaires humorales et cellulaires chez le nouveau-né contribue à sa susceptibilité accrue pour les infections à *S. agalactiae* (Kallman *et al.*, 1998).

En outre, *S. agalactiae* échappe à la réponse immunitaire de l'hôte, avec implication de la β -hémolysine/cytolysine (Liu *et al.*, 2004) et des protéines de surface qui inhibent les défenses de l'hôte en fixant les immunoglobulines et le complément, telles la β -protéine (Jerlstrom *et al.*, 1996), le facteur CAMP (Lang and Palmer, 2003), la protéine ScpB (Cheng *et al.*, 2002), la protéine CspA (Quach *et al.*, 2009), la protéine BibA (Santi *et al.*, 2009) et la protéine FbsA (Rosenau *et al.*, 2007)(Figure 13). *S. agalactiae* est également capable de se protéger des peptides antimicrobiens tels que les défensines et cathélicidines par l'incorporation de D-alanine au niveau des ALT (Saar-Dover *et al.*, 2012) et par la protéine de liaison aux pénicillines PBP1a (Hamilton *et al.*, 2006).

III. LES SYSTEMES DE REGULATION CHEZ *STREPTOCOCCUS AGALACTIAE*

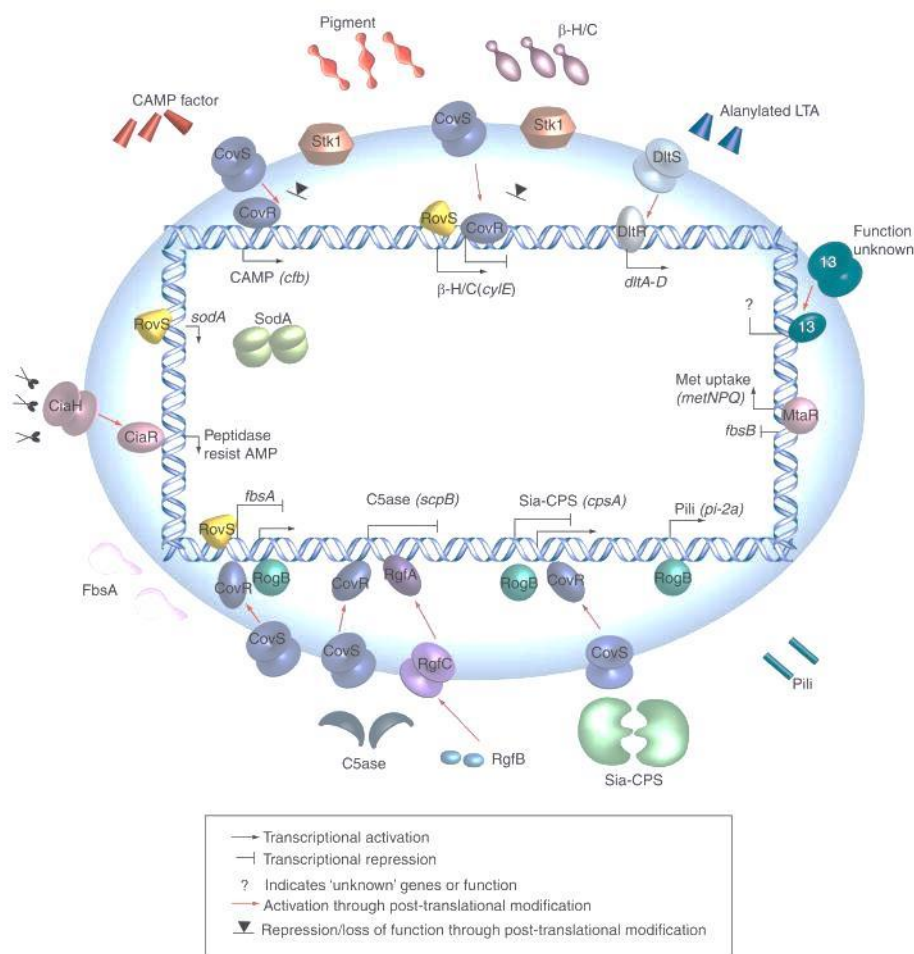


Figure 14 : Principaux régulateurs de facteurs de virulence chez *Streptococcus agalactiae*.

Les systèmes à deux composants formés par les régulateurs de réponse CovR, RgfA, CiaR et DltR et par les senseurs correspondants CovS, RgfC, CiaH et DltS permettent de contrôler l'expression des gènes impliqués dans l'adhérence, l'invasion cellulaire ainsi que dans l'évasion immunitaire par *S. agalactiae*; le système de signalisation eucaryote formé par la sérine thréonine kinase Stk1 et la phosphatase correspondante Stp1; et trois régulateurs transcriptionnels (simple composant) RogB, RovS et MtaR. Treize autres TCS restent de fonction inconnue ; β-H/C: β-hémolysine/cytolysine; AMP: peptide antimicrobien; C5ase: C5a peptidase; FbsA: protéine A liant le fibrinogène; LTA: acide lipotéichoïque; Met: méthionine; Sia-CPS: acide sialique capsulaire; SodA: Superoxyde dismutase.

D'après (Rajagopal, 2009)

III. Les systèmes de régulation chez *Streptococcus agalactiae*

S. agalactiae rencontre des conditions environnementales variées et sa survie nécessite une capacité à percevoir les changements du milieu externe et une capacité d'adaptation à ces différents environnements. La bactérie module l'expression des gènes impliqués dans le métabolisme et la virulence (l'adhérence, l'invasion, la résistance à la réponse immunitaire) suite à des variations de température, du pH, de nutriments disponibles ou du niveau d'oxygène. Chez les procaryotes, la régulation de l'expression des gènes, en réponse à l'environnement extérieur, est accomplie par les systèmes de transduction de signal (STS). Ces systèmes réagissent à un signal externe et le plus souvent régulent la fonction de facteurs de transcription. Parmi les STS les mieux caractérisés se trouvent les systèmes à deux composants (TCS). Les TCS sont constitués d'un senseur à fonction histidine kinase et d'un régulateur cytoplasmique (Stock *et al.*, 2000). Les mécanismes de communication entre le senseur et le régulateur impliquent des réactions de phosphorylation et de déphosphorylation. L'histidine kinase ancrée à la membrane perçoit les *stimuli* extérieurs et active par phosphorylation le régulateur de réponse. Une fois activé, le régulateur va se fixer grâce à son domaine de liaison avec l'ADN de manière spécifique et régule ainsi la transcription des gènes cibles.

Des études ont montré que le génome de *S. agalactiae* héberge entre 17 et 20 TCS (Glaser *et al.*, 2002; Tettelin *et al.*, 2005). Cependant, le rôle de seulement quatre TCS, CovRS (Lamy *et al.*, 2004), DltRS (Poyart *et al.*, 2001), RgfCA (Spellerberg *et al.*, 2002) et CiaRH (Quach *et al.*, 2009) a été étudié (Figure 14). Glaser et ses collaborateurs ont montré que la souche *S. agalactiae* NEM316 renferme 20 histidine kinases et 21 régulateurs cytoplasmiques, organisés en 20 TCS et un régulateur "orphelin" (Glaser *et al.*, 2002). *S. agalactiae* encode également au moins six systèmes de régulation à simple composant (*Stand-alone Systems*). Cette famille de régulateurs transcriptionnels permet l'activation ou l'inhibition des gènes cibles en réponse à un changement dans le cytosol (Ulrich *et al.*, 2005) ou du milieu extracellulaire grâce à des senseurs non identifiés (Kreikemeyer *et al.*, 2003). La fonction de trois de ces six régulateurs, MtaR (Shelver *et al.*, 2003), RogB (Gutekunst *et al.*, 2003) et RovS (Samen *et al.*, 2006) a été étudiée. On distingue également les facteurs sigma (σ) et les systèmes de régulation de type eucaryote comme systèmes de régulation transcriptionnelle.

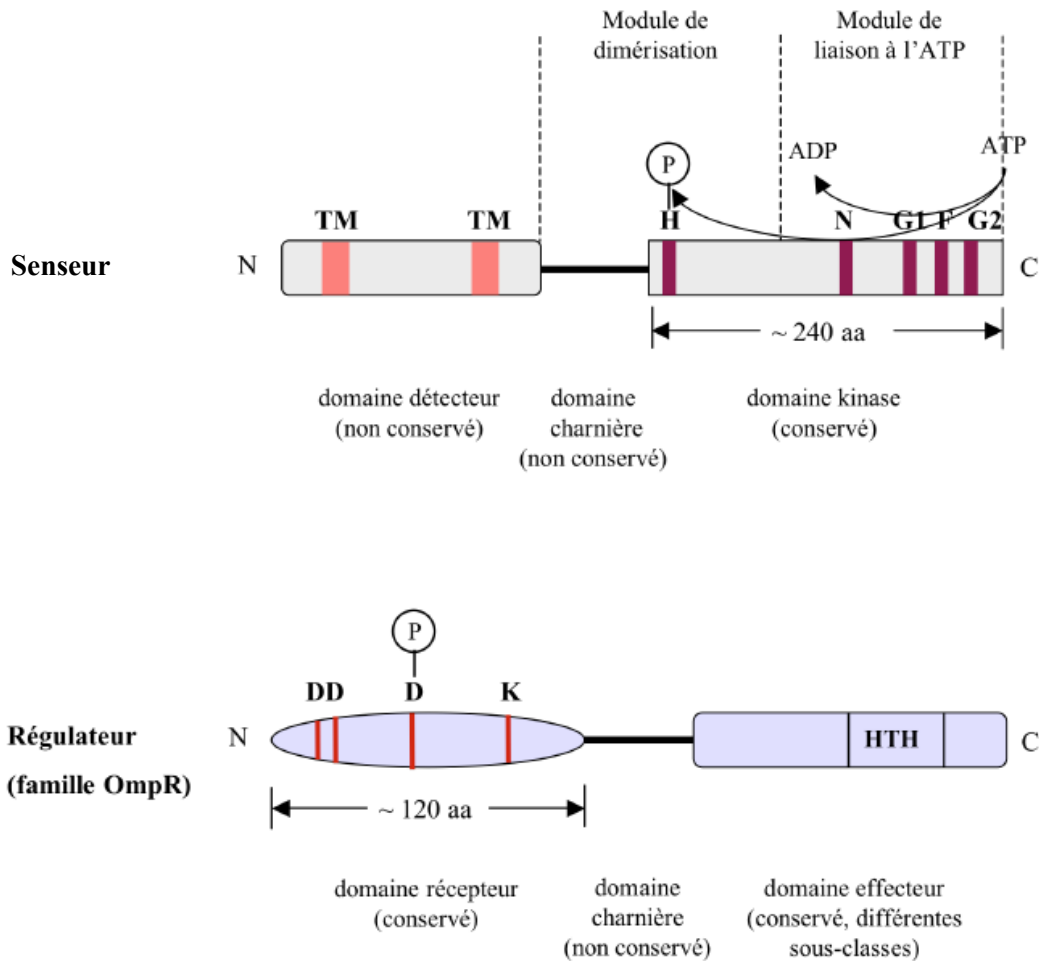


Figure 15 : Structure des systèmes de régulation à deux composants (TCS).

L'histidine kinase agit sous forme d'homodimères et est organisée en domaines fonctionnels: le domaine senseur ou détecteur N-terminal, variable d'une histidine kinase à une autre et comprenant généralement deux segments transmembranaires (TM) et le domaine transmetteur ou domaine kinase C-terminal composé de deux modules conservés: le module de dimérisation (contenant la boîte ou résidu H) et le module de liaison à l'ATP (contenant les boîtes N, G1, F et G2). Le domaine charnière lie les deux domaines. Le régulateur cytoplasmique comporte un résidu Asp (D), résidu cible du phospho-transfert, invariablement retrouvé en position N-terminale du domaine receveur de la protéine (vers le 50ème résidu). La phosphorylation de ce résidu Asp va influencer l'activité du domaine effecteur C-terminal présentant dans le cas des régulateurs de la famille OmpR, un motif de fixation à l'ADN de type « winged Helix-Turn-Helix » (HTH).

D'après (Stock *et al.*, 2000)

A. Systèmes de régulation à deux composants

Les TCS permettent la régulation de l'expression des gènes de virulence au sein de compartiments cellulaires infectés en réponse à des *stimuli* extra ou intracellulaires. Ils ont été impliqués dans l'adaptation métabolique à des changements de sources de carbone, d'azote et de phosphate, dans l'adaptation à des conditions de stress, dans le chimiotactisme, la sporulation et la pathogénicité (Stock *et al.*, 2000).

1. Le système CovS/CovR

Le TCS CovS/CovR (*control of virulence*), appelé également CsrS/CsrR (*capsule synthesis regulator*), a été décrit chez plusieurs espèces de streptocoques, notamment chez *S. pyogenes* (Graham *et al.*, 2002; Trevino *et al.*, 2009). Chez *S. agalactiae*, le locus *cov* code un régulateur cytoplasmique CovR de la famille OmpR/PhoB et une histidine kinase membranaire CovS (Figure 15). Le TCS CovS/R est un système pléiotropique qui régule la transcription d'approximativement 7 % du génome et fonctionne essentiellement comme un répresseur transcriptionnel (Lamy *et al.*, 2004). Initialement décrit comme inhibiteur de l'opéron *has* impliqué dans la synthèse de la capsule chez *S. pyogenes* (Levin and Wessels, 1998), le TCS CovS/R est également impliqué dans la régulation de nombreux gènes de virulence tels que l'opéron *cyf* responsable de la synthèse de l'hémolysine, le gène *cfb* codant le facteur CAMP ou encore le gène *scpB* codant la C5a peptidase ScpB (Lamy *et al.*, 2004; Jiang *et al.*, 2005). En outre, ce système de régulation permet la régulation des gènes codant des systèmes de transport de sucres, des acides-aminés, des gènes de réponse au stress oxydatif (Jiang *et al.*, 2005; Jiang *et al.*, 2008; Santi *et al.*, 2009), ou encore dans la résistance au peptide antimicrobien cationique humain LL-37 chez *S. pyogenes* (Froehlich *et al.*, 2009).

L'inactivation de locus *cov* dans la souche *S. agalactiae* NEM316 a conduit à l'obtention d'un mutant $\Delta covRS$ hyper-hémolytique et hyper-adhérent aux cellules épithéliales humaines (A549 et HeLa) et murines (Lamy *et al.*, 2004). Cependant, ce mutant était incapable de croître dans le sang humain et sa virulence était fortement altérée dans un modèle de sepsis du rat nouveau-né infecté par voie intrapéritonéale (Lamy *et al.*, 2004), soulignant le rôle primordial de ce système de régulation dans la pathogénicité. Par ailleurs,

il a été suggéré que le TCS CovS/R participerait à l'adaptation de la bactérie lors de la translocation d'un milieu acide tel que le milieu vaginal vers un milieu neutre comme le sang ou les poumons (Santi *et al.*, 2009). L'analyse comparative du transcriptome de la souche sauvage NEM316 et son mutant a montré que ce système de régulation activerait 63 gènes tandis qu'il agirait comme répresseur pour 76 autres (Lamy *et al.*, 2004). Dans la même étude, Lamy et ses collaborateurs ont identifié la séquence consensus (TATTTTAAT) située dans les régions régulatrices des gènes et à laquelle le régulateur CovR se fixerait.

Le système CovS/R présente de plus la particularité d'être lui-même régulé par un système de transduction sérine/thréonine kinase et sérine/thréonine phosphatase (Stk1/Stp1) de type eucaryote (Rajagopal *et al.*, 2006). Ce système est impliqué dans une large gamme de processus cellulaires allant du métabolisme à la réponse au stress et à la virulence de *S. agalactiae* (Rajagopal *et al.*, 2003). En effet, il a été montré que Stk1 phosphoryle la thréonine en position 65 (T65) du régulateur de réponse CovR entraînant ainsi une réduction de son activité. Enfin, la variation des régulons CovR est susceptible de jouer un rôle important dans l'adaptation des souches aux changements environnementaux et pourrait ainsi expliquer la différence de potentiel pathogène de différentes souches.

2. Le système CiaR/CiaH

Le système de régulation CiaR/CiaH a été identifié pour la première fois chez l'espèce *S. pneumoniae* (Guenzi *et al.*, 1994). Le régulateur CiaR de *S. agalactiae* appartient à la famille OmpR et présente 89 % d'identité peptidique avec son homologue chez *S. pneumoniae* (Quach *et al.*, 2009). Des mutations dans le locus *ciaRH* sont associées avec une augmentation de la résistance aux antibiotiques, tel que le céfotaxime (Giammarinaro *et al.*, 1999) ; ce phénotype est probablement associé à la régulation des gènes impliqués dans la synthèse de la paroi cellulaire. Le système CiaR/CiaH contribue également à la compétence par la régulation de l'opéron *comCDE* (Giammarinaro *et al.*, 1999), l'autolyse, la tolérance au stress oxydatif et à la virulence (Zahner *et al.*, 2002). Chez *S. agalactiae*, le régulateur CiaR est impliqué dans la survie aux polynucléaires neutrophiles, aux macrophages et dans les cellules endothéliales des capillaires cérébraux (Quach *et al.*, 2009). Le mutant Δ *ciaR* est plus rapidement éliminé dans un modèle murin d'infection systémique que la souche sauvage. Par ailleurs, des mutations dans le senseur CiaR ont permis

l'identification de 120 gènes dont la régulation est Cia-dépendante chez *S. pyogenes* (Riani *et al.*, 2007). Parmi ceux-ci, des gènes codant des protéines de transport de cations, de système PTS (*Phosphoenolpyruvate: carbohydrate phosphoTransferase System*), des hémolysines, des hyaluronidase et des DNases.

3. Le système DltS/DltR

Comme son nom l'indique, le TCS DltS/DltR contrôle l'expression des gènes de l'opéron *dltABCD* impliqués dans l'incorporation des résidus D-Alanyl dans les acides lipotéichoïques (ALT) (*cf.* I.B.3). Décrit pour la première fois chez *S. agalactiae*, le locus *dlt* code un régulateur cytoplasmique DltR appartenant à la famille des OmpR (Poyart *et al.*, 2001). Le gène *dltS* code une histidine kinase, dont la fonction est probablement la transduction du signal du milieu extracellulaire vers le DltR situé dans le cytoplasme. L'utilisation de fusions transcriptionnelles a permis de montrer, en présence de DltS, que DltR active la transcription de son propre promoteur. Cette analyse a montré également que les gènes de l'opéron *dltABCD* ne nécessitaient pas la présence de DltR et DltS pour leur expression. Les auteurs suggèrent que le rôle de ce TCS serait plutôt le maintien d'un niveau approprié de D-Alanine dans les ALT, en réponse à des changements dans l'environnement. Le signal capté par DltS reste inconnu.

4. Le système RgfC/RgfA

Parmi les treize autres systèmes conservés dans le génome de *S. agalactiae*, le TCS RgfC/RgfA est codé par l'opéron *rgfBDAC* (*regulator of fibrinogen binding*) (Spellerberg *et al.*, 2002). Ce TCS est composé d'une histidine kinase (RgfC) et d'un régulateur (RgfA) appartenant à la famille LytR. Cette famille de régulateurs contient à l'extrémité N-terminale un domaine de phosphorylation et un domaine de liaison à l'ADN en C-terminal. L'inactivation du gène *rgfC* chez une souche de *S. agalactiae* entraîne une réduction significative de sa liaison au fibrinogène (Spellerberg *et al.*, 2002). De plus, la transcription du gène *scpB*, codant une C5a peptidase, est augmentée dans le mutant *rgfC* (Spellerberg *et al.*, 2002). Le mécanisme exact de cette régulation par RgfC reste à identifier.

B. Régulateurs transcriptionnels

Les régulateurs transcriptionnels à simple composant (*stand-alone regulators*) sont des protéines se liant à l'ADN qui permettent d'activer ou de réprimer la transcription du gène cible, en réponse à un changement cytoplasmique ou du milieu extérieur. Ces régulateurs se fixent sur des séquences spécifiques (dites consensus) au niveau de la région promotrice du gène cible.

1. Le régulateur MtaR

Le régulateur MtaR (*methionine transport regulator*) appartient à la famille des régulateurs transcriptionnels LysR (Schell, 1993). Cette famille de régulateurs est constituée de protéines cytoplasmiques qui reconnaissent des signaux intracellulaires et qui possèdent à la fois un domaine senseur et un domaine effecteur de liaison à l'ADN (Schell, 1993). Chez *S. agalactiae*, le régulateur MtaR est impliqué dans la régulation du transport de la méthionine, la croissance bactérienne en plasma humain et dans la virulence (Shelver *et al.*, 2003). Ce régulateur active également la transcription du gène codant une sérine protéase (CspA) qui contribue à la virulence (Bryan *et al.*, 2008).

2. Le régulateur RogB

RogB (*Regulator of group B streptococcus*) appartient à la famille des régulateurs *RofA-like proteins* (RALP) et présente des similarités avec les régulateurs de la famille RofA (Fogg *et al.*, 1994). RogB possède un domaine hélice-tour-hélice (HTH) dans sa partie N-terminale permettant la liaison à l'ADN et présente 50,2% d'identité avec RofA et 49,9% d'identité avec Nra, un autre régulateur de la famille RofA présent chez *S. pyogenes*. Chez *S. agalactiae*, le gène *rogB* est situé à 284 nucléotides en aval du locus PI-2a, impliqué dans la biosynthèse des pili (Morello *et al.*, 2015). RogB contrôle positivement la transcription du gène *fbsA* (codant une protéine d'adhérence au fibrinogène) et régule négativement des gènes impliqués dans la synthèse du polysaccharide capsulaire (Gutekunst *et al.*, 2003). La délétion de RogB entraîne une diminution d'adhérence à la fibronectine, au fibrinogène et aux cellules épithéliales pulmonaires (A549) (Gutekunst *et al.*, 2003).

3. Le régulateur RovS

Le régulateur RovS (*Regulator of virulence in S. agalactiae*) est un homologue du régulateur Rgg identifié chez *Streptococcus gordonii*, chez lequel il régule l'expression d'une glycosyltransférase (Sulavik *et al.*, 1992). De même que RogB, RovS intervient dans l'adhérence de *S. agalactiae* au fibrinogène et aux cellules épithéliales pulmonaires (Samen *et al.*, 2006). L'analyse transcriptionnelle ainsi que des expériences de retard sur gel ont montré que RovS inhibe la transcription du gène *fbxA* et active celle des gènes *sodA* (codant la Superoxyde dismutase), *rogB* (codant le régulateur RogB) et *cyl* (β -hémolysine/cytolysine). RovS peut réguler directement l'expression de ces gènes ou *via* l'interaction avec d'autres régulateurs, tel que CovRS (Samen *et al.*, 2006). L'interaction directe se fait par la liaison à une séquence consensus (AWAAWVHTDAW-N_{6/7}-WTKWWAMDWAK) située dans les régions promotrices des gènes cibles. Une étude récente suggère également que RovS est impliqué dans la communication cellulaire et la pathogénicité de *S. agalactiae* *via* la co-régulation de SHP (short hydrophobic protein) et d'une protéine extracellulaire de fonction inconnue codée par le gène *gbs1556* (Perez-Pascual *et al.*, 2015). Ces différentes données suggèrent que RovS pourrait jouer un rôle majeur dans la virulence de *S. agalactiae*.

C. Les facteurs sigma

Les facteurs sigma (σ) sont des sous-unités de l'ARN polymérase holoenzyme qui lui permettent de reconnaître et de se fixer aux promoteurs des gènes. Il existe différents types de facteurs σ , les facteurs σ primaires et les facteurs σ alternatifs. Le facteur σ primaire est essentiel à la survie des bactéries et reconnaît des séquences consensus situées -35 (TTGACA) et -10 (TATAAT) nucléotides en amont du site d'initiation de la transcription. Les autres facteurs σ sont non essentiels à la survie cellulaire en conditions de croissance favorables et sont appelés facteurs σ alternatifs. Ces derniers sont impliqués dans des fonctions diverses comme par exemple la régulation de l'expression de gènes des cellules exposées à divers stress, la compétence, la production d'antibiotiques ou encore des facteurs de virulence (Mauri and Klumpp, 2014).

Chez *S. agalactiae*, il existe seulement trois facteurs σ : le facteur σ^A , un facteur primaire ; le facteur ComX, un facteur de compétence et le facteur σ^{ECF} , présent chez

d'autres bactéries à Gram positif (Glaser *et al.*, 2002; Tettelin *et al.*, 2005).

D. Les systèmes de signalisation de type eucaryote

L'espèce *S. agalactiae* code deux systèmes de signalisation de type eucaryote (Glaser *et al.*, 2002; Tettelin *et al.*, 2005). Le premier est constitué d'une tyrosine kinase (CspD) et de la phosphatase correspondante (CspB). Le deuxième est constitué d'une sérine/thréonine kinase (Stk1) et de la phosphatase correspondante (Stp1) ; ces deux protéines sont impliquées dans plusieurs processus cellulaires allant du métabolisme, résistance aux neutrophiles, à la réponse au stress et au développement ainsi qu'à la virulence de *S. agalactiae* (Rajagopal *et al.*, 2003).

Les bactéries régulent l'expression des gènes en utilisant ces systèmes de régulation mais aussi des ARN régulateurs. Dans le prochain chapitre, nous allons vous présenter : quelles sont les caractéristiques qui distinguent ces régulateurs ? Quels sont les mécanismes de régulation ? Et quel est leur rôle dans le pouvoir d'adaptation des bactéries et notamment des streptocoques ?

IV. LES ARN REGULATEURS BACTERIENS

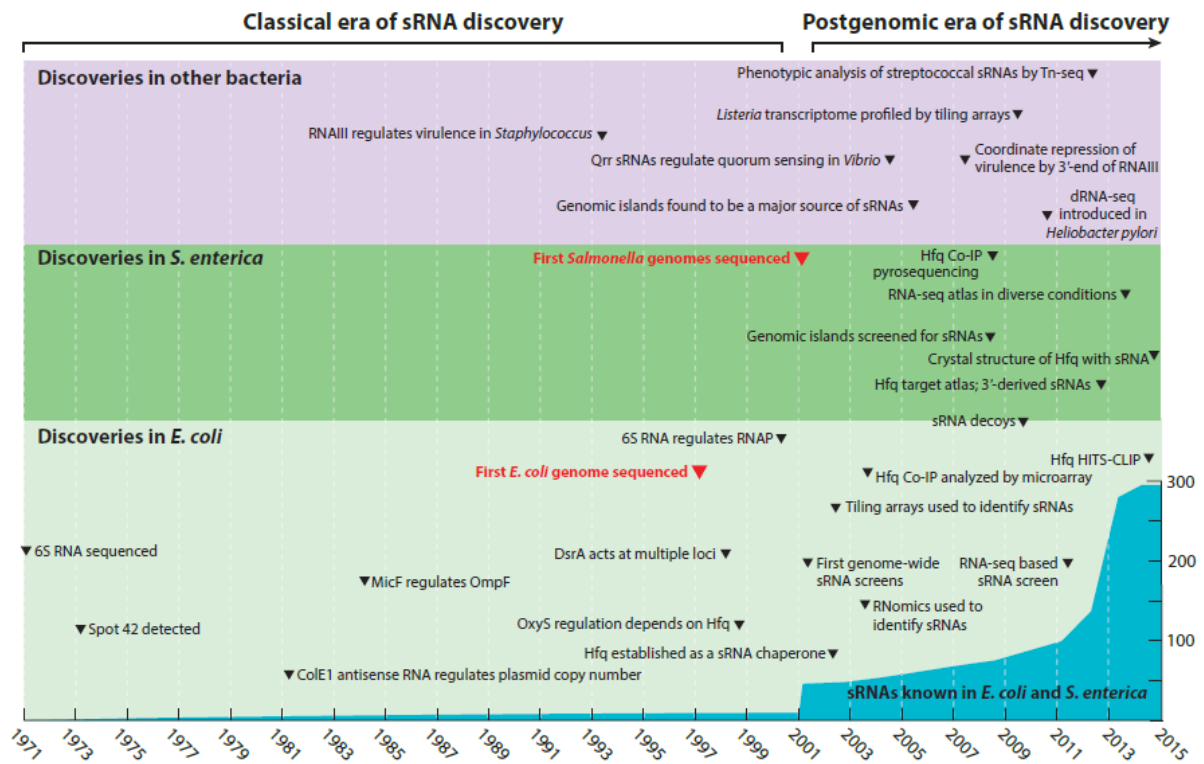


Figure 16 : Chronologie des études marquantes et les innovations technologiques dans l'étude des ARN régulateurs bactériens.

Co-IP, co-immunoprécipitation ; HITS-CLIP, séquençage à haut débit d'ARN isolé par réticulation immunoprécipitation.

D'après (Barquist and Vogel, 2015)

IV. Les ARN régulateurs bactériens

A. Découverte et caractéristiques des ARN régulateurs

Les bactéries pathogènes ont la capacité d'envahir les tissus de l'hôte pour établir un foyer d'infection. Le pathogène doit tout d'abord adhérer aux tissus *via* l'expression d'adhésines. Puis, l'expression de facteurs de virulence neutralise les défenses immunitaires. La synthèse de toxines détruit les cellules de l'hôte pour permettre l'accès à de nouvelles sources de nutriments et favoriser la dissémination des bactéries dans les tissus et les organes voisins. Cette machinerie infectieuse nécessite l'intervention d'un réseau de régulation spécifique à chaque étape. La découverte des ARN régulateurs (ARNrég), appelés également ARN non-codants (ARNnc) ou petits ARN (sRNA pour *small RNA*), et leur rôle dans la modulation de l'expression des gènes chez les bactéries, en réponse aux changements environnementaux, a permis d'améliorer nos connaissances sur ces réseaux de régulation et leur implication dans les processus infectieux.

Pendant de nombreuses années, la communauté scientifique a pensé que les ARN avaient pour seul rôle de servir de support intermédiaire à l'information génétique (ARNm) et d'assurer la synthèse protéique *via* les ARN ribosomiques (ARNr) et les ARN de transferts (ARNt). Cependant, dès les années 1960, le scientifique Hindley a été le premier à montrer que les ARN peuvent aussi avoir des fonctions dites « de ménage ou *housekeeping* » (Hindley, 1967). Ces ARN « de ménage » comprennent l'ARN M1 de la RNase P qui possède une fonction d'endoribonuclease sur l'extrémité 5' de la séquence précurseur de l'ARNt (Guerrier-Takada *et al.*, 1983) ; l'ARN 6S, qui inhibe la transcription en se liant directement à la forme holoenzyme de l'ARN polymérase et bloque son accès au promoteur ADN (Wassarman, 2007) ; ou encore l'ARNtm (RNAtm, *transfer-messenger RNA*) qui possède à la fois la fonction de l'ARNt et de l'ARNm et joue un rôle important dans le contrôle de la qualité de traduction (Withey and Friedman, 2003).

A partir des années 1980, de nombreuses équipes ont rapporté le rôle des ARNrég dans des phénomènes de régulation du nombre de copies de plasmide ou dans la maintenance de ces plasmides dans les bactéries hôtes (Stougaard *et al.*, 1981; Tomizawa and Itoh, 1981). En parallèle, Mizuno et ses collaborateurs s'interrogèrent sur un potentiel

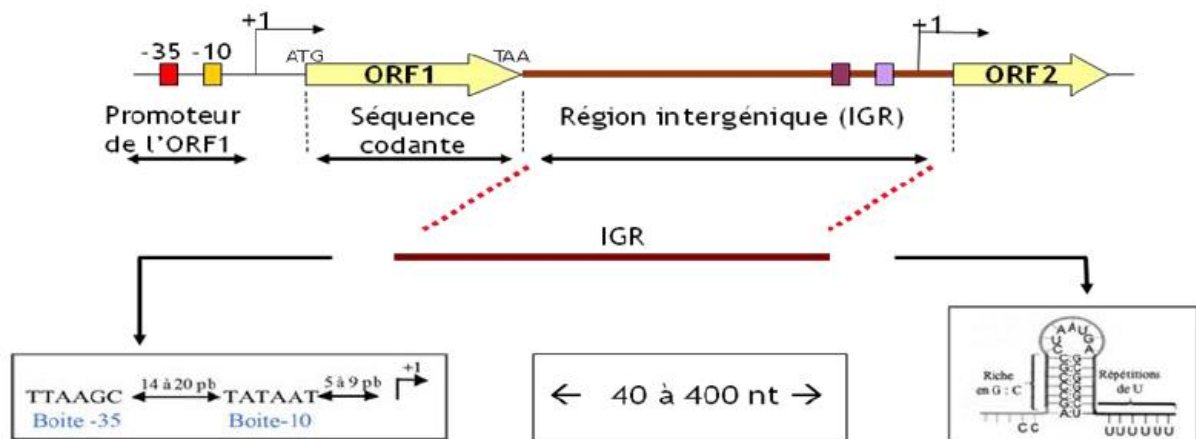


Figure 17 : Critères de prédiction des ARNrég par bio-informatique.

Les principales études bio-informatiques se sont concentrées sur les régions intergéniques (IGR) présentes entre deux ORF annotées. Les algorithmes ont été conçus pour détecter au sein de ces IGR des motifs promoteurs et des terminateurs rho-indépendants (formant une structure en tige-boucle), espacés par 40 à 400 nt en moyenne.

D'après (Reignier, 2010)

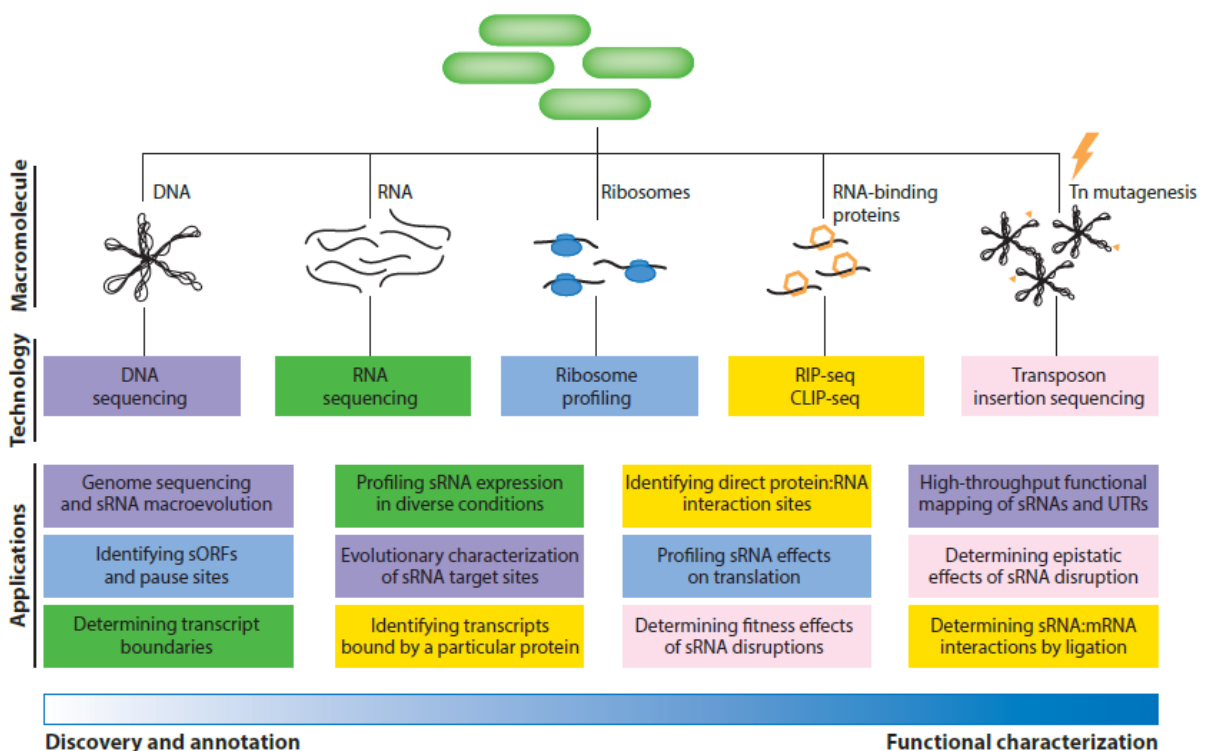


Figure 18 : Vue d'ensemble des méthodes à haut débit pour l'identification et la caractérisation des ARNrég bactériens.

CLIP, le « *cross-linking* » et immunoprécipitation ; RIP, immunoprécipitation d'ARN ; sORF, petit cadre de lecture ouvert ; UTRs, région non-traduite.

D'après (Barquist and Vogel, 2015)

rôle de ces ARNrég dans la régulation des gènes codés par les chromosomes bactériens (Mizuno *et al.*, 1984). En effet, ces chercheurs ont caractérisé un fragment d'ADN, exprimé en *trans* à partir d'un plasmide multi-copies, qui réprime l'expression de l'ARNm *ompF*. L'analyse de ce fragment d'ADN leur a révélé qu'il codait un ARNrég non-codant et que la répression de la synthèse d'OmpF était liée à l'hybridation de cet ARNrég, appelé MicF (pour « *mRNA-interfering complementary RNA* »), au site de fixation du ribosome (RBS ou *ribosome binding site*) de l'ARNm *ompF*, composé de la séquence de Shine-Dalgarno (SD) et du codon d'initiation de la traduction AUG.

Après cette découverte, il a fallu attendre 2001 pour que la vision de la communauté scientifique prenne un nouvel essor grâce à la découverte d'une nouvelle classe d'ARNrég exprimés dans les régions intergéniques (*intergenic region* ou IGR) des bactéries, principalement chez *Escherichia coli* (Figure 16) (Argaman *et al.*, 2001; Rivas *et al.*, 2001; Wassarman *et al.*, 2001). Ces travaux s'appuyaient sur des approches bioinformatiques de comparaison de génome, combinées à des tests d'expression *in vivo*. Trois caractéristiques étaient recherchées pour l'identification des ARNrég dans les IGR : (i) une taille entre 40 à 400 nucléotides (nts), (ii) la présence de terminateurs Rho-indépendants et de promoteurs orphelins, et (iii) la conservation de ces mêmes IGR dans des génomes de bactéries phylogénétiquement proches (Figure 17). L'autre caractéristique des ARNrég est leur structure secondaire. Les ARNrég possèdent généralement une proportion importante de bases complémentaires impliquées dans des interactions intra-moléculaires, permettant à la séquence ribonucléotidique primaire de se replier sur elle-même pour former des structures secondaires de type tige-boucle. Le niveau d'expression et la stabilité restent également des critères primordiaux pour l'activité régulatrice de ces ARNrég (Gottesman, 2005).

De nos jours, le développement de nouvelles techniques d'identification des ARNrég à grande échelle, tels que les méthodes à haut débit (Figure 18) ou les microarray, a permis la découverte de centaines de ces molécules codées par les génomes bactériens, qui représentent 3 à 5% du nombre total de gènes annotés quelles que soient les espèces étudiées (Siezen *et al.*, 2010).

B. Mécanismes et modes d'action

Ces ARNrég sont impliqués dans de très nombreux processus physiologiques bactériens, comme par exemple le métabolisme du carbone, l'homéostasie des métaux, la virulence, le quorum-sensing ou encore la réponse aux stress environnementaux tels que la température ou le pH (Vogel and Papenfort, 2006; Repoila and Darfeuille, 2009; Waters and Storz, 2009; Storz *et al.*, 2011). Chez les bactéries, les ARNrég peuvent être divisés en deux grandes classes : les éléments agissant en *cis* (*cis-acting* RNAs) et les éléments agissant en *trans* (*trans-acting* RNAs). Les ARNrég agissant en *cis* correspondent aux ARN qui agissent et/ou sont localisés aux extrémités 5' ou 3' non-traduites (*untranslated region* ou UTR) d'un ARNm (Gossringer and Hartmann, 2012; Ruiz de los Mozos *et al.*, 2013). La classe des ARNrég agissant en *trans*, d'origine plasmidique ou chromosomique, se divise en deux catégories : (i) les ARN antisens (ARNas) qui sont codés en *cis* sur le brin opposé de leur ARNm cible, avec laquelle ils présentent une complémentarité parfaite de séquence ; (ii) les ARN codés en *trans*, transcrits à partir d'un locus différent de leur cible (Zorgani *et al.*, 2016). A l'instar des ARNas, les ARN codés en *trans* peuvent activer/réprimer la traduction de leur(s) ARNm cible(s) ou bien affecter leur(s) stabilité (Gottesman and Storz, 2011). Les ARN codés en *trans* peuvent également former des complexes ribonucléo-protéiques (Storz *et al.*, 2011) ou encore participer à l'élimination des fragments d'ADN étrangers, cas des CRISPR RNAs (Clustered Regularly Interspaced Short Palindromic Repeats RNAs) (Karvelis *et al.*, 2013).

1. Les ARN régulateurs agissant en *cis*

Chez les bactéries, les éléments agissant en *cis* sont de véritables senseurs de l'état métabolique de la cellule. Sous des conditions spécifiques et afin de contrôler l'expression de leur gène respectif, ces ARNrég captent les signaux environnementaux tels que des métabolites (cas des *riboswitch*) (Breaker, 2012) ; les *T-box*, impliquées dans la liaison des ARNt et la modulation des besoins en amino-acyl ARNt synthétase (Green *et al.*, 2010) ; les ions métal et pH et ; les thermosenseurs capables de détecter les variations de température (Serganov and Nudler, 2013).

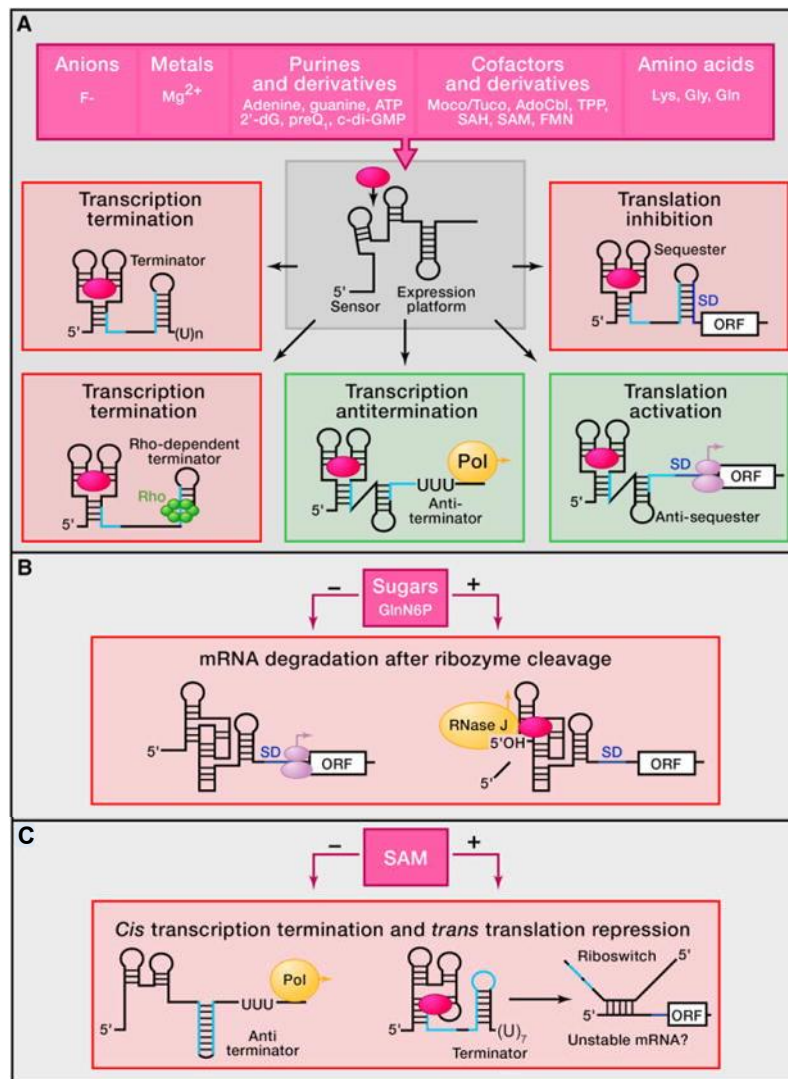


Figure 19 : Mécanisme d'action des *riboswitch*.

(A) Quand certains métabolites présents dans la cellule dépassent un seuil de concentration, ils peuvent être directement captés par le domaine senseur des *riboswitch*. La capture du métabolite (ligand) stabilise la conformation de senseur de *riboswitch* et induit généralement des changements structuraux dans la plateforme d'expression. Tandis que certains éléments structuraux stabilisent l'expression des gènes (carré vert), d'autres l'inhibent (carré rose). (B) La Glucosamine-6-phosphate (GlcN6P) induit le clivage du *riboswitch*-ribozyme *glmS* dans l'extrémité 5'UTR. Ainsi, le groupement 5'-OH stimule la dégradation du messager par la RNase J. (C) Chez *Listeria monocytogenes*, en absence du S-adenosyl-méthionine (SAM), les *riboswitch* SreA et SreB forment des tiges-boucles anti-terminatrices et permettent la transcription. La liaison du SAM provoque l'arrêt de la transcription. Ainsi, le fragment de l'ARNm se lie à la partie 5'UTR de l'ARNm cible et agit en *trans* pour déstabiliser ce dernier et réduire sa traduction.

D'après (Serganov and Nudler, 2013)

a- Les *riboswitch*

Le terme « *riboswitch* » reflète la capacité de ces ARN non-codants à fonctionner comme commutateurs génétiques. Parmi les métabolites auxquels les *riboswitch* sont capables de se fixer, il y a la flavine mononucléotide (FMN), la glycine, la lysine, le S-adénosyl-méthionine (SAM), la Glucosamine-6-phosphate (GlcN6P), l'adénosylcobalamine (AdoCbl), la S-adenosyl-L-homocystéine (SAH), l'adénine, la guanine, la pré-queuosine-1 (preQ1), le cyclic di-guanosyl-5'-monophosphate (c-di-GMP), les cofacteurs molybdenum et tungsten (Moco/Tuco), la thiamine pyrophosphate (TPP) et la vitamine B12 (Serganov and Nudler, 2013) (Figure 19).

Généralement, les *riboswitch* se composent de deux parties : i) la région dite aptamère (senseur), responsable de la reconnaissance spécifique du métabolite et ii) la plateforme d'expression qui permet la fixation du métabolite. Cette fixation conduit à des changements conformationnels, favorisant la formation d'un terminateur ou d'un antitermineur de transcription (Figure 19. A) qui favorise/bloque l'accès au site de fixation du ribosome (RBS) (Figure 19. A, B et C). D'autres structures plus complexes ont été décrites également, tel que le *riboswitch* en tandem chez *B. subtilis* (Mandal *et al.*, 2004). En outre, l'intervention de deux *riboswitch* indépendants pour le contrôle de l'expression du gène *tenA* chez *B. anthracis*, en fixant la TPP, permet de favoriser la formation d'un terminateur de transcription (Welz and Breaker, 2007).

Un mécanisme non conventionnel est utilisé par le *riboswitch*-ribozyme *glmS* qui couple la détection du métabolite et le clivage de l'ARNm (Figure 19. B). Ce type de *riboswitch* est fréquent chez les bactéries à Gram positif, il interagit avec la Glucosamine-6-phosphate (GlcN6P) et coupe l'ARNm *glmS* au sein même de la séquence de *riboswitch* (Winkler and Breaker, 2005). Ainsi, la RNase J dégrade le fragment de l'ARNm à partir de l'extrémité 5'-OH et inhibe sa traduction. De plus, ces ARN agissant en *cis* peuvent aussi avoir une action en *trans*. Ce mécanisme a été décrit chez *Listeria monocytogenes* pour les *riboswitch* SreA et SreB, qui fixent le S-adénosyl-méthionine (SAM) (Figure 19. C) (Loh *et al.*, 2009). Après la fixation du SAM et la terminaison de la transcription, ces *riboswitch* se lient à la partie 5'UTR de l'ARNm codant le facteur de virulence PrfA et inhibent sa traduction.

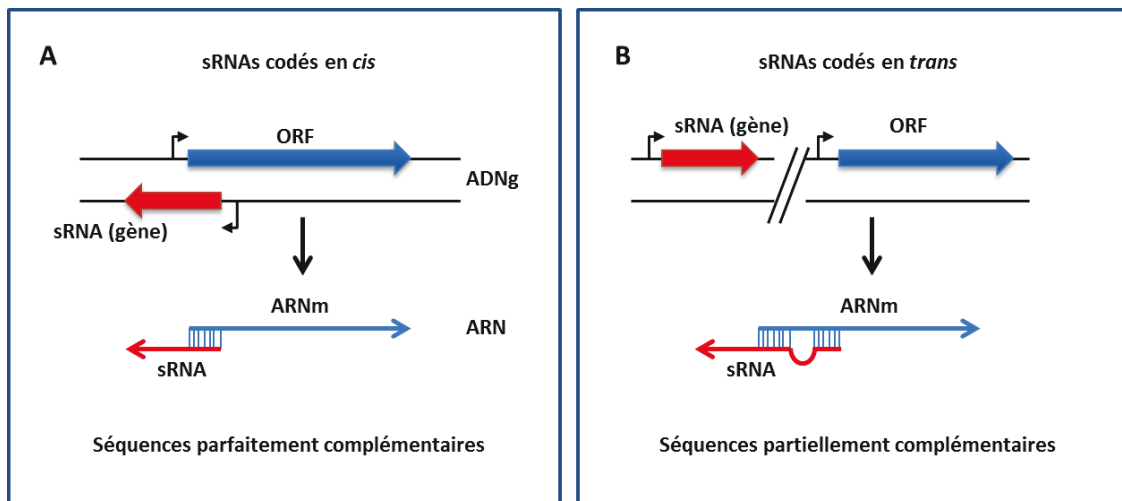


Figure 20 : ARNrég agissant en *trans*.

(A) ARNrég codés en *cis* ou asRNAs. (B) ARNrég codés en *trans*.

b- Les thermosenseurs

Outre la détection des métabolites, les ARN agissant en *cis* peuvent aussi répondre à des signaux environnementaux telle que la température. Les ARN thermosenseurs, appelés également thermomètres, constituent une classe unique des éléments agissant en *cis*. Ils ne nécessitent pas de domaine senseur, mais leur fonction est ajustée en réponse à un changement de température. En effet, le RBS est souvent bloqué à basse température (30°C) dans une structure secondaire qui est instable à plus haute température. Un exemple typique de ces ARN non-codants est le thermosenseur situé dans la région 5'UTR de l'ARNm *prfA* chez *L. monocytogenes* (Johansson *et al.*, 2002). *L. monocytogenes* produit le facteur de virulence PrfA uniquement à 37°C. A cette température, la structure secondaire qui piège le RBS se désagrège, rendant la traduction de l'ARNm *prfA* possible (Johansson *et al.*, 2002).

2. Les ARN régulateurs agissant en *trans* ciblant les ARNm

La très grande majorité des ARNrég caractérisés à ce jour agissent sur les ARNm par un mécanisme dit de type « antisens ». Ce mécanisme, à ne pas confondre avec les ARN antisens, met en jeu un appariement de bases de type Watson-Crick (A:T ; G:C) ou non canoniques (G:U) entre des régions simple brin antiparallèles portées par l'ARNrég et l'ARNm (Repoila and Darfeuille, 2009). De cette association ARNrég/ARNm (antisens/sens) résulte alors la formation d'un duplex bi-moléculaire qui conduit à une modification de l'efficacité de traduction et/ou de la stabilité de l'ARNm correspondant. En se basant sur la localisation du gène codant l'ARNrég, deux classes peuvent être distinguées. D'une part, le gène codant l'ARNrég est localisé sur le brin opposé d'un cadre ouvert de lecteur (*open reading frame* ou ORF), et par conséquent ces ARNrég, dits ARN antisens (ARNas), sont qualifiés de ARNrég codés en « *cis* » de leur cible d'ARNm (Figure 20. A). Dans ce cas, le couple ARNas/ARNm est strictement complémentaire. A l'inverse, les ARNrég codés par des régions intergéniques (IGR) ou par un autre locus que l'ARNm cible peuvent tolérer des mésappariements dans leur liaison (Figure 20. B). Ces ARNrég d'origine intergénique agissent quant à eux sur des ARNm cibles transcrits depuis des *loci* génomiques distants du gène de l'ARNrég. Dans ce cas, nous parlons d'ARNrég codés en « *trans* » (Waters and Storz, 2009).

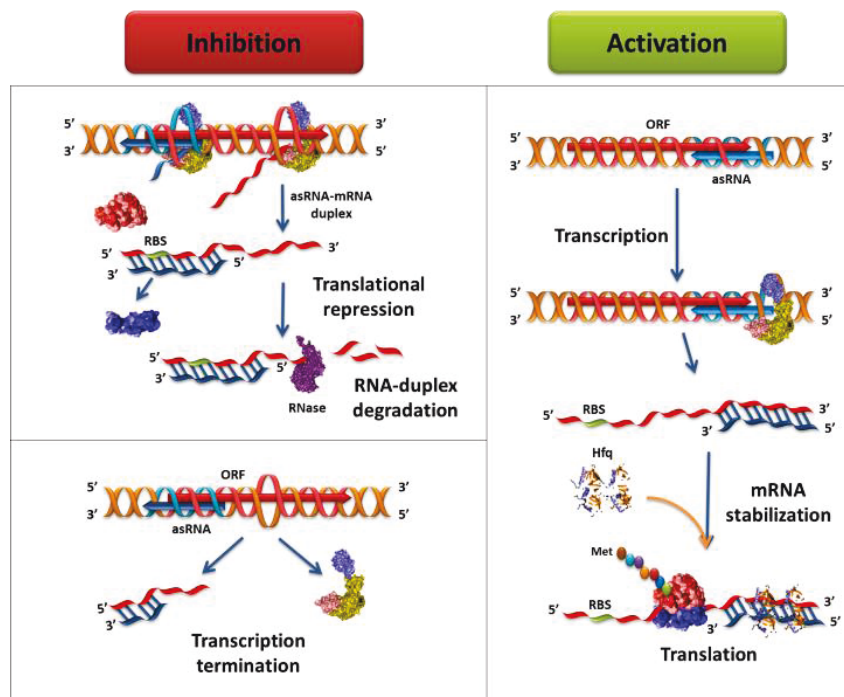


Figure 21 : Mécanisme d'action des ARN codés en *cis* ou antisens.

(Partie gauche) La liaison de l'ARNas (en bleu) au RBS inhibe la traduction. Le duplex ARNas-ARNm est ensuite dégradé par la RNase. L'ARNas peut également se lier à l'extrémité 5' de l'ARNm cible conduisant à la terminaison de la transcription. (Partie droite) Le complexe ARNas-Hfq peut également stabiliser l'ARNm cible pour activer sa traduction.

(Zorgani *et al.*, 2016)

a- Les ARN codés en *cis* ou antisens

Chez les procaryotes, la traduction des ARNm en protéines est assurée par le ribosome, composé d'une petite sous-unité 30S (contenant l'ARNr 16S) et d'une grande sous-unité 50S (contenant les ARNr 23S et 5S). Les ARNm portent une séquence codante, définie par un codon d'initiation (AUG mais aussi GUG et plus rarement UUG) et un codon stop. L'initiation de la traduction est assurée par la petite sous-unité 30S qui se fixe sur l'ARNm au niveau d'une région appelée « Shine-Dalgarno » (SD) ou « *ribosome binding site* » (RBS), située à une dizaine de bases en amont du codon d'initiation. La fixation de la sous-unité 30S au niveau de la séquence SD est médiée par un appariement de bases complémentaires entre l'ARN 16S et un court motif nucléotidique présent sur l'ARNm comportant 4 à 9 nucléotides consécutifs. Une fois que la sous-unité 30S est fixée sur l'ARNm, elle permet le recrutement et le positionnement de la grande sous-unité 50S de telle sorte que le site A du ribosome se trouve au niveau du codon d'initiation. Le ribosome ainsi assemblé recouvre physiquement une portion de l'ARNm, désignée sous le terme de région d'initiation de la traduction (TIR, pour « *translation initiation region* ») et qui correspond à une séquence de 40 nucléotides répartis de part et d'autre du codon d'initiation de la traduction. Cette région TIR doit être structurellement accessible pour que le ribosome puisse se fixer et que la traduction puisse être initiée, l'élément déterminant étant avant tout l'accessibilité à la séquence SD.

La très grande majorité des ARNas cible cette région TIR, souvent au niveau du RBS. Celui-ci, engagé dans un duplex avec l'ARNas n'est plus accessible à la reconnaissance par la sous-unité 30S qui, par conséquent, ne peut initier la traduction, conduisant à une inhibition de l'expression de la protéine codée par l'ARNm correspondant (Figure 21, partie gauche) (Repoila and Darfeuille, 2009). Le duplex ARNas-ARNm est souvent dégradé par des RNases, telles que RNase E, RNase Y, PNPase ou RNase III (*cf* dégradation des ARN). Ce mécanisme d'inhibition de la traduction a été décrit pour la première fois dans les éléments génétiques mobiles, tels que les plasmides et les phages (Wagner *et al.*, 2002). La réplication du plasmide R1 est sous le contrôle de l'ARNas CopA (Nordstrom, 2006). CopA est complémentaire de la région 5'UTR de l'ARNm *tap*. La traduction du peptide Tap, appelé peptide leader, est nécessaire pour favoriser la traduction de l'ARNm codant RepA qui initie la réplication du plasmide. En interagissant avec l'ARNm *tap*, CopA empêche la fixation du

ribosome sur le RBS de *tap* et, grâce au couplage traductionnel, inhibe indirectement la synthèse de RepA. Le duplex CopA-ARNm *tap* est ciblé par la RNase III qui dégrade les deux ARN.

Les ARNas peuvent également altérer l'expression de leur ARNm cibles par un mécanisme d'atténuation de la transcription (Figure 21, partie gauche) (Thomason and Storz, 2010). Dans certains cas fréquemment retrouvés chez les plasmides, la liaison de l'ARNas à l'ARNm induit la formation d'un terminateur sur l'ARNm et provoque ainsi l'arrêt prématuré de la transcription. Il a été montré que le long ARNas RNA β (427 nts), codé par le plasmide pJM1 du pathogène de poisson *Vibrio anguillarum*, permet l'atténuation de la transcription de l'ARNm polycistronique *fatDCBA-angRT* impliqué dans la biosynthèse du transporteur du fer (Stork *et al.*, 2007). L'interaction, entre la structure en tige-boucle de l'IGR de *fatA-angR* et celle du RNA β , provoque l'arrêt de la transcription et ainsi une augmentation dans le niveau de transcrit *fatA* par rapport à *angRT*.

D'autres études ont montré que l'appariement des ARNas sur l'extrémité 3'UTR permet de stabiliser l'ARNm cible pour une traduction efficace (Figure 21, partie droite). Le gène codant l'ARNas GadY, impliqué dans l'adaptation au stress acide, se situe dans l'IGR de l'opéron *gadXW* et est complémentaire à la région 3'UTR de *gadX* (Opdyke *et al.*, 2004). L'ARNm *gadXW* est instable et rapidement dégradé, mais la liaison du GadY induit une maturation qui génère deux ARNm distincts et plus stables, conduisant à une traduction des protéines GadX et GadW.

Les systèmes toxine/antitoxines (TA) de type I. Dans certains cas, les ARNrég codés en *cis* inhibent l'expression des gènes codant des peptides de courte taille (< à 60 acides aminés), généralement hydrophobes et toxiques pour la bactérie (Silvaggi *et al.*, 2005). Ces modules d'expression, transcrits à partir d'un même locus génomique, comprennent un ARNm codant la « toxine » et un ARNas répresseur codant « l'antitoxine ». A ce jour, plusieurs systèmes TA de type I ont été découverts, tels que le système TpxA-RatA chez *B. subtilis* (Silvaggi *et al.*, 2005), les systèmes OhsC-ShoB (Fozo *et al.*, 2008), TisB-IstR-1 (Argaman *et al.*, 2001) et SymE-SymR (Kawano *et al.*, 2007) chez *E. coli*. De façon générale, l'ARN-antitoxine est transcrit face à l'ARNm codant la toxine, au niveau de la TIR, formant

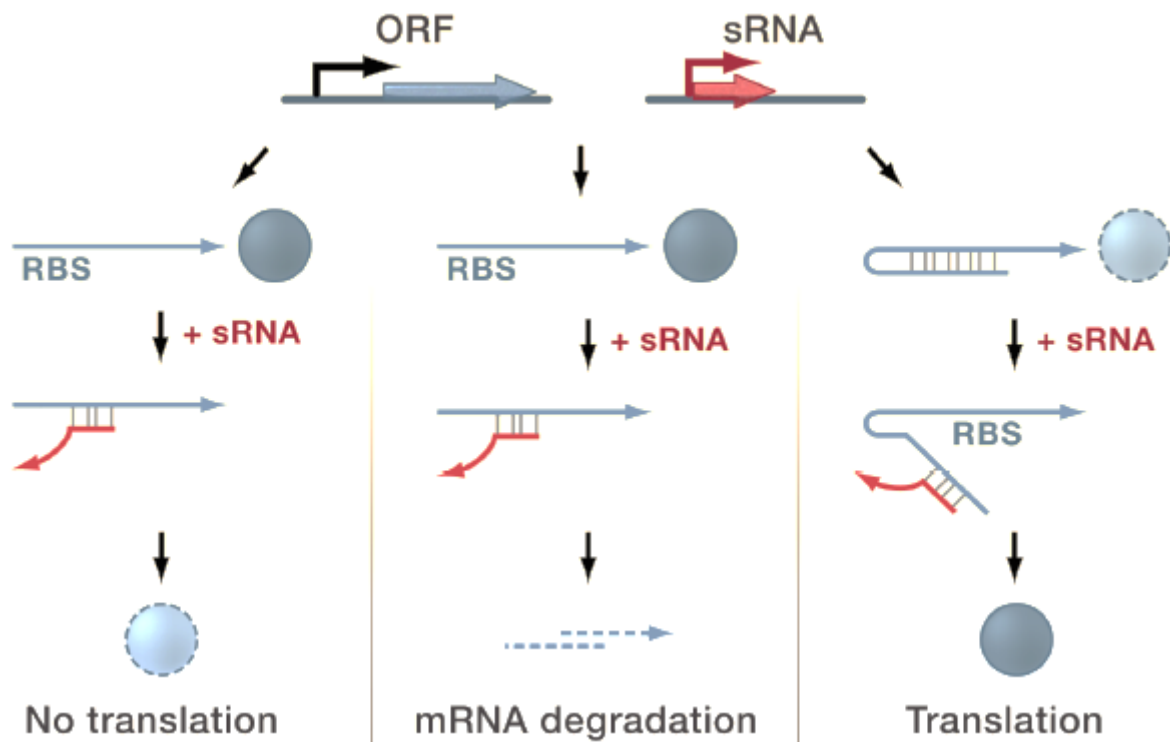


Figure 22 : Mécanisme d'action d'ARN codés en *trans*.

Le gène codant l'ARN codé en *trans* (rouge) est génétiquement distant du gène codant sa cible (bleu) et présente une faible complémentarité de séquence. Ces ARNrég peuvent inhiber la traduction en bloquant le RBS par liaison à l'extrémité 5'UTR et/ou la dégradation du duplex ARNrég-ARNm par des RNases. Les ARN codés en *trans* peuvent également activer la traduction en empêchant la formation des structures inhibitrices qui piègent le RBS.

D'après (Waters and Storz, 2009)

ainsi un duplex inhibiteur de la traduction.

b- Les ARN codés en *trans*

Les ARNrég appartenant à cette classe présentent une complémentarité partielle avec la séquence de leurs ARNm cibles. La liaison entre l'ARNrég et l'ARNm conduit généralement à une baisse du niveau de la protéine cible par inhibition de la traduction, dégradation de l'ARNm, ou les deux (Figure 22) (Gottesman and Storz, 2011). Bien que la plupart des ARN codés en *trans* se lient au niveau de 5'UTR de l'ARNm et piègent le RBS, certains ARNrég, tel que RyhB, inhibent la traduction de leur cibles par liaison en amont de l'AUG du gène cible (Masse *et al.*, 2005), ou encore au niveau de la séquence codante tel que l'ARNrég SgrS, impliqué dans la régulation de l'expression de l'ARNm polycistronique *manXYZ* (Rice *et al.*, 2012).

Pour les ARNrég codés en *trans*, une faible corrélation existe entre la localisation du gène codant l'ARNrég et celui codant l'ARNm cible. En effet, généralement un seul ARNrég peut cibler plusieurs messagers (Gottesman, 2005). Cette capacité d'interaction avec plusieurs ARNm résulte du fait que les ARN codés en *trans* nécessitent l'implication de peu de nucléotides pour inhiber la traduction de leur cible. Par exemple, bien que 23 nucléotides soient engagés dans l'interaction entre l'ARNrég SgrS et l'ARNm *ptsG*, seulement quatre mutations ponctuelles dans SgrS ont significativement affecté la régulation du *ptsG* (Kawamoto *et al.*, 2006). Afin de faciliter l'interaction ARN-ARN, à cause de la faible complémentarité de séquence, une grande majorité des ARN codés en *trans* agissent en concert avec la protéine chaperonne Hfq.

De plus, à l'inverse des ARN codés en *cis* exprimés de façon constitutive, les ARN codés en *trans* sont synthétisés dans des conditions spécifiques (stress, phase de croissance, densité cellulaire...). Les ARNrég MicA et RyhB régulent négativement l'expression des transcrits *yfeK* et *htrG* (Gogol *et al.*, 2011). La surexpression de ces deux protéines devient létale pour la bactérie et conduit à la mort cellulaire programmée, lorsque l'homéostasie membranaire atteint un niveau trop important. La mort programmée des bactéries diminue la taille de la population et la compétition pour l'assimilation des nutriments en période de stress sera plus faible.

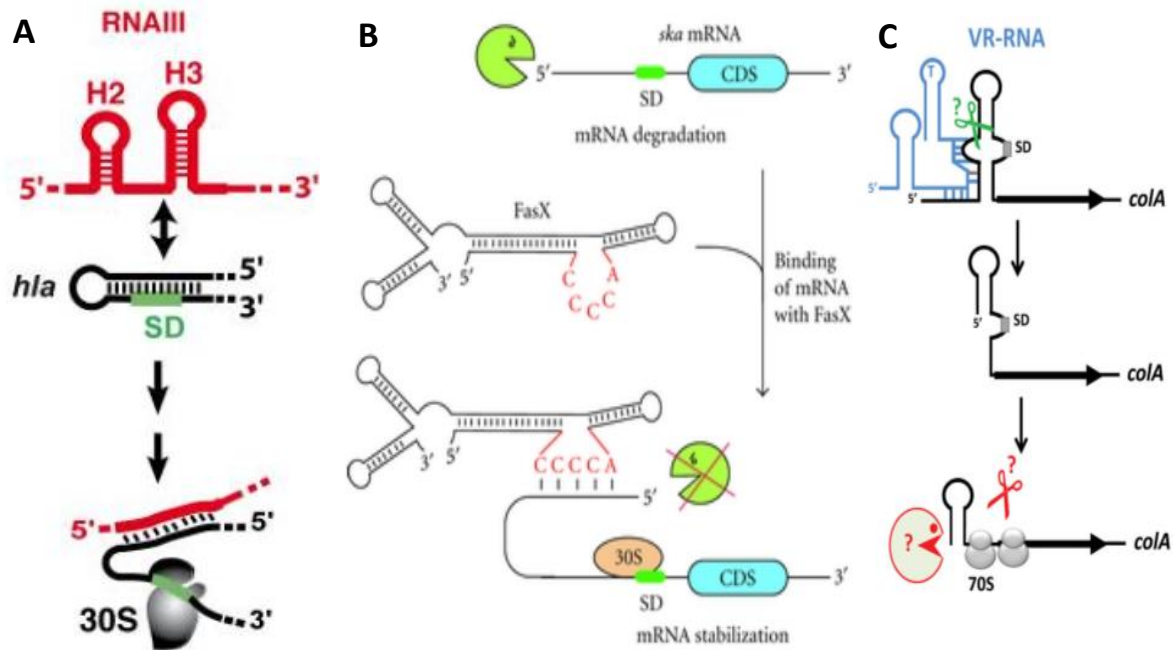


Figure 23 : Activation de la traduction par les ARN codés en *trans*.

(A) RNAIII cible l'ARNm *hla* et libère le SD, piégé dans une région en double brin, activant ainsi la traduction de l'ARNm *hla* (Felden *et al.*, 2011). (B) L'ARNrég FasX se lie à l'extrémité 5' de l'ARNm *ska*, le protégeant de la dégradation par la RNase J1 (Jester *et al.*, 2012). (C) La liaison du l'ARNrég VR à l'ARNm *colA* va induire le clivage par une endoribonucléase (en rouge) (Obana *et al.*, 2010).

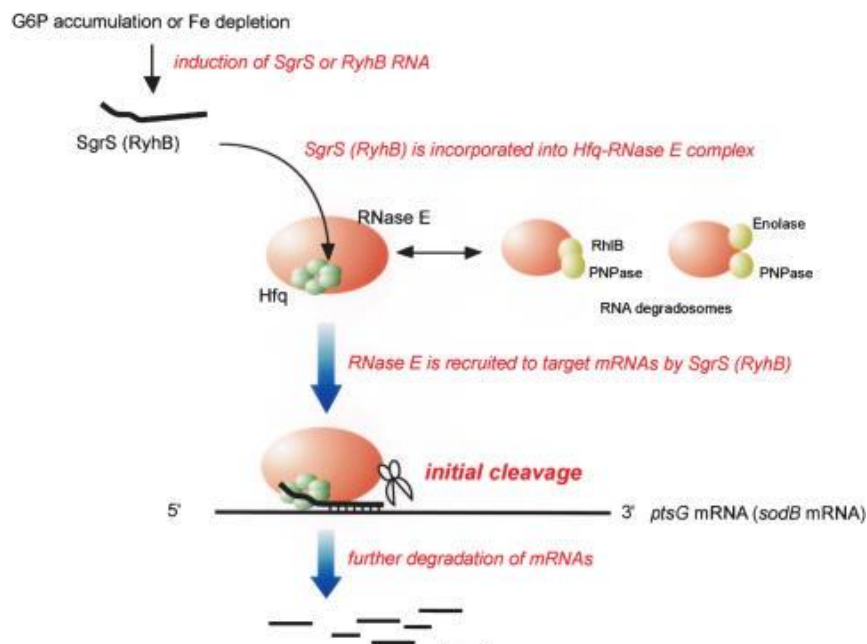


Figure 24 : Mode d'action de la protéine Hfq.

Rôle de la protéine Hfq dans l'action de l'ARNrég SgrS (RyhB) sur la traduction de sa cible ARNm *ptsG* (ARNm *sodB*).

D'après (Morita *et al.*, 2005)

Les ARN codés en *trans* peuvent également activer la traduction par divers mécanismes (Figure 22). Le mécanisme le plus courant consiste à empêcher la formation d'une structure inhibitrice séquestrant le RBS d'un ARNm. Chez *S. aureus*, le RNAIII possède plusieurs domaines de régulation, dont le domaine 5' qui interagit en amont de l'ARNm *hla*, codant l'alpha hémolysine (Felden *et al.*, 2011). Le RNAIII empêche la formation d'une région en double brin pour activer la traduction de l'ARNm (Figure 23. A).

L'ARNrég FasX active la traduction de l'ARNm *ska*, qui code un facteur de virulence (streptokinase) de *S. pyogenes* (Ramirez-Pena *et al.*, 2010). La fixation de FasX n'a pas d'effet direct sur la traduction, mais protège l'extrémité 5' de l'ARNm *ska* probablement contre la dégradation par l'activité exoribonucléase de la RNase J1 (Figure 23. B). Chez *Clostridium perfringens*, l'ARNrég VR se fixe à l'extrémité 5' de l'ARNm *colA*, codant la collagénase, piégée dans une structure en tige-boucle stable et inaccessible au ribosome (Obana *et al.*, 2010). La liaison de l'ARNrég VR induit une maturation spécifique de l'ARNm *colA* par une endoribonucléase inconnue qui permet de générer un ARNm doté d'une région 5'-UTR plus courte et un RBS libre et accessible, permettant d'initier la traduction (Figure 23. C).

Les ARN codés en *trans* requièrent souvent l'action d'une protéine chaperonne Hfq, afin de stabiliser le duplex ARN-ARN. De plus, les endoribonucléases (RNase III, RNase E, PNPase, RNase J1/J2, RNase Y) sont également des partenaires essentiels qui interviennent souvent dans l'étape de la dégradation de ce duplex.

L'ARN chaperonne Hfq. La protéine Hfq (*Host factor protein q*), homologue aux protéines Sm chez les eucaryotes, joue un rôle important dans l'épissage et la dégradation des ARNm (Vogel and Luisi, 2011). Cette protéine, capable de lier l'ARN, est conservée dans la moitié des bactéries séquencées (Gram positif et négatif confondus), mais absente dans d'autres comme les streptocoques (Moller *et al.*, 2002). Hfq s'assemble en un homohexamère, avec une forme en anneau, le recrutement de la RNase lui permet de stimuler la dégradation des ARNm cibles appariés aux sRNA (Figure 24) (Morita *et al.*, 2005). A l'inverse, la liaison de Hfq à RyhB sur le site reconnu par la RNase E permet de protéger l'ARNrég de la dégradation par la RNase (Moll *et al.*, 2003).

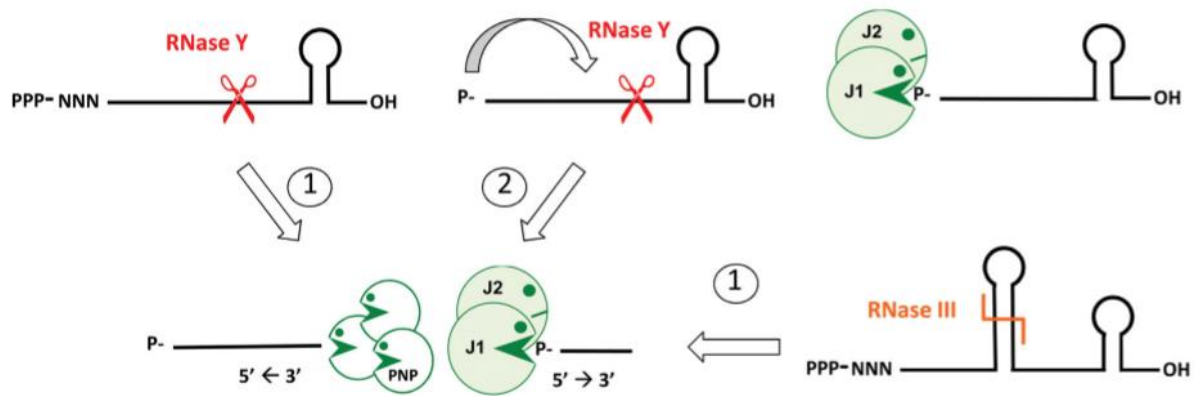


Figure 25 : Représentation schématique des voies de dégradation des ARN chez les bactéries à Gram positif.

Chez les bactéries, les ARNm sont protégés par leur extrémité 5' triphosphate. Cette extrémité peut être ciblée par des exoribonucléases telles que la RNase J1/J2 ou la RNase Y. L'extrémité 3', quant à elle, est principalement dégradée par l'action des PNPases. La RNase III possède une action spécifique aux ARN double brin.

D'après (Durand *et al.*, 2015)

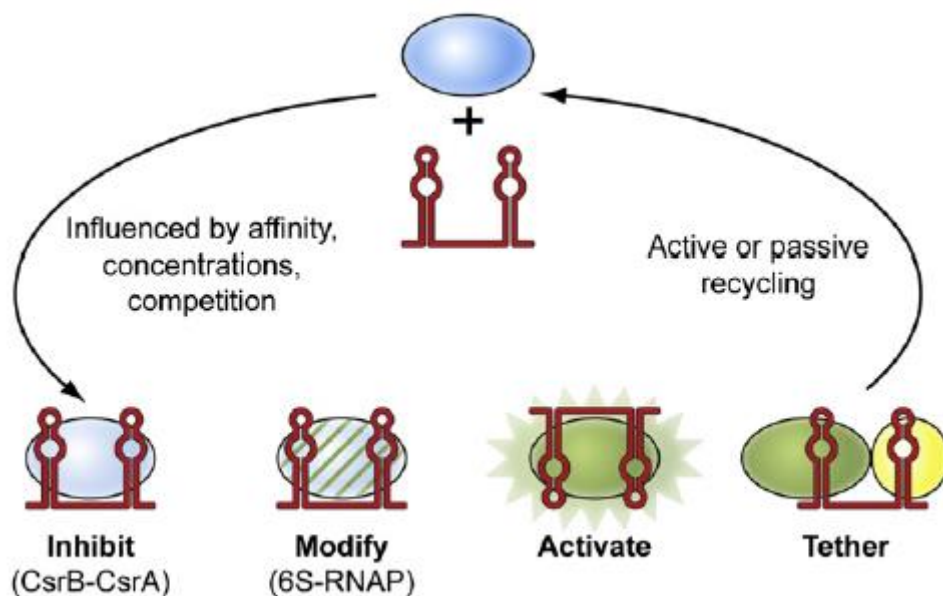


Figure 26 : Propriétés générales des sRNA séquestrant des protéines.

D'après (Storz *et al.*, 2011)

Dégradation par les RNases. De nombreuses ribonucléases sont impliquées dans la dégradation et le recyclage des ARN. La RNase E possède une action endonucléotidique, qui est spécifique des régions simple-brin avec une préférence pour les séquences riches en A/U (Arraiano *et al.*, 2010). Cette protéine agit au sein d'un complexe multi-protéique appelé dégradosome. L'inhibition du recrutement des ribosomes par la fixation de l'ARNrég au niveau de la séquence SD conduirait à un dénudement de l'ARNm, devenant ainsi vulnérable à la dégradation par la RNase E. L'activité de la RNase E est sensible à l'état de phosphorylation de l'extrémité 5' de l'ARNm : elle est plus active en présence d'une extrémité 5' monophosphate que triphosphate. Cette RNase est peu conservée chez les bactéries à Gram positif (Jester *et al.*, 2012). Chez ces dernières, la dégradation des ARNm fait intervenir d'autres enzymes, telles que la RNase III, la RNase Y, la PNPase ou le complexe RNase J1/J2 (Durand *et al.*, 2015). Par exemple, le duplex RNAIII/*spa* chez *S. aureus* est ciblé et dégradé par la RNase III, qui a une affinité pour les séquences d'ARN double brin (Figure 25) (Felden *et al.*, 2011). Les extrémités des ARNm clivées par la RNase Y sont ciblées par deux types d'exoribonucléases : de 3' vers 5' *via* PNPase, et de 5' vers 3' *via* les RNases J1 et J2 (Figure 25). Les RNases J1 et J2 possèdent en plus une activité endoribonucléase (Durand *et al.*, 2015).

3. Les ARN régulateurs ciblant les protéines

Certains ARNrég exercent leurs fonctions en interagissant directement avec des protéines et en modulant leur activité (Figure 26). Parmi ces ARNrég, l'ARNrég CsrB module l'activité de CsrA chez *E. coli* (Babitzke and Romeo, 2007). La protéine CsrA, impliquée dans la régulation du métabolisme du carbone, se lie à des motifs GGA dans la région 5'-UTR des ARNm ciblés, affectant ainsi leur stabilité et leur traduction (Liu *et al.*, 1997). L'ARNrég CsrB contient de nombreux motifs GGA (de 13 à 22), spécifiques à CsrA. Dans le cytoplasme de la bactérie, CsrB séquestre la protéine CsrA afin de diminuer sa disponibilité, et ainsi limiter sa fixation sur les régions 5'-UTR des ARNm ciblés.

Un autre ARNrég caractérisé dans cette classe est l'ARN 6S. Chez *E. coli*, l'ARN 6S est un ARN non-codant qui s'accumule en phase stationnaire de croissance (Wassarman, 2007). L'ARN 6S co-précipite avec l'ARN polymérase, et plus précisément avec l'holoenzyme associée au facteur σ^{70} . Ainsi le rôle de cet ARNrég est de diminuer la transcription de

Tableau 3 : Liste non-exhaustive d'ARNrég impliqués dans le contrôle de la virulence chez plusieurs bactéries pathogènes pour l'Homme, d'après (Reignier, 2010)

Espèce	ARNrég	Taille (nt)	Cible	Mode d'action	Phénotype/Fonction
ARNrég codés en <i>trans</i>					
<i>S. aureus</i>	RNA III	514	<i>rot</i> (repressor of toxins), <i>spa</i> , <i>sa1000</i> et <i>coa</i>	Inhibition de la traduction	Régulateur global du quorum-sensing et de l'expression des gènes de virulence
			<i>hla</i> (α -hemolysine)	Activation de la traduction	
	SprD	142	<i>Sbi</i> , facteur d'échappement à la réponse immunitaire	Inhibition de la traduction	Virulence chez la souris
	RsaA	139	<i>mgrA</i>	Inhibition de la traduction	Atténue la virulence chez la souris: diminue la formation de la capsule, favorise la persistance et augmente la formation du biofilm
<i>S. pyogenes</i>	RivX	220	<i>mga</i>	n.d	Régulation de l'expression de facteurs de virulence
	FasX	205	<i>ska</i>	Activation de la traduction	
<i>S. Typhimurium</i>	MgrR	98	<i>eptB</i>	Inhibition de la traduction	Régulation des modifications du LPS
<i>V. cholerae</i>	Qrr-1	100	<i>hapR</i> , facteur de transcription	Inhibition de la traduction	Contrôle du quorum-sensing
	VrrA	140	<i>ompA</i>		Modulation de la colonisation de l'intestin, chez la souris
ARNrég codés en <i>cis</i>					
<i>S. Typhimurium</i>	IsrC	288	<i>msgA</i> (macrophage survival gene)	Antisens probable (dégradation mutuelle <i>IsrC/msgA</i>)	Virulence et survie dans les macrophages
<i>M. tuberculosis</i>	<i>AsDes</i>	110	<i>desA1</i> (désaturase des acides gras)	Mécanisme antisens (prédiction)	Régulation de la biosynthèse des lipides (induite lors de la colonisation de la souris)
ARNrég ciblant les protéines					
<i>L. pneumophila</i>	6S	182	ARN polymérase sigma70	Séquestre l'ARNpol en mimant un promoteur	Régule le système de sécrétion de type IV
	RsmY, Z	132	RmsA, régulateur traductionnel	Séquestration de RsmA sur plusieurs motifs de fixation	Régulateurs clés pour la réplication dans les macrophages
<i>S. Typhimurium</i>	CsrB, C	300	CsrA, régulateur traductionnel	Séquestration de CsrA	Régulateurs clés pour la réplication dans les macrophages

certaines gènes par le facteur σ^{70} . Sa structure secondaire mime une bulle de transcription qui recrute l'ARN polymérase lorsqu'elle est liée au facteur σ^{70} (Wassarman, 2007). Certaines hypothèses suggèrent que l'ARN 6S pourrait piéger l'ARN polymérase chargée avec le facteur σ^{70} en mimant la structure d'un promoteur bactérien pour permettre le recyclage de l'ARN polymérase lors de l'entrée en phase stationnaire (Wassarman, 2007).

C. Exemple de certains ARN régulateurs

1. ARN régulateurs impliqués dans la virulence chez *Staphylococcus aureus*

Certains ARN régulateurs sont au cœur des événements de régulation touchant les facteurs de virulence. Ces facteurs de virulence sont des armes fatales que la bactérie exprime au cours de processus infectieux, souvent en modulant leur expression en fonction des conditions environnementales. Ainsi, les études d'ARNrég chez des bactéries pathogènes pour l'homme ont démontré un rôle crucial de ces régulateurs dans l'adaptation à la niche humaine, la défense contre le système immunitaire et le contrôle des gènes de virulence (Tableau 3). Par exemple, l'ARNrég RNAIII de *S. aureus* contrôle l'expression de plusieurs facteurs de virulence tels que *rot* (*repressor of toxins*), *spa* (surface protein A), *sa1000* (protéine liant le fibrinogène) ou *coa* (coagulase) (Boisset *et al.*, 2007). L'ARNrég SprD régule l'expression de l'ARNm *sbi*, codant le facteur d'adhésion, la protéine G (Chabelskaya *et al.*, 2010; Chabelskaya *et al.*, 2014). Dans un modèle murin, l'expression de SprD module la virulence de *S. aureus*, mais indépendamment de l'expression de la protéine Sbi. L'ARNrég RsaA est impliqué dans la virulence, la persistance et la formation du biofilm chez *S. aureus* (Lays *et al.*, 2014). Les avancées sur l'étude des ARNrég chez *S. aureus* a permis à Felden et ses collaborateurs de proposer RNAIII et SprD comme des potentielles biomarqueurs des infections du système sanguin (Bordeau *et al.*, 2016). La liste du tableau 3, présentant des ARNrég impliqués dans la virulence chez l'Homme, est loin d'être exhaustive car de nombreux autres ARNrég sont décrits et d'autres restent à découvrir (Papenfort and Vogel, 2010).

Tableau 4 : Les ARNrég étudiés chez les streptocoques d'après (Le Rhun and Charpentier, 2012)

Espèces	ARNrég	Cible et mécanisme d'action	Fonction	Régulation de l'expression de sRNA
<i>S. pneumoniae</i>	csRNA1	?	Restore le retard de croissance du mutant <i>ciaR</i> Affecte la longueur des chainettes du mutant <i>ciaR</i>	Activé par CiaR et le peptide stimulateur de compétence
	csRNA2, 3	?	?	Activé par CiaR
	csRNA4	?	Influence l'autolyse en phase stationnaire	Activé par CiaR
	spd-sr17, spd-sr37	?	Impliqué dans la réponse au stress	?
<i>S. pyogenes</i>	FasX	Stabilisation de l'ARNm <i>ska</i> par liaison à l'extrémité 5'	Améliore la traduction de Ska Affecte les transcrits <i>fbp</i> et <i>mrp</i> et le sRNA Pel Améliore l'adhérence et l'internalisation des cellules Hep-2	activé par: FasBCA, les acides aminés, le système LuxS/AI-2
	Pel	?	Affecte la transcription des ARNm <i>emm</i> et <i>speB</i> , l'expression de Ska, d'immunoglobuline et de la protéine de liaison à la fibronectine Affecte le niveau de transcription des gènes <i>nga</i> et <i>sic</i> et la traduction du <i>SpeB</i> Affecte la virulence	Activateurs: Acides aminés, CodY, Irr, SLS, StpR Inhibiteurs: CcpA, CovRS, FasX, LuxS, Mga, Nra, Ralp3, RofA, RopB/Rgg
	RivX	?	Affecte l'expression des ARNm <i>emm</i> , <i>scpA</i> , <i>mga</i> et <i>speA</i>	Activé par CovRS
	tracrRNA	Clivage de crRNA en présence de Rnase III et Cas9	Active le système CRISPR/Cas impliqué dans l'immunité bactérienne	?
	crRNA (type 2)	Cible l'ADN étranger	Système CRISPR/Cas impliqué dans l'immunité bactérienne	?
	SQ18 et SQ485	<i>gbs0031 (sip)</i> et <i>gbs1588</i>	Régulent négativement l'expression des gènes cibles	?
<i>S. agalactiae</i>	SQ893	<i>gbs1263</i>	Active l'expression	?

1. ARN régulateurs chez *Clostridium*

Plusieurs ARNrég ont été également identifiés chez le pathogène humain *Clostridium*. Chez *C. difficile*, 251 ARNrég ont été identifiés par RNA-seq (94 agissant en *trans*, 91 ARNas et 66 *riboswitch*) (Soutourina *et al.*, 2013). L'expression de 35 de ces ARNrég a été confirmée par northern blot et/ou RT-PCR. Cette étude a permis d'identifier l'ARNrég dit « *housekeeping* », l'ARN 6S ; des ARNrég régulés par la phase de croissance ; des *riboswitch* fixant le cyclic di-guanosyl-5'-monophosphate (c-di-GMP) ; ou encore le locus CRISPR.

D'autres exemples d'ARNrég existent chez d'autres espèces de *Clostridium* responsables d'intoxications alimentaires. L'expression de l'ARNrég VR-RNA (VirR-regulated RNA) de *C. perfringens* est régulée par le TCS VirSR (Shimizu *et al.*, 2002). VR-RNA active l'expression de nombreux gènes, notamment des gènes codant des toxines, et réprime l'expression d'autres, notamment *luxS* codant une autoinducteur 2 synthase.

2. ARN régulateurs chez *Enterococcus faecalis*

Enterococcus faecalis est un pathogène opportuniste de l'Homme, responsable de certaines infections nosocomiales (infections des voies urinaires, des plaies et des tissus mous). La délétion de l'ARNrég *ef0408-0409*, homologue de RNAlI composant du système TA de type I, a permis d'augmenter la virulence du mutant, sa survie dans les macrophages et la colonisation des organes dans un modèle d'infection de souris. D'autres phénotypes ont été observés pour le mutant *ef0408-0409*, notamment la résistance aux stress acide, oxydatif et osmotique (Michaux *et al.*, 2014). De nos jours, plus de 100 ARNrég ont été identifiés chez *E. faecalis* (Innocenti *et al.*, 2015), dont la fonction de nombreux d'entre eux reste à déterminer.

3. ARN régulateurs impliqués dans la virulence de *Streptococcus*

La plupart des études menées sur les ARNrég chez *Streptococcus* concerne trois espèces pathogènes : *S. pyogenes*, *S. pneumoniae* et *S. agalactiae*. Bien que plusieurs ARNrég ont été identifiés à ce jour, très peu sont caractérisés (Tableau 4). A ce jour, 75 ARNrég ont été identifiés chez *S. pyogenes* (Perez *et al.*, 2009; Patenge *et al.*, 2012; Tesorero *et al.*, 2013) et 179 ARNrég chez *S. pneumoniae* (Kumar *et al.*, 2010; Tsui *et al.*, 2010; Acebo

et al., 2012; Mann *et al.*, 2012). Chez *S. agalactiae*, 197 ARNrég ont été prédits *in silico*, et dix d'entre eux ont été validés par northern blot (Pichon *et al.*, 2012). De plus, 125 ARNrég ont été identifiés par une approche de séquençage massif d'ARN différentielle (dRNA-seq) avec 39 sRNA agissant en *cis*, 39 asRNA et 47 sRNA codés en *trans* (Rosinski-Chupin *et al.*, 2015).

a- **ARN régulateurs chez *Streptococcus pyogenes***

Chez *S. pyogenes*, deux gènes forment un opéron dont le transcrit comporte deux éléments : le facteur de transcription RivR et un ARNrég non-codant RivX, qui semble traité au niveau post-transcriptionnel pour former un ARN régulateur à part entière. L'absence de RivR et de RivX altère de manière identique le niveau d'expression de nombreux gènes dont le régulateur transcriptionnel Mga, qui assure un contrôle direct de la transcription des gènes de virulence. Cette réduction de l'expression des gènes de virulence en absence de RivR/RivX est restaurée à un niveau identique par une complémentation soit de RivR, soit de RivX, démontrant le caractère redondant de ces deux régulations, l'une exercée par une protéine et l'autre par un sRNA (Roberts and Scott, 2007).

L'ARNrég FasX a été identifié la première fois par Kreikemeyer et ses collaborateurs (Kreikemeyer *et al.*, 2001). Cet ARNrég appartient au locus *fasBCAX* codant deux histidines kinases putatives (FasB et FasC), un régulateur de réponse de liaison à l'ADN (FasA) et un ARNrég de 205 nts (FasX). Deux gènes codant des facteurs de virulence ont été confirmés comme cibles directes et le mécanisme de régulation a été étudié. La première cible est le gène *ska* codant la streptokinase (Ramirez-Pena *et al.*, 2010). La streptokinase est une protéine sécrétée qui favorise la conversion du plasminogène humain vers une protéase plasmine. La formation d'une structure secondaire à l'extrémité 5', suite à la liaison de l'ARNrég FasX, permet d'augmenter la stabilité de l'ARNm *ska*. La deuxième cible de FasX est l'ARNm *cpa* codant le pili (Liu *et al.*, 2012). Chez *S. pyogenes*, le pili favorise l'adhérence du pathogène aux cellules de l'hôte et à la formation du biofilm. L'ARNrég FasX régule négativement l'expression de l'ARNm *cpa* par l'inhibition de l'accès au RBS par le ribosome (Liu *et al.*, 2012). FasX module également l'expression d'autres cibles ayant un rôle dans la virulence, notamment *sagS* (streptolysine), *fbp54* et *mrp* (protéines de liaison à la fibronectine).

Possédant des fonctions similaires à celles de RNAlII de *S. aureus*, l'ARNrég Pel (*pleiotropic effect locus*) est un ARNrég bi-fonctionnel, étant un effecteur de l'expression de facteurs de virulence et codant en même temps l'expression d'une hémolysine (Mangold *et al.*, 2004). Pel est également un régulateur pléiotrope de l'expression de facteurs de virulence à la fois au niveau transcriptionnel (*emm* codant la protéine M et *sic* codant l'inhibiteur du complément chez les streptocoques) et au niveau post-transcriptionnel (*speB* codant une cystéine protéase). L'expression de l'ARNrég Pel est modulée par de nombreux régulateurs de virulence, notamment CcpA et CovRS (Shelburne *et al.*, 2008).

b- ARN régulateurs chez *Streptococcus pneumoniae*

S. pneumoniae, appelé également pneumocoque, est un pathogène opportuniste, responsable d'un grand nombre de maladies humaines, allant de l'otite moyenne aiguë à des infections plus graves, telles que la méningite, la septicémie et l'endocardite (O'Brien *et al.*, 2009). Une combinaison entre des approches bioinformatiques, analyse de transcriptome et de RNA-seq ont été utilisées afin d'identifier les ARNrég codés par *S. pneumoniae*. Tandis qu'une étude *in silico* a permis d'identifier 128 ARNrég localisés dans des régions intergéniques (Livny *et al.*, 2008), une analyse par microarray a permis d'identifier 50 ARNrég putatifs de la souche TIGR4 (Kumar *et al.*, 2010). En outre, Acebo et ses collaborateurs ont identifié 88 ARNrég par RNA-seq (Acebo *et al.*, 2012) ; et Mann et ses collaborateurs ont augmenté le nombre total d'ARNrég identifiés chez *S. pneumoniae* à 178, dont certains sont impliqués dans la virulence (Mann *et al.*, 2012).

c- ARN régulateurs chez *Streptococcus agalactiae*

Malgré son pouvoir pathogène dans les infections néonatales, seulement trois études ont été effectuées sur les ARNrég de *S. agalactiae* (Marx *et al.*, 2010; Pichon *et al.*, 2012; Rosinski-Chupin *et al.*, 2015). Ces travaux ont permis l'identification et la validation de nombreux ARNrég de la souche NEM316. Marx et ses collaborateurs ont montré *in silico* que chaque espèce de streptocoques héberge de 2 à 6 ARNrég associés au TCS CiaRH (csRNA pour « *cia-associated small RNA* »). Afin de réaliser cette prédiction bio-informatique, une recherche des promoteurs activés par le régulateur CiaR, localisés dans des régions intergéniques et suivis par un terminateur de transcription a été effectuée. Quatre csRNA

potentiels ont été prédits pour interagir avec le TCS CiaRH de la souche *S. agalactiae* NEM316 (Marx *et al.*, 2010). Dans une autre étude, l'utilisation des RIT « *Rho-independent terminators* » comme unique signature pour la prédiction *in silico* des ARNrég a permis de prédire 197 ARNrég dans le génome de la souche NEM316 de *S. agalactiae*. Vingt-six candidats sur les 30 testés ont été validés par RT-PCR. Parmi ces 26 candidats, l'expression de 10 d'entre eux a été confirmée par northern blot (Pichon *et al.*, 2012).

Récemment, M.F. Lartigue, en collaboration avec l'équipe de P. Bouloc et celle de P. Glaser, a identifié les ARNrég de la souche NEM316 par dRNA-seq « *differential RNA-sequencing* » (Rosinski-Chupin *et al.*, 2015). Cette étude a permis, entre autre, l'identification de 47 ARNrég localisés dans des régions intergéniques et agissant potentiellement en *trans*. L'expression de sept d'entre eux a été étudiée par northern blot en fonction de la cinétique à différentes phases de croissance et conditions de stress.

L'expression de ces sept ARNrég, dont trois régulés par CiaR, est modifiée en réponse à des variations de pH. Ces modifications suggèrent que ces ARNrég jouent un rôle dans l'adaptation de *S. agalactiae* aux changements environnementaux, notamment en conditions acides. Puisque cette adaptation est primordiale en cas de colonisation vaginale et afin de survivre dans le phagosome des macrophages lors d'un processus infectieux, ces ARNrég peuvent potentiellement être impliqués dans la virulence de *S. agalactiae*.

Dans les prochains chapitres, nous exposerons les résultats du travail expérimental.

OBJECTIFS DU TRAVAIL

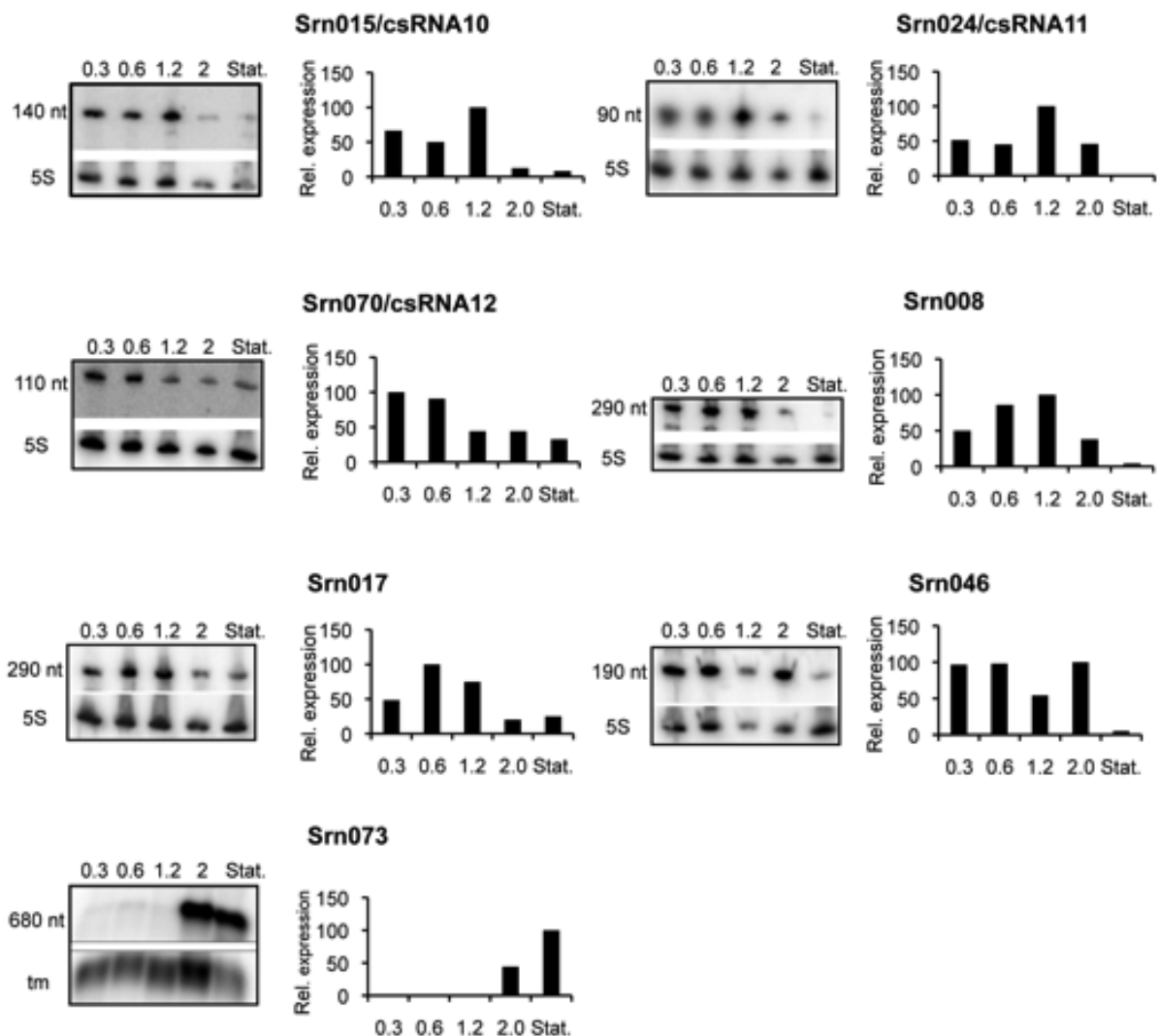


Figure 27 : Profil d'expression northern blot de sept ARNrég de *S. agalactiae* NEM316 en fonction de la phase de croissance.

Les ARN totaux ont été extraits à partir de cultures en milieu THY (milieu TH additionné d'extrait de levure) centrifugées à la DO_{600nm} : 0,3 ; 0,6 ; 1,2 ; 2 ; et en phase stationnaire tardive ou « overnight » (Stat.). L'expression constitutive du transcrit des gènes ARN 5S et ARNtm a été utilisée comme contrôle.

D'après (Rosinski-Chupin *et al.*, 2015)

Objectifs du travail

A. Introduction

Les évènements de régulation touchant les facteurs de transcription permettent généralement aux bactéries de passer d'un état physiologique à un autre, en remaniant rapidement leur arsenal de gènes d'adaptation et de virulence pour répondre à une condition environnementale spécifique. *S. agalactiae* est une bactérie qui colonise le tube digestif et le tractus génital. Etant donné l'omniprésence des régulations ARNrég-dépendantes chez les procaryotes et les conditions changeantes des niches digestive et génitale, nécessitant une régulation fine et rapide des gènes d'adaptation, il est difficile d'imaginer que *S. agalactiae* puisse se passer de ce mode de régulation médiée par les ARNrég.

B. Données préliminaires

1. Identification des ARNrég dans la souche NEM316

Au sein de l'équipe Bactéries et risque materno-fœtal, dont le responsable est le Professeur Laurent MEREGHETTI, les recherches effectuées par Marie-Frédérique LARTIGUE sont orientées vers la caractérisation des petits ARN régulateurs (ARNrég) chez *S. agalactiae*. Ce pathogène opportuniste est la première cause d'infections néonatales. Jusqu'à présent, aucune étude n'a caractérisé le rôle des ARNrég pour ce pathogène. Pendant une mobilité au sein de l'équipe Signalisation et Réseaux de Régulations Bactériens, dirigée par Philippe BOULOC, M.F. LARTIGUE a mis en évidence l'expression de sept ARNrég dans la souche *S. agalactiae* NEM316 en collaboration avec l'équipe de Philippe Glaser (Figure 27). Ces ARNrég présentent des caractéristiques très variables, telles que la taille, le profil d'expression en fonction de la phase de croissance ou encore en fonction de stress (Figure 28).

L'ARNrég Srn008 a une taille d'environ 290 nucléotides (Figure 27). Son gène, *srn008*, est situé entre *gbs0220* qui code une protéine de fonction inconnue et *gbs0221* qui code une protéine hypothétique de la famille du système toxine-antitoxine RelE/ParE. L'expression de Srn008 augmente pendant la phase de croissance exponentielle puis décroît

en phase stationnaire. Il s'accumule en stress acide (Figure 28).

Srn015 a une taille d'environ 140 nucléotides (Figure 27). Il correspond au csRNA10 (Marx *et al.*, 2010). Son gène, *srn015*, est localisé entre *gbs0437* et *gbs0438* qui codent respectivement une glucose-6-phosphate isomérase et une 5-formyltetrahydrofolate cyclo-ligase. L'expression de Srn015 diminue fortement en phase stationnaire. En revanche, il est fortement exprimé en stress acide, mais également de façon beaucoup plus modérée, lors d'un stress oxydant (H₂O₂) et en présence de lysozyme (Figure 28).

Srn017 a une taille d'environ 290 nucléotides (Figure 27). Il correspond au SQ154 (Pichon *et al.*, 2012). Son gène, *srn017*, est localisé entre *gbs0476* qui code une protéine de membrane et *gbs0477* qui code une protéine de fonction inconnue. Le brin complémentaire de *srn017* code également l'ARNrég Srn018. Le système Srn017-Srn018 présente une homologie avec un système toxine-antitoxine. L'expression de Srn017 diminue en phase stationnaire et s'accumule lors d'un stress oxydant et en présence de lysozyme (Figure 28).

Srn024 a une taille d'environ 90 nucléotides (Figure 27). Il correspond au csRNA11 (Marx *et al.*, 2010). Son gène, *srn024*, est situé entre *gbs0600* et *gbs0601* qui codent des protéines hypothétiques de la superfamille des protéines CsbD qui sont des protéines de réponse au stress. L'expression de Srn024 est corrélée à celle de Srn015 dans les conditions testées (Figure 28).

Srn046 a une taille d'environ 190 nucléotides (Figure 27). Son gène, *srn046*, est situé entre *gbs1087* et *gbs1088* qui codent respectivement une protéine d'adhésion au fibrinogène FbsA et une cardiolipine synthétase. L'expression de Srn046 diminue fortement en phase stationnaire et en stress oxydant et salin (Figure 28).

Srn070 a une taille d'environ 110 nucléotides (Figure 27). Il correspond au csRNA12 (Marx *et al.*, 2010) et au SQ517 (Pichon *et al.*, 2012). Son gène, *srn070*, est localisé entre *gbs1697* et *gbs1698* qui codent respectivement une protéine favorisant l'absorption du glycérol et une protéine hypothétique de la superfamille des Cupin_2. L'expression de Srn070 diminue pendant la cinétique bactérienne, mais est augmentée lors d'un stress oxydant et en présence de lysozyme (Figure 28).

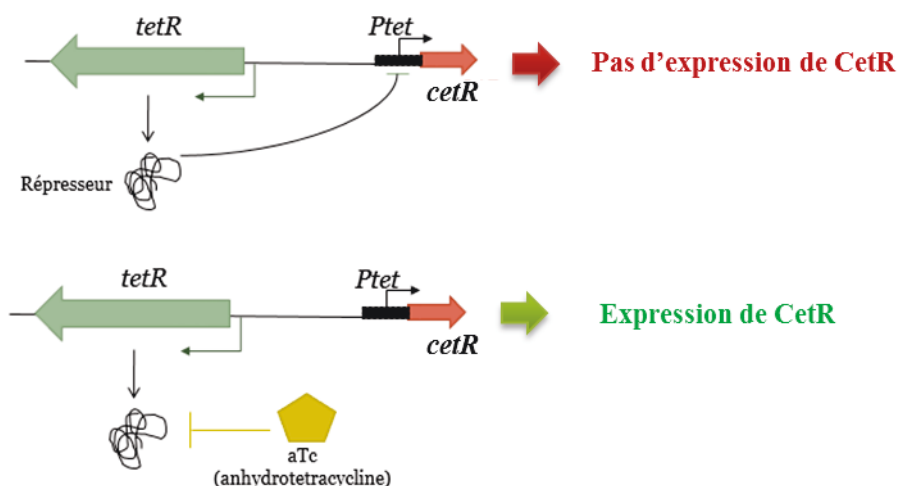


Figure 29 : Mécanisme d'action du vecteur d'expression pG+off portant un promoteur inducible par la tétracycline.

Le plasmide pG+off exprime le répresseur TetR qui se lie à l'opérateur tetO ; dans ce cas, il n'y a pas d'expression de *cetR*. L'expression génique est induite par l'ajout de l'anhydrotétracycline (aTc) dans le milieu, dont l'action est d'inhiber le répresseur.

D'après (Lartigue and Boulloc, 2014)

Tableau 5 : Gènes différentiellement exprimés dans le mutant *cetR* induit (I) par rapport au mutant *cetR* non-induit (NI), évalués par microarray

Opéron	gène	Microarray		Fonction prédite (MicroCyc pathways) (http://www.genoscope.cns.fr/microme/microcyc-wsc/)
		Fold Change (I vs NI)	p-value	
1	<i>gbs0896</i>	1,87	0,005	Glycolyse / Gluconéogenèses, métabolisme du pyruvate, biosynthèses des métabolites secondaires
	<i>gbs0897</i>	1,86	0,006	Cycle citrate (cycle TCA), métabolisme du butanoate, biosynthèses de la valine, leucine et isoleucine
2	<i>gbs1112</i>	0,55	0,007	Biosynthèses de la thiamine lipoprotéine
	<i>gbs1113</i>	0,51	0,004	NADPH-dépendent FMN réductase domain-containing protein
	<i>gbs1114</i>	0,53	0,005	
3	<i>gbs1202</i>	2,01	0,001	Protéine de stress similaire à la famille Gls24
	<i>gbs1203</i>	1,92	0,004	MF3-like protéine
	<i>gbs1204</i>	1,83	0,007	Métabolisme de la purine
	<i>gbs1205</i>	1,78	0,013	Protéine membranaire
	<i>gbs1206</i>	1,8	0,01	Protéine exportée
	<i>gbs1207</i>	1,84	0,007	Transglycosylase
	<i>gbs1208</i>	1,69	0,054	
4	<i>gbs1504</i>	1,89	0,005	Protéine rhodanese-like
5	<i>gbs1786</i>	1,82	0,008	Phosphorylation oxydative, système à deux composants
	<i>gbs1787</i>	1,87	0,006	Cytochrome bd-II oxydase, sous-unité I 1.10.3.-
	<i>gbs1788</i>	1,81	0,008	Protéine NADH dehydrogenase-like yjID

Enfin, l'ARNrég Srn073, que nous avons renommé CetR pour « *cell-envelope-targeting RNA* », a une taille d'environ 680 nucléotides (Figure 27); c'est le plus long ARNrég qui a été identifié lors de cette étude, il ne contient pas de séquence codante. Il correspond au SQ968 prédit par Pichon *et al.* (2012). Son gène est localisé entre *rpsF* et *mutY* qui codent respectivement la sous-unité 30S du ribosome et une adénine glycosylase. Contrairement aux six autres ARNrég précédemment décrits, CetR est très fortement exprimé en phase stationnaire, mais également en stress salin (Figure 28).

2. Identification des cibles de CetR par microarray

La majorité des ARNrég dont la fonction a été caractérisée à ce jour agissent par appariement de bases avec leur ARNm cibles. Afin d'identifier les cibles potentielles de CetR, nous avons combiné plusieurs stratégies (*cf*, chapitre 2). Une des approches utilisées est l'étude par analyse transcriptomique des effets de la surproduction de CetR sur le transcriptome [en utilisant le plasmide pG+off qui permet l'induction régulée de CetR chez *S. agalactiae* (Lartigue and Boulloc, 2014)]. Le plasmide pG+off est un vecteur d'expression portant un promoteur inducible par la tétracycline. Ce vecteur permet une expression génique étroitement contrôlée chez *S. agalactiae*. Il exprime le répresseur TetR qui se lie à l'opérateur *tetO* ; l'expression génique est induite en ajoutant de l'anhydrotétracycline (aTc) dans le milieu de culture (Figure 29). Cette approche favorise l'identification des cibles directes de l'ARNrég. Le microarray ou analyse transcriptomique permet d'analyser le niveau d'expression des gènes dans un organisme à un temps donné par rapport à un échantillon de référence. Les expériences de microarray ont été réalisées en collaboration avec I. Rosinski-Chupin, équipe P. Glaser, Institut Pasteur, Paris.

L'accumulation de CetR a conduit à l'activation de 13 gènes et à la régulation négative de trois autres gènes (Tableau 5). L'ensemble des gènes modulés significativement, suite à la surexpression de CetR, appartient à cinq opérons distincts. L'opéron 1 (*gbs0886* et *gbs0897*) code des gènes impliqués dans le métabolisme du carbone ou la biosynthèse des acides aminés. L'opéron 2 (*gbs1112-4*) est le seul opéron dont la régulation est négative. Les gènes appartenant à l'opéron 3 codent différentes protéines de fonctions inconnues ; cependant, la recherche dans la base de données Pfam (<http://pfam.xfam.org/>) a montré que les protéines codées par les gènes *gbs1202* et *gbs1204* présentent une forte homologie

de séquence avec les protéines GlS24 et GlSB, respectivement (environ 60 % d'identité de séquence). Une étude a montré que l'inactivation des gènes de l'opéron *gls24* conduit à des effets pléiotropes, notamment sur la morphologie cellulaire, la sensibilité aux stress et l'expression des gènes chez *E. faecalis* (Giard *et al.*, 2000). La fonction des gènes appartenant aux opérons 4 et 5 n'a pas été déterminée expérimentalement, cependant la prédiction bioinformatique montre que ces gènes sont impliqués dans le stress oxydatif.

La validation de l'expression de ces cibles potentielles sera présentée dans le chapitre 2. Ces cibles seront ensuite comparées à celles identifiées par d'autres approches *in silico* et *in vivo*.

C. Objectifs du travail de thèse

L'objectif de ma thèse est la caractérisation des ARNrég par l'étude de leur fonction, régulation, mécanisme d'interaction avec leurs cibles et de leur implication dans la physiologie et dans la virulence de la souche *S. agalactiae* NEM316.

L'objectif principal de cette thèse consistait à caractériser plus particulièrement l'ARNrég CetR d'un point de vue fonctionnel et mécanistique. Afin d'apporter une meilleure compréhension des mécanismes liés à la physiologie et à la virulence de la bactérie, nous avons identifié par différentes approches les cibles de CetR et nous avons caractérisé le mécanisme d'interaction de l'une d'entre elles (Publication N°1, soumis).

Le second objectif était de réaliser une analyse comparative globale du protéome sur des extraits totaux de protéines entre la souche sauvage et une souche délétée de l'ARNrég CetR. Cette approche a été conduite dans un milieu chimiquement défini dans le but d'identifier d'autres cibles et de définir les processus biologiques régulés par CetR dans cette condition (Publication N°2, en préparation).

La dernière partie de ma thèse consistait à : i) étudier la prévalence des sept ARNrég (Srn008, Srn015, Srn017, Srn024, Srn047, Srn070 et Srn073) chez différentes souches de *S. agalactiae* ; ii) étudier le système toxine-antitoxine (Srn017-Srn018) et iii) construire des mutants de délétion de différents ARNrég et identifier leurs cibles potentielles *in silico*.

RESULTATS

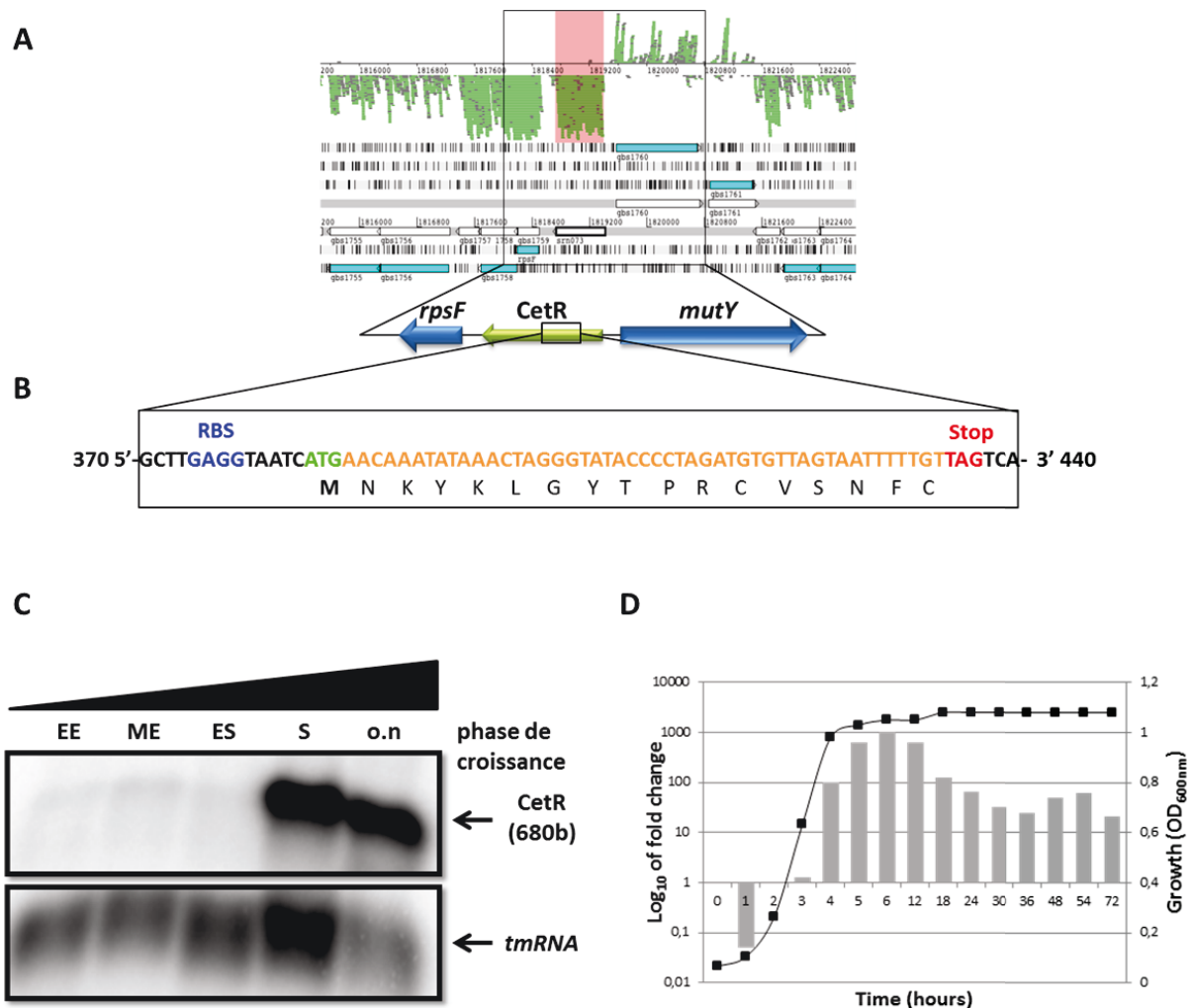


Figure 30 : Identification et validation de l'expression de l'ARNrég CetR.

(A) Identification de l'ARNrég CetR par RNA-seq et son environnement génétique. (B) Séquence du peptide prédit à partir de la séquence de *cetR*. (C) Profil northern blot révélant la quantité de transcrit CetR, en milieu THY (Todd-Hewitt-Yeast), dans la souche sauvage. L'ARNtm a été utilisé comme témoin. (D) Profil de qRT-PCR et de cinétique de croissance. L'expression de CetR est rapportée au gène de ménage *gyrB*. Abréviations, EE: exponentiel précoce, ME: mi-exponentiel, ES: stationnaire précoce, S: stationnaire, o.n: overnight, tmRNA: ARN thermometer.

Chapitre 1 : Caractérisation de l'ARNrég CetR : rôle dans la résistance aux PAM et dans la virulence de *S. agalactiae*

A. Introduction

La première partie de ma thèse s'inscrit dans la poursuite des travaux initiés par M.F. Lartigue, en collaboration avec les équipes de P. Boulloc et P. Glaser, sur l'identification et la caractérisation des ARNrég de *S. agalactiae*. Au début de cette collaboration, M.F. Lartigue a choisi de travailler sur la souche *S. agalactiae* NEM316, qui est la première souche dont le génome était entièrement séquencé et annoté. Cette souche a été isolée d'un cas de bactériémie néonatale fatale. Etant donné que le transcriptome est un ensemble dynamique dont la composition varie en fonction des conditions physiologiques, le RNA-seq de *S. agalactiae* NEM316 a été effectué dans plusieurs conditions de culture afin d'avoir une vision plus large sur l'expression des ARN et de leur régulation. Ces conditions englobent des cultures issues de différentes phases de croissance (exponentielle et exponentielle tardive), des mutants de TCS (CiaRH, CovRS et RelRS) ou des cultures en condition acide.

La première partie de ma thèse aborde ainsi la caractérisation de l'ARNrég CetR en mettant en évidence, pour la première fois, le rôle d'un ARNrég dans la résistance aux PAM et dans la virulence de *S. agalactiae*. Nous présentons tout d'abord les caractéristiques de CetR, puis les phénotypes liés à sa délétion.

B. Caractéristiques de l'ARNrég CetR

Les données issues de séquençage massif du transcriptome de *S. agalactiae* NEM316 ont révélé la présence de 47 ARNrég localisés dans des régions intergéniques, notamment Srn073, renommé CetR. Le choix d'étudier CetR a été initialement basé sur sa grande taille (680 nt) et son expression abondante (Figure 30. A). Une recherche *in silico*, effectuée via le serveur ExPASy (www.expasy.org), a prédit la présence d'un peptide de 17 acides aminés (Figure 30. B). Ce peptide est hydrophobe avec un poids moléculaire de 2,02 kD. Nous avons montré que, même si ce peptide est exprimé, il n'impacte pas le rôle de CetR dans les phénotypes décrits dans la publication N° 01.

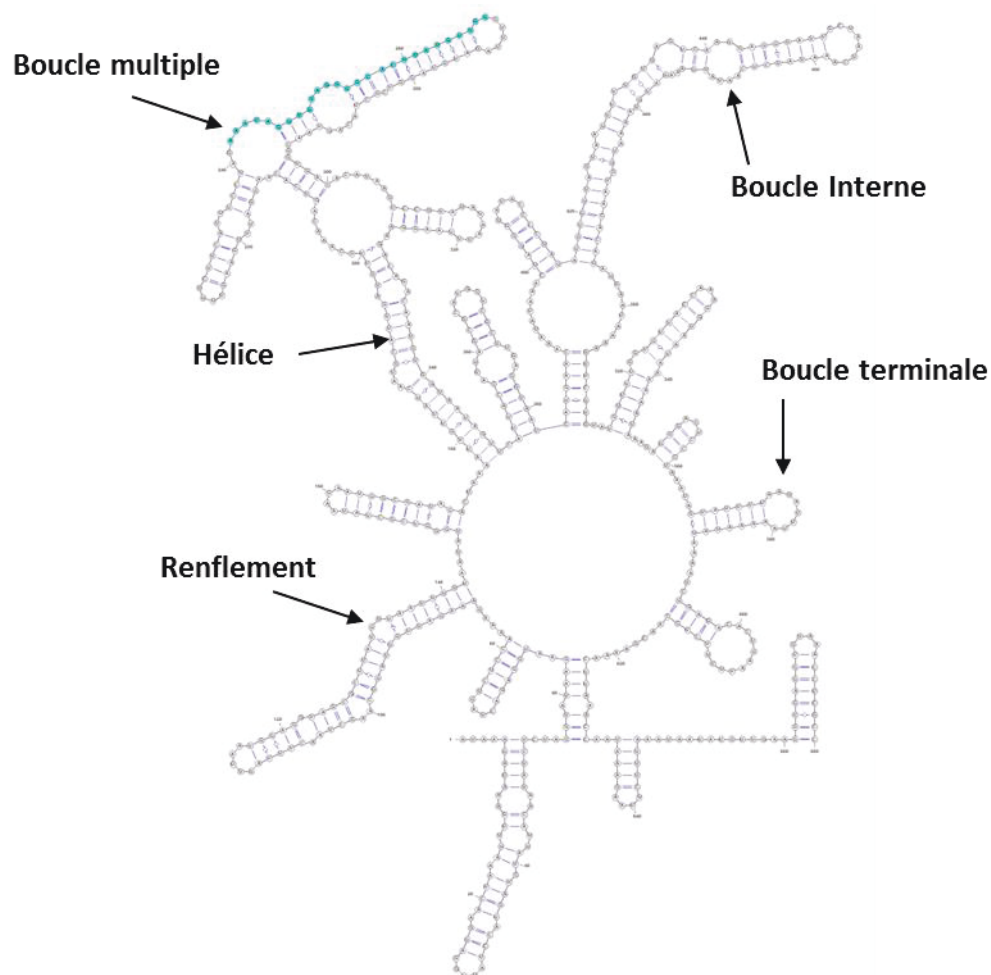


Figure 31 : Structure secondaire potentielle de l'ARNrég CetR.

La structure secondaire de CetR a été prédite en utilisant le serveur interne de RNashapes (<http://bibiserv.techfak.uni-bielefeld.de/rnashapes/>). Le paramétrage pour le MFE (Minimum Free Energy) était de 0,01 Kcal/mol.

Nous avons montré, par northern blot (en collaboration avec P. BOULOC) et par qRT-PCR, que l'expression de CetR dans la souche *S. agalactiae* NEM316 est modulée par la phase de croissance (Figure 30. C et 30. D). Les deux profils d'expression montrent que CetR est principalement exprimé au début de la phase stationnaire de croissance et que son expression est maintenue jusqu'à la phase stationnaire tardive. Le niveau d'expression de CetR est très supérieur à celui de *gyrB*, utilisé comme gène de référence pour la qPCR, soulignant son rôle possible dans les mécanismes d'adaptation de la bactérie lors de son passage de la phase exponentielle à la phase stationnaire de croissance. Nous avons montré également que l'expression de CetR est augmentée après un stress d'une durée de 30 min en NaCl 1M (Figure 37). Ce type d'induction en réponse à un stress est souvent révélateur des ARNrég potentiellement impliqués dans la réponse de la bactérie au changement des conditions environnementales.

Nous avons également prédit la structure secondaire de l'ARNrég CetR *via* le serveur interne RNashapes (Figure 31, collaboration avec Stefan Janssen). L'algorithme de ce serveur est basé sur la programmation dynamique, où tous les appariements possibles sont considérés, pour déduire ceux minimisant l'énergie thermodynamique globale. L'idée est que la structure réelle soit celle de l'énergie libre minimale. Cette structure secondaire correspond à la forme (ou topologie) induite par l'ensemble des appariements A-U, G-C (liaison hydrogène Watson-Crick) et G-U (liaison hydrogène Wooble) de la molécule simple brin. La structure secondaire de CetR est composée de plusieurs régions appariées, appelées hélices (ou tiges), et des régions non-appariées appelées boucles. On distingue quatre types de boucles : 1) des boucles internes, qui relient deux hélices, peuvent être soit symétriques (les deux brins composant la boucle sont de tailles égales), soit non-symétriques ; 2) les renflements, reliant deux hélices par un seul brin non-apparié, les deux hélices étant côte à côte sur l'autre brin ; 3) les boucles terminales, situées à l'extrémité d'un "bras" de la structure ; et 4) les boucles multiples, qui relient plusieurs hélices (trois ou plus) (Figure 30). On retrouve ainsi plusieurs motifs dans cette structure, notamment des motifs en « tige-boucle » ou « épingle de cheveux ». Il est possible que certains de ces motifs de structure secondaire puissent jouer des rôles importants dans certains processus de régulation.

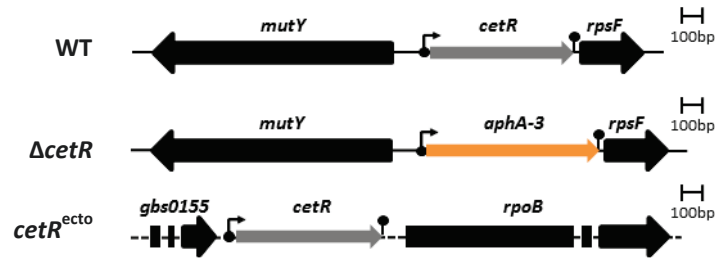


Figure 32 : Diagramme représentant l'environnement génétique de *cetR*. La délétion de *cetR* ($\Delta cetR$) et son remplacement avec une cassette de résistance à la Kanamycine (*aphA-3*), et sa complémentation ectopique sur le chromosome (*cetR^{ecto}*) de *S. agalactiae* NEM316.

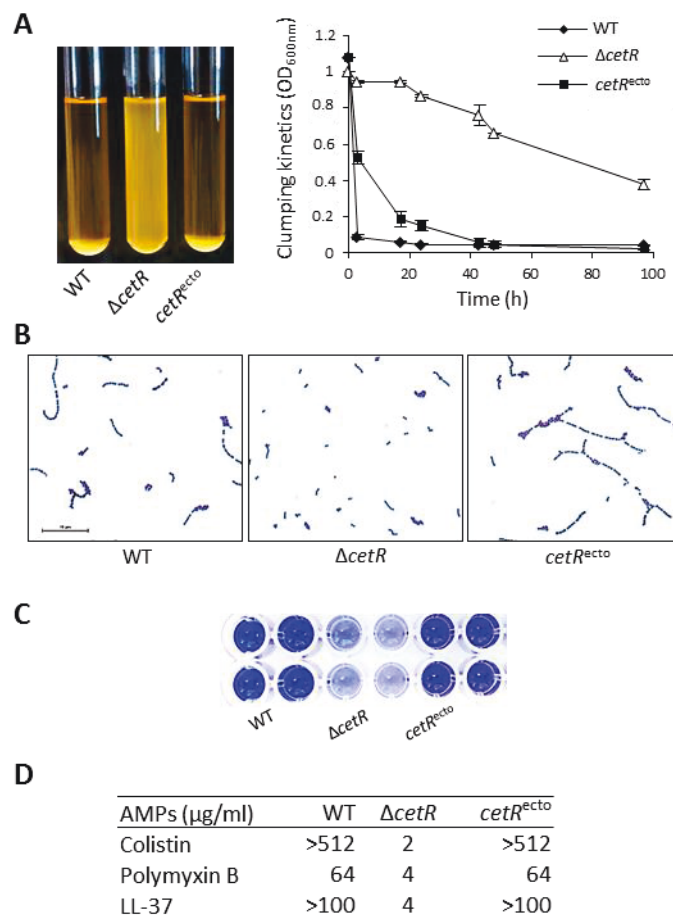


Figure 33 : Rôle de CetR dans la physiologie de *S. agalactiae*.

(A) Images de cultures bactériennes après 24h d'incubation statique (partie gauche) et suivi de la cinétique de sédimentation en culture statique en TH et à 37°C (partie droite). (B) Images en microscopie optique de la coloration de Gram des cultures après 24h d'incubation. (C) Image de la plaque 96 puits issue du suivi de la formation de biofilm après 24h d'incubation à 37°C. (D) Concentration minimale inhibitrice (CMI) de la colistine, la polymyxine et le LL-37.

C. Rôle de CetR dans l'intégrité cellulaire de *S. agalactiae*

Cette partie de ma thèse vise à comprendre le rôle de CetR dans la physiologie de *S. agalactiae*, son adaptation à son environnement et sa virulence. Nous avons d'abord construit le mutant de délétion *cetR* ($\Delta cetR$) en le remplaçant par une cassette de résistance à la Kanamycine (Figure 32). Afin de vérifier que la délétion de *cetR* n'a eu aucun effet polaire sur l'expression du gène *rpsF* (30S ribosomal subunit protein S6) situé en aval et orienté dans le même sens que *cetR*, nous avons réalisé la complémentation du mutant de délétion *cetR* par deux méthodes différentes. Nous avons réalisé : 1) une complémentation plasmidique en clonant *cetR* avec son promoteur endogène dans le plasmide à haut nombre de copies, pJIM2246 (Renault *et al.*, 1996), 2) une complémentation chromosomique en position ectopique en insérant *cetR* sous le contrôle de son promoteur endogène, dans la région intergénique située entre les gènes *gbs0155* et *rpoB* (Figure 32).

La délétion de *cetR* a conduit à la perte de sédimentation de la souche dans le milieu TH (Figure 33. A, partie gauche). Le suivi de la cinétique de sédimentation en fonction du temps montre que la souche sauvage et la souche complémentée sédimentent beaucoup plus rapidement que la souche $\Delta cetR$ (Figure 33. A, partie droite). L'analyse de la morphologie cellulaire, en microscopie optique, a montré que la perte de sédimentation du mutant *cetR* est due principalement à une anomalie dans la formation des chaînettes (Figure 33. B). La souche sauvage (wild-type, WT) de *S. agalactiae* forme de longues chaînettes en milieu liquide conduisant à sa sédimentation, tandis que le mutant *cetR* forme des mono- ou diplocoques. Cet impact sur la morphologie de la bactérie pourrait être lié à une fragilisation de la paroi cellulaire, un dysfonctionnement des enzymes impliquées dans la voie de biosynthèse de la membrane, ou dans la division cellulaire. Cette hypothèse implique que l'ARNrég CetR pourrait avoir plusieurs cibles potentielles liées au maintien de l'intégrité cellulaire de *S. agalactiae* et donc à sa virulence.

Nous avons étudié également la capacité de *S. agalactiae* à former du biofilm. Le test de cinétique de formation du biofilm est basé sur la quantification de l'absorbance de bactéries colorées avec le cristal violet. Nos résultats ont montré que le mutant *cetR* forme significativement moins de biofilm que la souche sauvage et la souche *cetR*^{ecto} (Figure 33. C).

Comme nous l'avons cité précédemment, l'espèce *S. agalactiae* est naturellement résistante aux PAM. Cette résistance lui confère la capacité d'échapper au système immunitaire de la cellule hôte. Nous avons testé la résistance de *S. agalactiae* à différents PAM, tels que la colistine, la polymyxine ou le LL-37. Les concentrations minimales inhibitrices (CMI) obtenues montrent que la souche $\Delta cetR$ est plus sensible à la colistine (CMI = 2µg/mL) que la souche WT ou la souche *cetR*^{ecto} (Figure 33. D ; CMI ≥ 256µg/mL). Des résultats similaires ont été obtenus avec la polymyxine et le LL-37. Chez *S. agalactiae*, la résistance aux PAM est conférée par la voie de D-alanylation des acides lipoteichoïques (ALT). Les gènes impliqués dans cette voie appartiennent à l'opéron *dltABCD*. Nous pouvons ainsi supposer que la diminution de la résistance aux PAM, observée pour le mutant *cetR*, pourrait être due à un dysfonctionnement dans la voie de D-alanylation des ALT.

D. Publication N° 1

Dans cette publication, nous avons détaillé, entre autre, l'identification de la cible de CetR liée aux phénotypes observés.

Virulence and antimicrobial peptide resistance in *Streptococcus agalactiae* promoted by CetR, a long regulatory RNA

Mohamed Amine Zorgani^a, Tatiana Rochat^{b,c}, Asmaa Tazi^{d,e}, Chantal Bohn^c, Aurélia Hiron^a, Prishila Ponien^f, Isabelle Rosinski-Chupin^g, Patrick Trieu-Cuot^h, Philippe Glaser^g, Eric Jacquet^f, Roland Quentin^{a,i}, Claire Poyart^{d,e}, Philippe Boulloc^{c,1}, Marie-Frédérique Lartigue^{a,c,i,1}

^a ISP, Université François Rabelais Tours, INRA, UMR1282, F-37032 Tours, France.

^b VIM, INRA, Université Paris-Saclay, 78350, Jouy-en-Josas, France

^c Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Université Paris-Sud, Université Paris-Saclay, 91400 Orsay, France

^d Institut Cochin, U1016 Inserm/UMR8104, CNRS/UMR-S8104, 75014 Paris, France.

^e Université Paris Descartes, 75005 Paris, France

^f Institut de Chimie des Substances Naturelles (ICSN), CNRS, UPR 2301, Université Paris-Saclay, 91190 Gif-sur-Yvette, France.

^g Institut Pasteur, Unité Ecologie et Evolution de la Résistance aux Antibiotiques, CNRS ERL3526, 75015 Paris, France.

^h Institut Pasteur, Unité de Biologie des Bactéries Pathogènes à Gram-Positif, CNRS UMR3525, 75015 Paris, France.

ⁱ Centre Hospitalier Régional Universitaire de Tours, Service de bactériologie virologie et hygiène hospitalière, F-37044 Tours, France.

¹To whom correspondence should be addressed. Email: philippe.boulloc@i2bc.paris-saclay.fr and marie-frederique.lartigue@univ-tours.fr

Running title: CetR RNA promotes virulence and AMP resistance

Virulence and antimicrobial peptide resistance in *Streptococcus agalactiae* promoted by CetR, a long regulatory RNA

Mohamed Amine Zorgani^a, Tatiana Rochat^{b,c}, Asmaa Tazi^{d,e}, Chantal Bohn^c, Aurélia Hiron^a, Prishila Ponien^f, Isabelle Rosinski-Chupin^g, Patrick Trieu-Cuot^h, Philippe Glaser^g, Eric Jacquet^f, Roland Quentin^{a,i}, Claire Poyart^{d,e}, Philippe Bouloc^{c,1}, Marie-Frédérique Lartigue^{a,c,i,1}

^a ISP, Université François Rabelais Tours, INRA, UMR1282, F-37032 Tours, France. ^b VIM, INRA, Université Paris-Saclay, 78350, Jouy-en-Josas, France. ^c Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Université Paris-Sud, Université Paris-Saclay, 91400 Orsay, France. ^d Institut Cochin, U1016 Inserm/UMR8104, CNRS/UMR-S8104, 75014 Paris, France. ^e Université Paris Descartes, 75005 Paris, France. ^f Institut de Chimie des Substances Naturelles (ICSN), CNRS, UPR 2301, Université Paris-Saclay, 91190 Gif-sur-Yvette, France. ^g Institut Pasteur, Unité Ecologie et Evolution de la Résistance aux Antibiotiques, CNRS ERL3526, 75015 Paris, France. ^h Institut Pasteur, Unité de Biologie des Bactéries Pathogènes à Gram-Positif, CNRS UMR3525, 75015 Paris, France. ⁱ Centre Hospitalier Régional Universitaire de Tours, Service de bactériologie virologie et hygiène hospitalière, F-37044 Tours, France.

Submitted to Proceedings of the National Academy of Sciences of the United States of America

Antimicrobial peptides (AMPs) are essential components of the innate immune response against bacterial infections in human and most eukaryotic hosts. The human opportunistic pathogen *Streptococcus agalactiae* avoids AMPs toxicity by increasing D-alanylation of lipoteichoic acids and thickness of the cell wall. Here, we show that a 680 nt-long regulatory RNA, CetR (cell-envelope-targeting RNA), modulates AMPs resistance and virulence through posttranscriptional regulation of *dltD* mRNA which encodes a D-alanyl-lipoteichoic acid biosynthesis protein. Deletion of *cetR* leads to cell morphology changes, reduced biofilm formation and AMPs resistance. A 27 nt-long CetR-*dltD* interacting region is predicted *in silico*. Compensatory base pair exchanges in *S. agalactiae* demonstrate that CetR interacts directly with *dltD* mRNA and that disruption of this RNA pairing is sufficient to observe the CetR-associated phenotypes. By quantifying both mRNA and protein, we demonstrate that CetR enhances *dltD* translation and disruption of the CetR/*dltD* mRNA interaction results in a dramatic decrease in DltD protein. Finally, using an infection murine model and quantifying bacterial survival in macrophages, we observe that both CetR and DltD are crucial for *S. agalactiae* virulence.

Antibiotic resistance | bacterial virulence | cell wall | posttranscriptional regulation | colistin

Introduction

The virulence of pathogens relies on their ability to adapt to their hosts, escape defense mechanisms and antimicrobial treatments. A multilayer control ensures tight regulation of these processes to optimize pathogen fitness through the different encountered environments (1). A first line of control is transcriptional with regulators such as two-component systems (TCSs) that respond to physical and chemical perturbations. It is followed by posttranscriptional regulations in which regulatory RNAs may affect RNA elongation, processing/degradation and translation (2).

Streptococcus agalactiae, a Group B *Streptococcus*, is a chain-forming gram-positive opportunistic pathogen for farm animals and humans. It colonizes asymptotically the digestive and genitourinary tracts of 10 to 30% of the human population (3). *Streptococcus agalactiae* is the leading cause of invasive infections in neonates, causing pneumonia, sepsis and meningitis mainly *via* maternal transmission and has emerged as an increasing cause of invasive diseases in immune-compromised and elderly adults (4, 5). Intrapartum antibiotic prophylaxis is used for pregnant women tested positive for *S. agalactiae*, but the emergence of antibiotic resistant strains is of concern (6).

Antimicrobial peptides (AMPs) are produced by certain bacteria, and also by higher organisms as a part of the innate immune

response (7, 8). They have broad spectrum antimicrobial activity leading to disruption of the bacterial cell envelope and leakage of intracellular contents (9). However, as AMPs are mainly cationic, their effects on Gram-positive bacteria are limited by the global positive charge of the cell wall. This charge is conferred by the incorporation of D-alanine residues in lipoteichoic acids (LTAs) and wall teichoic acids (WTAs), which are major cell envelope constituents of Gram-positive bacteria. The *dlt* operon encodes four enzymes (DltA, DltB, DltC, and DltD) that catalyze LTA D-alanylation. The *S. agalactiae* *dlt* operon comprises, in addition to the *dltABCD* genes, two regulatory genes, designated *dltR* and *dltS*, located upstream of *dltA*. In several species, including *S. agalactiae*, inactivation of *dlt* genes results in reduced virulence and higher sensitivity to AMPs (10-12).

Here, we characterize CetR, a newly identified *S. agalactiae* long regulatory RNA that is required for Dlt activity. Deletion of its corresponding gene leads to numerous phenotypes including AMP sensitivity and reduced virulence. We demonstrate that CetR RNA base-pairs to *dltD* mRNA to stimulate its translation. We propose that drugs targeting regulatory RNAs such as CetR

Significance

Emergence of antimicrobial resistance is an increasing threat worldwide and failure of treatments urges the discovery of new means to combat against multi-resistant strains. *Streptococcus agalactiae* is an opportunistic pathogen responsible for neonatal infections. We report that CetR, a long regulatory RNA of *S. agalactiae*, contributes to the regulation of virulence and antimicrobial peptides (AMP) resistance. CetR binds to the *dltD* mRNA coding region to stimulate its translation. The inability of *dltD* mRNA to associate by base-pairing to CetR prevents *dltD* expression. DltD contributes to the final step of D-alanine transfer onto lipoteichoic acid and its absence alters the bacterial envelope leading to AMP sensitivity. Preventing the action of this regulatory RNA may significantly increase the activity of antimicrobial compounds.

Reserved for Publication Footnotes

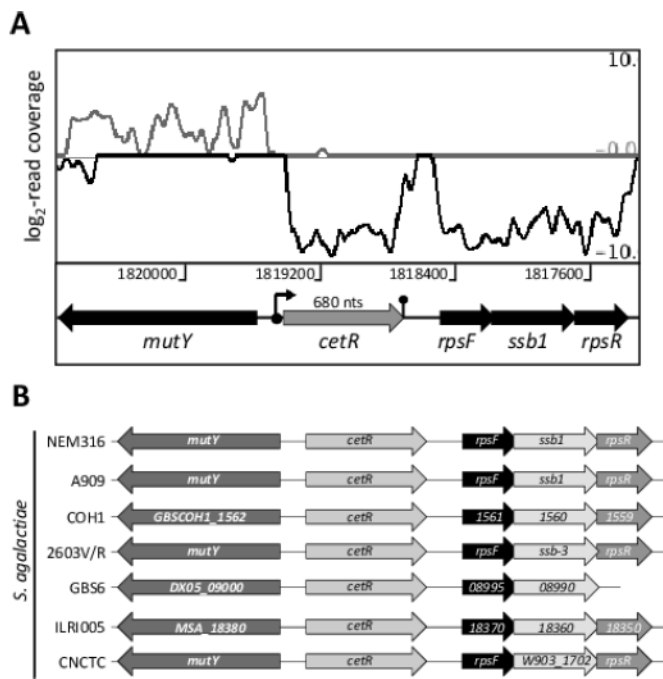


Fig. 1. CetR, a long and abundant RNA. (A) Transcriptome profile of *S. agalactiae* showing CetR genetic context and abundance. NEM316 strain (NC.004368) RNA-seq results are visualized using the Artemis genome browser (35). The *cetR* NEM316 genomic coordinates are: 1.818.732-1.819.411. (B) Diagram representing the alignments of NEM316 *cetR* region with several other *S. agalactiae* strain genomes. Homologous genes in different *Streptococcus* strains are represented by the same great scale, whereas unrelated genes are indicated in white.

would affect pathogens by reducing their virulence and increasing their susceptibility to AMPs.

Results

CetR is a Long RNA Conserved in *Streptococcus agalactiae*

We reported the identification of 47 putative *trans*-acting regulatory RNAs in our previous differential RNA-seq (dRNA-seq) study of *S. agalactiae* strain NEM316 (13). Among them, Srn073 hereafter called CetR for cell-envelope-targeting RNA, is highly expressed and remarkably long (680 nucleotides [nts]), Fig. 1A). Its corresponding gene, *cetR*, is located between *rpsF* and *mutY*. The longest putative open reading frame predicted in *cetR* sequence is a 17 amino acid peptide that is not conserved in other species (Fig. S1A). As discussed hereafter, the CetR-associated phenotypes that we report are not due to peptide expression; we therefore consider that CetR is likely a long regulatory RNA. Sequence homology searches using all available *S. agalactiae* shotgun and complete genome sequences reveal that *cetR* is conserved in *S. agalactiae*, with a low degree of polymorphism (99% nt identity; Figs. 1B and S1). Thus, *cetR* belongs to *S. agalactiae* regulatory RNA core genes. A long non-coding sequence harboring a predicted terminator is also present between *rpsF* and *mutY* genes of other *Streptococcus* species. The sequence displays weak homologies with *cetR* at about 150 nts downstream of a potential consensus promoter (Fig. S2), suggesting that CetR-like RNAs may exist in a subset of streptococci.

CetR Promotes AMP Resistance, Chain and Biofilm Formation

To identify putative functions associated with CetR, we constructed a strain in which the *cetR* gene was inactivated (Δ *cetR*). In rich medium, growth rate and final biomass are not affected by the absence of CetR (Fig. S3A). The minimum inhibitory concentrations (MICs) for different antibiotics of the Δ *cetR* and parental

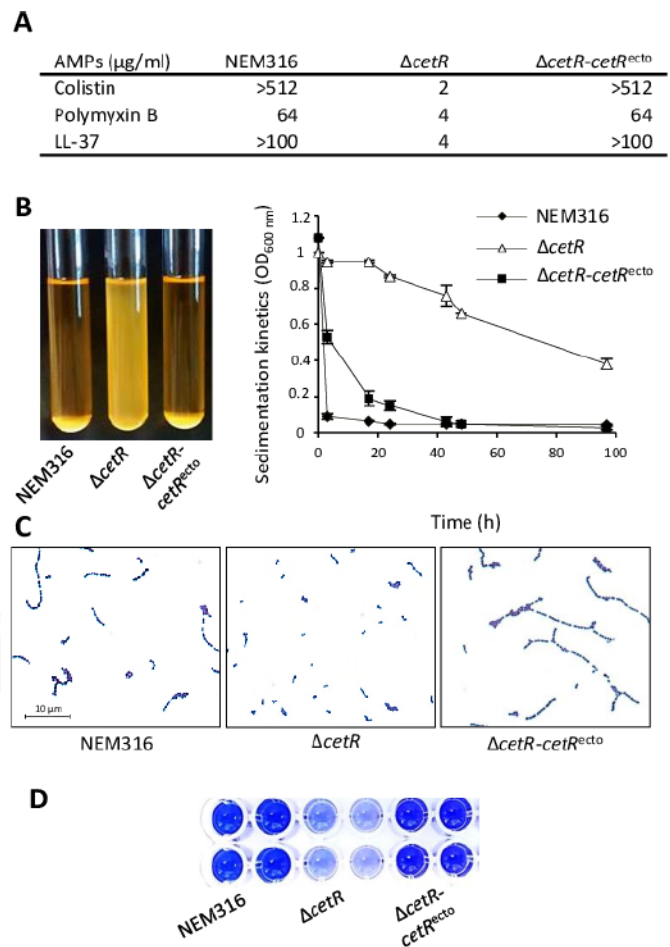


Fig. 2. CetR promotes AMP resistance, chain formation and biofilm. (A) MICs measured for colistin, polymyxin B and LL-37 in NEM316, Δ *cetR* and Δ *cetR*-*cetR*^{ecto}. (B) Bacterial sedimentation on standing tubes of liquid cultures after 24 hours (left panel), and sedimentation kinetics monitored as described in Material and Methods (right panel). (C) Microscopic fields of Gram-colored strains. The Δ *cetR* cells present a reduced chain lengths compared to the wild-type and Δ *cetR* *cetR*^{ecto} strains. (D) Biofilm formation assay of wild-type, Δ *cetR* and Δ *cetR* *cetR*^{ecto} strains on 96-well plastic plates. Bacterial sessile biomass was measured at 24h with crystal violet coloration.

strains were also tested (Figs. 2A, S4 and Table S1). As observed for many Gram-positive bacteria, NEM316 is highly resistant to three tested AMPs, LL-37, colistin and polymyxin B (e.g. colistin MIC > 512 μ g/ml). In contrast, a dramatic decrease of MICs was observed in Δ *cetR* (e.g. > 256-fold for colistin). To determine if reduced AMP resistance in the Δ *cetR* strain was not due to a polar effect of the mutation, we constructed a Δ *cetR*-derivative strain carrying a chromosomal ectopic copy of wild-type *cetR* (referred as *cetR*^{ecto}) inserted between *gbs0155* and *rpoB* (Fig. S3B). In this strain, the AMP resistance phenotype was restored to the wild-type level (Fig. 2A). The CetR requirement for antibiotic resistance was exclusive to AMPs, as MICs of all other antibiotic categories tested were similar for bacteria expressing CetR or not (Table S1). These results demonstrate that CetR promotes a specific resistance to AMPs in *S. agalactiae*.

Interestingly, reduced sedimentation was observed in liquid culture for the Δ *cetR* mutant as compared to the parental strain (Fig. 2B). This phenotype correlates with a morphological change observed for Δ *cetR*, which forms shorter chains, mainly diplococci, as compared to the parental strain (Fig. 2C).

Biofilm formation depends on adhesion of microorganisms to each other and to biotic or abiotic surfaces. The development of

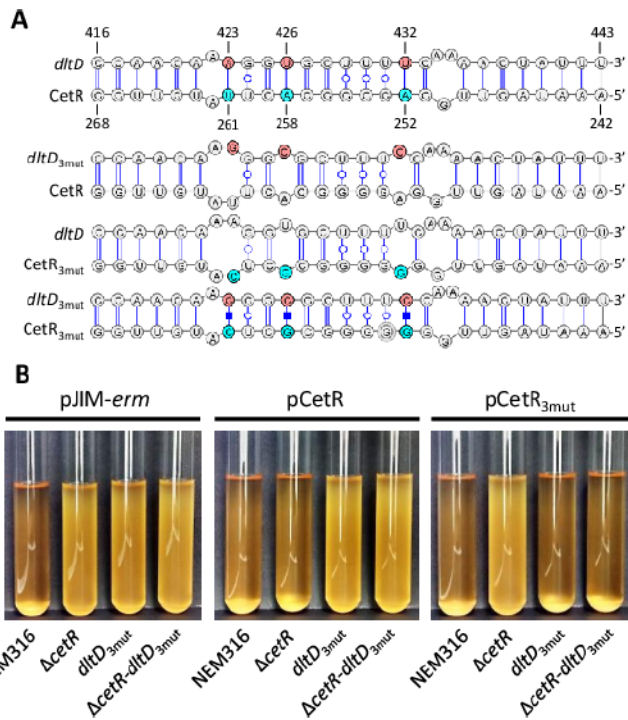


Fig. 3. CetR targets *dltD* mRNA by direct RNA/RNA pairing. (A) Predicted pairing between *dltD* mRNA and CetR, mutated *dltD*_{3mut} and CetR, *dltD* mRNA and mutated CetR (CetR_{3mut}) and, *dltD*_{3mut} and CetR_{3mut}. Mutated *dltD* nucleotides, pink; mutated CetR nucleotides, blue. (B) Sedimentation phenotype of the wild-type, Δ *cetR*, *dltD*_{3mut} and Δ *cetR*-*dltD*_{3mut} strains with the indicated plasmids.

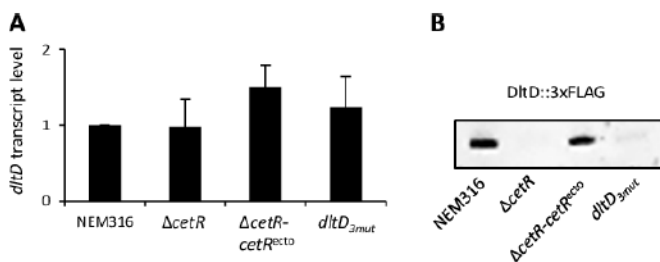


Fig. 4. CetR activates *dltD* mRNA translation. (A) *dltD* mRNA was quantified by quantitative RT-PCR in the indicated strains relative to NEM316. (B) *dltD* translation is evaluated by Western blot with anti-FLAG antibodies against the 3xFLAG epitope on strains having the indicated mutations and the chromosomal copy of *dltD* replaced by a translational *dltD*::3xflag gene fusion.

surface-associated bacteria contributes to virulence and antibiotic recalcitrance (14). The NEM316 strain forms biofilms on 96-well plastic plates. However, the strain deficient for CetR (Δ *cetR*) showed strongly reduced biofilm formation, which was restored by introducing the *cetR*^{ecto} allele (Figs. 2D and S5).

Our results suggest that the long RNA CetR may have a regulatory function affecting the expression of a gene (or genes) involved in cell morphology, cell wall biogenesis and AMP resistance.

CetR Targets the D-Alanyl-Lipoteichoic Acid Biosynthesis Pathway by RNA/RNA Pairing

To identify putative mRNA targets of CetR, a genome-wide *in silico* prediction of CetR-mRNA interactions was performed using RNAPredator (15). The best candidate identified by ranking putative targets by energy was *dltD* mRNA (-21.6 KJ/mol, see pairing region in Fig. 3A).

DltD is a membrane-anchored protein that contributes to cell wall synthesis by mediating D-alanine transfer to the LTA polyglycerophosphate chain (16). This modification is associated with AMP resistance and sedimentation of *S. agalactiae* (11, 17).

In silico prediction and Δ *cetR*-associated phenotypes strongly suggests that CetR modulates *dltD* expression by RNA/RNA pairing interaction. To test this hypothesis, we first examined the phenotypic consequences of disrupting the putative pairing region by mutagenesis. The morphological changes associated with Δ *cetR* are restored in the presence of a plasmid expressing CetR (pCetR) (Figs. S6A and S6B). However, a plasmid derivative carrying a 27-nt deletion of the sequence corresponding to the CetR-*dltD* mRNA predicted interaction region (pCetR _{Δ 27}) could not restore the parental phenotypes (Fig. S6). This experiment supported the *in silico*-predicted CetR-*dltD* mRNA interaction.

To better define the pairing region, we introduced three-nt changes in *dltD*, in *cetR*, or in both genes. Three nts on the *dltD* chromosomal gene were mutated (A423G, T426C and T432C) without changing the DltD amino acid sequence; the resulting allele was named *dltD*_{3mut} (Fig. 3A). Remarkably, the *dltD*_{3mut} strain displayed the same phenotypes as were observed for Δ *cetR*: reduced sedimentation, morphology changes, reduced biofilm formation, and increased AMP sensitivity (Figs. 3B, S4 and S7, and Table S2). These results suggest that the *dltD*_{3mut} allele does not interact with wild-type CetR. When expressed from a plasmid (pCetR), CetR restored sedimentation of the Δ *cetR* mutant to the parental level, but as expected, not of the *dltD*_{3mut} mutant (Fig. 3B). To confirm that the *dltD*_{3mut} phenotype was associated with the inability of *dltD*_{3mut} mRNA to base-pair with CetR, we constructed a mutated *cetR* allele (A252G, A258G and T261C), named *cetR*_{3mut}, restoring base-pair complementarity with *dltD*_{3mut} (Fig. 3A). CetR_{3mut} expressed from the plasmid pCetR_{3mut} restored sedimentation of the *dltD*_{3mut} and Δ *cetR*-*dltD*_{3mut} mutants to the parental level, but not of Δ *cetR* (Fig. 3B). Similarly, the morphology phenotype of short chains and increased AMP sensitivity observed for the Δ *cetR* and *dltD*_{3mut} mutants were lost in the presence of pCetR and pCetR_{3mut}, respectively (Fig. S8 and Table S2).

These experiments demonstrate that CetR is crucial for *dltD*-associated phenotypes (*i.e.* AMP resistance, chain-forming, and biofilm). The CetR activity is mediated by RNA/RNA interaction involving the 242-268 region of CetR and 416-443 of *dltD* mRNA.

CetR Activates *dltD* mRNA Translation

Bacterial regulatory RNAs interacting by base-pairing to mRNAs modulate mRNA stability and translation; in most studied cases, they act as negative regulators (2). Δ *cetR* phenotypes suggest that the absence of CetR decreases *dlt* activity; we therefore considered that the presence of CetR would upregulate *dltD* expression on transcriptional and/or translational levels.

The amount of *dltD* transcript was determined by qRT-PCR in strains NEM316, Δ *cetR*, Δ *cetR* *cetR*^{ecto} and *dltD*_{3mut}. Transcription was not significantly different in the four strains (Fig. 4A). We conclude that the role of CetR is not likely to affect the *dltD* mRNA transcription and stability and therefore, CetR may act translationally.

To evaluate DltD translation, the endogenous *dltD* gene was replaced by a *dltD*-3xflag allele. The latter gene expresses a DltD-3xFLAG fusion protein, which is detected by specific anti-FLAG antibodies. Four strains harboring the chromosomal *dltD*-3xflag allele were constructed (see SI Methods) and expression of DltD-3xFLAG was monitored by Western blot. The DltD-3xFLAG tagged-protein was detected with a strong signal in the NEM316 and Δ *cetR* *cetR*^{ecto} strains but not in Δ *cetR* and *dltD*_{3mut} strains (Fig. 4B). Altogether, our results demonstrate that pairing of CetR with *dltD* mRNA is required for efficient DltD synthesis. Synthesis remains efficient with the duplex formed by CetR_{3mut}/*dltD*_{3mut} mRNA. The RNA/RNA interaction does

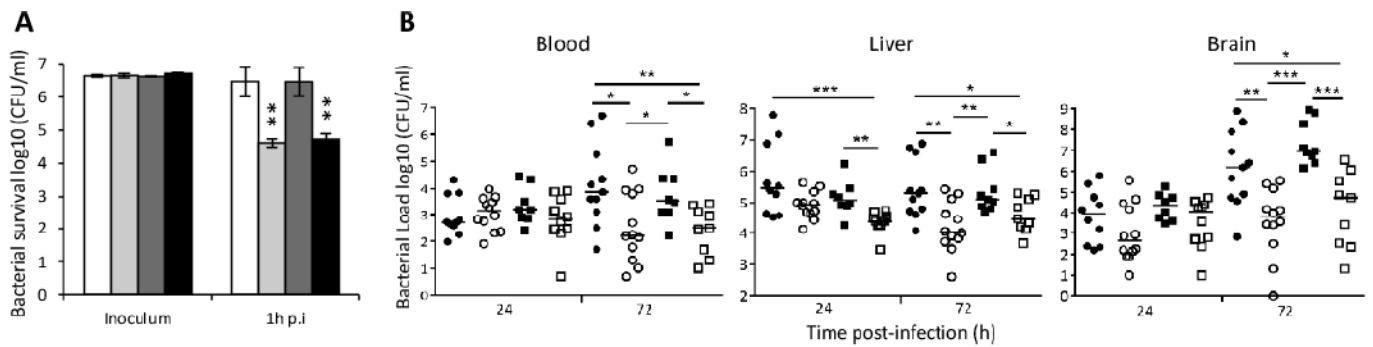


Fig. 5. CetR, a Long RNA Promoting Bacterial Survival in Macrophages and Virulence. (A) Bacterial survival in macrophages RAW 264.7 cells. Bacterial CFUs were counted 1-hour post-infection (1h p.i) and compared to starting inoculums. The error bars represent the standard deviation calculated from at least three independent experiments. The asterisks indicate *p*-values obtained using unpaired Student *t*-test (** *P*<0.01). (B) Mouse virulence assays. Bacterial load in blood, liver and brain of 3 weeks-old mice (*n* ≥ 8) infected with 3×10^6 CFU via intravenous route. NEM316, black circles; Δ *cetR*, white circles; Δ *cetR*-*cetR*^{ecto}, black squares; *dltD*_{3mut}, white squares. Asterisks indicate significant differences as assessed by the Mann-Whitney test (*, *P*<0.05; **, *P*<0.01; ***, *P*<0.001).

not alter the *dltD* mRNA level, but is required to stimulate its translation.

CetR is Required for *S. agalactiae* Survival in Macrophages and Virulence

Production of AMPs is an important component of effective macrophage bactericidal activity. However, expression of LTAs by Gram-positive pathogens leads to AMP resistance. Our findings establishing the direct effects of CetR on DltD expression led us to investigate the roles of CetR on *S. agalactiae* survival in phagocytic cells and in infection. Survival of NEM316, Δ *cetR*, Δ *cetR*^{ecto} and *dltD*_{3mut} exposed to RAW 264.7 macrophages was compared at 1-hour post-infection. Survival was more than 55-fold lower for Δ *cetR* and *dltD*_{3mut} than for NEM316 and *cetR*^{ecto} strains (Fig. 5A).

We then investigated virulence of the Δ *cetR* mutant compared to Δ *cetR*^{ecto}, NEM316, and *dltD*_{3mut} strains in an intravenous mouse infection model. Animals were sacrificed 24 and 72 hours post-infection, and bacterial load was measured in infected tissues. At 72 hours post-infection, bacterial counts in blood, liver and brain were significantly lower in animals infected with Δ *cetR* or *dltD*_{3mut} compared to those infected with NEM316 and Δ *cetR*^{ecto} (Fig. 5B).

We conclude that CetR, likely through its interaction with *dltD* mRNA, contributes to *S. agalactiae* resistance to macrophage killing and survival during infection in a murine model.

Discussion

Regulatory RNAs in bacteria are classically referred to as sRNAs, for small RNAs, because their size is usually between 70 to 300-nts. However, few of them are exceptionally long and RNAIII from *Staphylococcus aureus* is their paradigm (18). This 514-nts-long RNA expresses the 26 amino acid δ -hemolysin and has 14 hairpin structures. It regulates positively and negatively several genes via RNA/RNA interactions involving substrate-dedicated hairpins. We characterized CetR, the longest regulatory RNA from *S. agalactiae* (13) and identified *dltD* mRNA as one of its target. All reported phenotypes associated with Δ *cetR* are explained by a *dltD* down-regulation and involves a critical short pairing region. Since regulatory-RNA pairing regions with targets are usually short, it is likely that different CetR regions nested in its 680-nts interact with other, so far unknown, targets. We do not exclude the possibility that *cetR* encodes a putative small peptide; however, all presented phenotypes are mediated by a RNA/RNA interaction.

Regulatory RNAs can modulate all biological processes and several of them, contribute to virulence. A large scale study on *Streptococcus pneumoniae* indicates that several of them are likely

involved at different steps of pneumococcal pathogenesis (19). Inactivation of these sRNAs alters expression of genes related to transport and metabolism; but, no direct regulation by these sRNAs on putative targets was demonstrated. In *Streptococcus pyogenes*, the three sRNAs Pel, FasX and RivX are implicated in virulence: i) as RNAIII, Pel codes for an hemolysin and regulates the expression of virulence factors via RNA/RNA interaction, ii) FasX binds the 5' end of *ska* mRNA to stimulate the streptokinase expression, and iii) RivX, issued from the RivR mRNA processing, upregulates a regulator of virulence (20). *Streptococcus agalactiae* sRNAs were inventoried, but their targets remain unknown (13, 21). We provide the first evidence of a regulatory RNA involved in the virulence of *S. agalactiae*. Previous reports demonstrated that incorporation of D-alanine residues is important for virulence in animal model (22-24). Here, we demonstrate that the *dltD* mRNA encoding the final step of D-alanine transfers onto LTA (25) is a direct target of CetR. Indeed, mutations within the *dltD* open reading frame keeping intact the DltD amino-acid sequence but abolishing its interaction with CetR (*dltD*_{3mut}) mimic the Δ *cetR* phenotype. We therefore conclude that the attenuated virulence associated with Δ *cetR* is due to its regulatory effect on the *dlt* operon.

Host response to *S. agalactiae* infection includes activation of macrophages and other component of the innate immune system (26). Mouse macrophages produce cathelin-related antimicrobial peptides (27). The efficient clearance of Δ *cetR* and *dltD*_{3mut} mutants in mice is possibly a consequence of their increased susceptibility to AMPs and indeed, these mutants are also more susceptible to killing by macrophages than the wild-type strain.

So far, only few regulatory RNAs were reported to affect antibiotic resistance, but the emergence of multidrug resistant strains raised an interest for them (28). Gram-positive bacteria are usually resistant to AMPs through the increase of the global positive charge of the cell envelope; interestingly, the inactivation of *cetR* results in a dramatic increase in AMP susceptibility. Since defects in D-alanyl-LTA biosynthesis affect the membrane charge and increase AMP susceptibility (10), we conclude that the Δ *cetR* AMP-susceptibility is associated to a down-regulation of *dlt* gene expression. This result is supported by the observation that *dltD*_{3mut} mimics the Δ *cetR* AMP phenotype.

Most *trans*-acting regulatory RNAs downregulate gene expression; however, there is growing number of them contributing to gene activation (29). They can liberate sequestered ribosome binding sites (e.g. DsrA in *Escherichia coli*) and antagonize mRNA processing to increase the stability of targeted mRNAs (e.g. FasX). In most cases, these regulatory RNAs bind to the untranslated regions (UTRs) of their targets. CetR-dependent phenotypes indicate that CetR is a *dltD* activator. It does not

affect significantly the *dltD* mRNA quantity, but stimulates its translation. Unusually, the CetR activity involves a pairing within the *dltD* coding region. One hypothesis is that in the absence of CetR, the structure of the polycistronic *dlt* mRNA prevents *dltD* translation. Thus, CetR binding to the *dltD* coding sequence may generate an alternate *dlt* mRNA structure that free a ribosome binding site (*dltD* or upstream sequences that would act via translational couplings) allowing *dltD* translation and LTA D-alanylation.

Artificial molecules tampering with riboswitches can have antimicrobial activities (30) and RNA-based molecules (e.g. artificial *trans*-encoded sRNA) can silence specific bacterial genes (31, 32). Such compounds are one step forward to the development of new therapeutic agents. Thereby, finding a molecule preventing the binding of CetR to *dltD* mRNA would be a mean to generate a reduced virulence and susceptibility to AMP of *S. agalactiae*. This concept applies to other regulatory RNAs in different bacteria.

Material and Methods

Bacterial Strains, Growth Conditions, Plasmids and Primers

Serotype III *S. agalactiae* strain NEM316 from a human case of bacteremia was the parental strain used for this study (33). Mutant derivatives were constructed as described (SI Material and Methods). *E. coli* was used for cloning purposes. All *S. agalactiae* and *E. coli* strains are listed (SI Table S3).

Streptococcus agalactiae and *E. coli* strains were routinely grown at 37°C statically in Todd Hewitt broth (TH, Difco) and under agitation in Luria Bertani broth (LB, MP Biomedicals), respectively. When necessary, erythromycin was added to the media at final concentrations: 10 µg/mL and 150 µg/mL for *S. agalactiae* and *E. coli*, respectively.

Plasmid constructions is described (SI Material and Methods). Plasmids and DNA primers are listed (SI Table S4 and S5, respectively).

Sedimentation Kinetics and Morphology Observation

Bacterial sedimentation was measured according to previous report with modification (34). Cells were grown overnight, diluted 1:100 into 45 mL of fresh medium and grown statically at in 50 mL

conical tubes to stationary phase. Shaking, pipetting or vortexing were avoided to allow bacterial sedimentation. Samples of 500 µL were taken from the top of each tube at different time points for OD_{600 nm} measurement. To observe morphological changes or bacterial aggregation, Gram coloration and photonic microscopy observations were performed after 24 h of incubation. Results reported are from three independent experiments.

Biofilm Assay

Streptococcus agalactiae strains were grown overnight and diluted 100-fold in the same medium. Cells were then prepared for biofilm assay as described previously (34). At OD_{600 nm}=0.5, 200 µL culture aliquots were dropped into 96-well plates. Plates were then incubated at 37 °C, and bacterial sessile development was evaluated with crystal violet coloration at different time points. Negative control wells were filled with sterile medium.

Antibacterial Activity

Antibiotic MICs were measured according to EUCAST guidelines (www.eucast.org). Bacteria were plated on Mueller Hinton agar supplemented with defibrinated horse blood and 20 mg/L β-NAD (MHF, BioMérieux). Antibiotic susceptibility was tested by disks containing antibiotics (Bio-Rad) or Etest strips (BioMérieux). For human cathelicidin AMP (LL-37, InvivoGen), MICs were determined by a dilution method in 96-well microplates (Argos Technologies). Bacteria (10⁶ CFU) were added to wells containing increasing LL-37 concentrations and incubated overnight at 37 °C; bacterial growth was determined by OD using a microplate reader (BioTec).

Acknowledgments

We thank Sandy Gruss for critical reading of the manuscript and helpful discussions. We thank Emilie Camiade, Philippe Gilot, Daniel Niquet, Laurent Mereghetti, Annick Jacq for helpful discussions and Paul Aner for warm support. We thank Guillaume Desoubreaux and Eric Morello for technical advice and providing materials. This work was supported by the Agence Nationale pour la Recherche (http://www.agence-nationale-recherche.fr/, ANR-12-BSV6-0008 'ReadRNA') and was funded in part by R.Q. Valorisation Contrats - Université François Rabelais. M.A.Z. was the recipient of a PhD fellowship from the French Ministère de l'Éducation et de l'Enseignement Supérieur.

- Vasil ML & Darwin AJ (2013) *Regulation of bacterial virulence* (ASM Press, Washington, DC) pp xvi, 603 p.
- Storz G, Vogel J, & Wassarman KM (2011) Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell* 43(6):880-891.
- van der Mee-Marquet N, et al. (2008) Molecular characterization of human-colonizing *Streptococcus agalactiae* strains isolated from throat, skin, anal margin, and genital body sites. *J Clin Microbiol* 46(9):2906-2911.
- Stoll BJ, et al. (2011) Early onset neonatal sepsis: the burden of group B Streptococcal and *E. coli* disease continues. *Pediatrics* 127(5):817-826.
- Farley MM (2001) Group B streptococcal disease in nonpregnant adults. *Clin Infect Dis* 33(4):556-561.
- Poyart C, Jardy L, Quesne G, Berche P, & Trieu-Cuot P (2003) Genetic basis of antibiotic resistance in *Streptococcus agalactiae* strains isolated in a French hospital. *Antimicrob Agents Chemother* 47(2):794-797.
- Hassan M, Kjos M, Nes IF, Diep DB, & Lotfpour F (2012) Natural antimicrobial peptides from bacteria: characteristics and potential applications to fight against antibiotic resistance. *Journal of applied microbiology* 113(4):723-736.
- Kuroda K, Okumura K, Isogai H, & Isogai E (2015) The Human Cathelicidin Antimicrobial Peptide LL-37 and Mimics are Potential Anticancer Drugs. *Frontiers in oncology* 5:144.
- Brogden KA (2005) Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat Rev Microbiol* 3(3):238-250.
- Malanovic N & Lohner K (2016) Gram-positive bacterial cell envelopes: The impact on the activity of antimicrobial peptides. *Biochim Biophys Acta* 1858(5):936-946.
- Poyart C, Lamy MC, Boumaila C, Fiedler F, & Trieu-Cuot P (2001) Regulation of D-alanyl-lipoteichoic acid biosynthesis in *Streptococcus agalactiae* involves a novel two-component regulatory system. *J Bacteriol* 183(21):6324-6334.
- Poyart C, et al. (2003) Attenuated virulence of *Streptococcus agalactiae* deficient in D-alanyl-lipoteichoic acid is due to an increased susceptibility to defensins and phagocytic cells. *Molecular Microbiology* 49(6):1615-1625.
- Rosinski-Chupin I, et al. (2015) Single nucleotide resolution RNA-seq uncovers new regulatory mechanisms in the opportunistic pathogen *Streptococcus agalactiae*. *BMC Genomics* 16:419.
- Lebeaux D, Ghigo JM, & Beloin C (2014) Biofilm-related infections: bridging the gap between clinical management and fundamental aspects of recalcitrance toward antibiotics. *Microbiology and molecular biology reviews : MMBR* 78(3):510-543.
- Eggenhofer F, Tafer H, Stadler PF, & Hofacker IL (2011) RNApredator: fast accessibility-based prediction of sRNA targets. *Nucleic Acids Res* 39(Web Server issue):W149-154.
- Reichmann NT, Cassona CP, & Grundling A (2013) Revised mechanism of D-alanine incorporation into cell wall polymers in Gram-positive bacteria. *Microbiology* 159(Pt 9):1868-1877.
- Saar-Dover R, et al. (2012) D-alanylation of lipoteichoic acids confers resistance to cationic peptides in group B streptococcus by increasing the cell wall density. *PLoS Pathog* 8(9):e1002891.
- Felden B, Vandenesch F, Boulou P, & Romby P (2011) The *Staphylococcus aureus* RNome and its commitment to virulence. *PLoS Pathog* 7(3):e1002006.
- Mann B, et al. (2012) Control of virulence by small RNAs in *Streptococcus pneumoniae*. *PLoS Pathog* 8(7):e1002788.
- Brantl S & Bruckner R (2014) Small regulatory RNAs from low-GC Gram-positive bacteria. *RNA Biol* 11(5):443-456.
- Pichon C, du Merle L, Caliot ME, Trieu-Cuot P, & Le Bouguenec C (2012) An *in silico* model for identification of small RNAs in whole bacterial genomes: characterization of antisense RNAs in pathogenic *Escherichia coli* and *Streptococcus agalactiae* strains. *Nucleic Acids Res* 40(7):2846-2861.
- Peschel A, et al. (1999) Inactivation of the *dlt* operon in *Staphylococcus aureus* confers sensitivity to defensins, protegrins, and other antimicrobial peptides. *J Biol Chem* 274(13):8405-8410.
- Abachin E, et al. (2002) Formation of D-alanyl-lipoteichoic acid is required for adhesion and virulence of *Listeria monocytogenes*. *Mol Microbiol* 43(1):1-14.
- Collins LV, et al. (2002) *Staphylococcus aureus* strains lacking D-alanine modifications of teichoic acids are highly susceptible to human neutrophil killing and are virulence attenuated in mice. *J Infect Dis* 186(2):214-219.
- Percy MG & Grundling A (2014) Lipoteichoic acid synthesis and function in gram-positive bacteria. *Annu Rev Microbiol* 68:81-100.
- Maisey HC, Doran KS, & Nizet V (2008) Recent advances in understanding the molecular basis of group B *Streptococcus* virulence. *Expert Rev Mol Med* 10:e27.
- Gallo RL, et al. (1997) Identification of CRAMP, a cathelin-related antimicrobial peptide expressed in the embryonic and adult mouse. *J Biol Chem* 272(20):13088-13093.
- Laloua D, Eyraud A, Chabelskaya S, Felden B, & Masse E (2014) Regulatory RNAs involved in bacterial antibiotic resistance. *PLoS Pathog* 10(8):e1004299.
- Papenfors K & Vanderpool CK (2015) Target activation by regulatory RNAs in bacteria.

FEMS Microbiol Rev 39(3):362-378.

30. Kim JN, *et al.* (2009) Design and antimicrobial action of purine analogues that bind Guanine riboswitches. *ACS chemical biology* 4(11):915-927.
31. Man S, *et al.* (2011) Artificial trans-encoded small non-coding RNAs specifically silence the selected gene expression in bacteria. *Nucleic acids research* 39(8):e50.
32. Park H, Bak G, Kim SC, & Lee Y (2013) Exploring sRNA-mediated gene silencing mechanisms using artificial small RNAs derived from a natural RNA scaffold in *Escherichia coli*. *Nucleic acids research* 41(6):3787-3804.
33. Glaser P, *et al.* (2002) Genome sequence of *Streptococcus agalactiae*, a pathogen causing invasive neonatal disease. *Mol Microbiol* 45(6):1499-1513.
34. Chagnot C, *et al.* (2013) In vitro colonization of the muscle extracellular matrix components by *Escherichia coli* O157:H7: the influence of growth medium, temperature and pH on initial adhesion and induction of biofilm formation by collagens I and III. *PLoS One* 8(3):e59386.
35. Carver T, Harris SR, Berriman M, Parkhill J, & McQuillan JA (2012) Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* 28(4):464-469.

Submission PDF

Additional Methods

Strain constructions

Mutant strains were constructed by allelic exchanges via a two-step homologous recombination process using pG+host5 derivatives as described (1).

The *cetR* deletion (Fig. S3A) was obtained using p Δ *cetR*::*aphA-3*, a pG+host5 derivative constructed as follow. The *aphA-3* gene conferring kanamycin resistance was PCR-amplified from pBT2 with primers kan-F/kan-R. The *cetR* upstream and downstream regions were PCR-amplified with primers Srn073-UpF/Srn073-UpR and Srn073-downF/Srn073-downR, respectively. pG+host5 was PCR-amplified using primers pG_{host}-F/pG_{host}-R. The four PCR products were assembled by Gibson assembly (2). The Δ *cetR*::*aphA-3* allele in NEM316 generated Δ *cetR* (MFL100) strain.

The ectopic complementation of Δ *cetR* strain was performed by inserting a *cetR* wild-type copy in the chromosome at a different position than its endogenous copy using the pG+host5 derivative pCetR^{ecto}. The position between 178811-178812nts of the NEM316 strain (NC_004368) was chosen for the absence of transcriptional signals by RNA-seq. pCetRecto was constructed as follow. *cetR* was PCR-amplified with primers cetR-ectoF/cetR-ectoR. The amplified sequence includes the *cetR* native promoter and Rho-independent terminator (Fig. S3A). The upstream (*gbs0155*) and downstream (*rpoB*) regions of position 178811 were PCR-amplified with primers IGR-UpF/ IGR-UpF R and IGR-DownF/IGR-DownR, respectively. pG+host5 was PCR amplified using primers pG_{host}-F/pG_{host}-R. The three PCR products were assembled by Gibson assembly (2). The *cetR*^{ecto} allele introduced into the Δ *cetR* strain generated Δ *cetR cetR*^{ecto} (MFL101).

The *dltD* chromosomal gene was substituted by the *dltD*_{3mut} allele using a pG+host5 derivative, pDltD3mut, constructed as follow. Two fragments of *dltD* were PCR-amplified with primers *dltD*_{3mut}-UpF/ *dltD*_{3mut}-UpR and *dltD*_{3mut}-DownF/*dltD*_{3mut}-DownR. The primers were designed to introduce three point mutations, A423G, T426C and T432C. These mutations are conservative and therefore do not affect the *dltD* coding sequence. pG+host5 was PCR amplified using primers pG_{host}-F/pG_{host}-R. The three PCR products were assembled by Gibson assembly (2). The *dltD*_{3mut} allele in NEM316 and Δ *cetR* strains generated *dltD*_{3mut} (MFL102) and Δ *cetR-dltD*_{3mut} (MFL103) strains.

The translational expression of *dltD* and *dltD*_{3mut} was evaluated by means of a DltD::3xFlag fusion protein in which 3xFlag is at the C-term end of the protein. The chromosomal copy of *dltD* was substituted by *dltD3xflag* using pDltD3xFlag, a pG+host5 derivative constructed as follow. The *3xflag* sequence was PCR-amplified from pTOOX with primers 3xFLAG-F/3xFLAG-R. Fragment of about 1kb upstream and downstream from the *dltD* stop codon were PCR-amplified from NEM316 genomic DNA with primers *gbs1829*-F/ *gbs1829*-R and *dltD*-F/ *dltD*-R, respectively; pG+host5 was PCR-amplified using primers pGhost-F/pGhost-R. The three PCR products were assembled by Gibson assembly (2). The *dltD3xflag* allele introduced into the Δ *cetR* and Δ *cetR cetR^{ecto}* strain generated Δ *cetR dltD3xflag* (MFL110) and Δ *cetR cetR^{ecto} dltD3xflag* (MFL111), respectively. The chromosomal copy of *dltD* was substituted by *dltD*_{3mut}*3xflag* using pDltD_{3mut}3xFlag, a pG+host5 derivative constructed as pDltD3xFlag, except that upstream and downstream sequences from the *dltD* stop codon were PCR-amplified from *dltD*_{3mut} (MFL102) genomic DNA. The *dltD*_{3mut} *3xflag* allele introduced into NEM316 generated *dltD*_{3mut} *3xflag* (MFL112).

Plasmids were constructed into *E. coli*, verified by DNA sequencing of the cloned regions and transferred into the appropriate NEM316 *S. agalactiae* derivative. For strain constructions, the genomic DNA of mutants were isolated and the presence of appropriate mutations was checked by PCR-amplification of the concerned regions and sequencing (for primers, see Table S5).

Construction of pCetR, pCetR_{Δ27} and pCetR_{3mut} plasmids

Plasmid pCetR expressing wild-type CetR was constructed as follow. The *cetR* gene with 81 nts upstream its TSS and 75 nts downstream its terminator was PCR-amplified from NEM316 chromosomal DNA using pJIM-*cetR*-F and pJIM-*cetR*-R primers. plasmid pJIM-*erm* (3) was PCR amplified using pJIM-F and pJIM-R primers. The two resulting PCR products were assembled by the Gibson assembly (2). To investigate the CetR and *dltD* mRNA interaction *in vivo*, pCetR derivatives with site directed deletion (pCetR_{Δ27}) and mutations (pCetR_{3mut}) in the plasmid *cetR* sequence were constructed. pCetR_{Δ27} was generated by the ligation (by Gibson assembly method) of a PCR product resulting for the PCR-amplification of pCetR using CetR_{Δ27}-upF and CetR_{Δ27}-upR primers. pCetR_{3mut} was generated in the same way, but using CetR_{3mut}-F and CetR_{3mut}-R primers. The CetR_{3mut} mutations are the following: A252G, A258G and T261C (see Fig. 3A). They were chosen to disrupt CetR-*dltD* interaction with a

minimum change in the overall RNA structures and to be complementary of three *dltD* mutations designed not to alter the *dltD* coding sequence.

RNA Preparation

For RNA preparation, overnight cultures of *S. agalactiae* strains grown in TH medium were diluted 100-fold in the same medium and grown at 37°C. Samples were then harvested at different growth phases, mixed with 0.5 volumes stop mix [95% (vol/vol) EtOH/5% (vol/vol) acetone], and immediately stocked at -20°C. Cells pellets were mixed with resuspension buffer (0.5 (vol/vol) 20% glucose, 0.5 (vol/vol) 25 mM Tris pH 7.6 and 10 mM EDTA), mechanically lysed with glass beads in FastPrep®-24. Total RNA was extracted using the phenol/trizol method as described (4). The quality of total RNA was checked on agarose and bioanalyser (Agilent).

Quantitative Real-Time PCR Analysis

RNA samples were treated with TURBO DNase (Ambion) and absence of contaminating genomic DNA was checked by PCR (for primers see Table S5). Gene selections for normalization, qRT-PCR experiments and data analysis were performed as described (5). The geometric mean of three genes (*gyrA*, *recA* and *rpoB*) chosen among ten genes was used to normalize the samples. Presented data were obtained from duplicate measures of three independent RNA sample extractions.

Western Blot Analysis

Overnight cultures of *Streptococcus* strains were diluted 100-fold in fresh TH media and grown at 37°C. Culture samples were then taken during exponential growth phase centrifuged for 8 min at 1, 250 x g at 4°C. Pellets were resuspended in lysis buffer (10 mM Tris-HCl, 1 mM EDTA, 150 mM NaCl, and 1 mg/ml lysozyme) to a final concentration of 20 OD units and incubated 30 min at 37°C. Lysed cells were mixed with the same volume of 2x SDS gel-loading buffer (100 mM Tris-HCl pH 6.8, 4 % SDS, 0.2 % bromophenol blue, 20 % glycerol, and 1 mM DTT) and denatured for 5 min at 95 °C. Western blot analysis of DltD::3xFLAG fusion proteins were performed using Life Technology western blotting devices, according to manufacturer's instructions. Novex™ WedgeWell™ Tris-Glycine gels were used for protein electrophoresis, iBlot® 2 Dry Blotting System for protein transfer, and the iBind™ Flex Western Device for antibodies hybridization and protein detection. Monoclonal anti-FLAG

(SIGMA) was used at 1:2,000 dilutions. Goat anti-rabbit secondary antibody (ImmunoReagents) was used at 1:2,000 dilutions.

Macrophage Survival Assay

Bacteria and human macrophages RAW 264.7 cells were prepared separately for survival assays. Bacteria were grown in TH medium at 37 °C and resuspended in RPMI 1640 medium (Gibco) to 10^6 CFU/ml. Macrophage cells were grown to confluence in RPMI 1640 supplemented with 10 % of Fetal Calf Serum (Gibco). Macrophages were then incubated anaerobically with bacteria at 10 MOI (multiplicity of infection) in RPMI 1640 at 37 °C for 1 hour to allow bacterial phagocytosis. The supernatants were recovered and infected macrophages were washed with RPMI (Gibco) and lysed by incubation with pure ice-cold water for 30 min. Total bacterial CFUs were determined by serial dilutions and compared to the starting inoculum. Survival assays were performed in duplicates and repeated at least three times.

Sequence Homology Searches

Sequence homology searches were performed with Blastn (<http://blast.ncbi.nlm.nih.gov>) implemented in Geneious R9.1.5 software (Biomatters LTD) using default parameters. Clonal complexes (CCs) were attributed to *cetR* sequences according to Patron *et al.* (6) and aligned using ClustalW using default parameters (7). Single nucleotide polymorphisms (SNPs) were determined using *S. agalactiae* NEM316 *cetR* sequence as reference.

Animal Experiments

Three weeks-old Swiss female mice were infected intravenously with 3×10^6 CFU of exponentially growing bacteria. One and three days after infection, mice were sacrificed; blood, brain, and liver were collected. Serial dilutions of blood were immediately plated on solid TH medium for CFU counts. Organs were homogenized in 0.9% NaCl using a tissue homogenizer, and serial dilutions were plated on solid TH medium.

All animal experiments were conducted in accordance with guidelines of Paris Descartes University, in compliance with the European animal welfare regulation (http://ec.europa.eu/environment/chemicals/lab_animals/home_en.html) and were approved by the Institut Cochin University Paris Descartes animal care and use committee (n° 12–145).

References

1. Biswas I, Gruss A, Ehrlich SD, & Maguin E (1993) High-efficiency gene inactivation and replacement system for gram-positive bacteria. *J Bacteriol* 175(11):3628-3635.
2. Gibson DG, *et al.* (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6(5):343-345.
3. Renault P, Corthier G, Goupil N, Delorme C, & Ehrlich SD (1996) Plasmid vectors for gram-positive bacteria switching from high to low copy number. *Gene* 183(1-2):175-182.
4. Lamy MC, *et al.* (2004) CovS/CovR of group B *streptococcus*: a two-component global regulatory system involved in virulence. *Mol Microbiol* 54(5):1250-1268.
5. Bury-Mone S, *et al.* (2009) Global analysis of extracytoplasmic stress signaling in *Escherichia coli*. *PLoS genetics* 5(9):e1000651.
6. Patron K, Gilot P, Camiade E, & Mereghetti L (2015) An homolog of the Frz Phosphoenolpyruvate:carbohydrate phosphoTransferase System of extraintestinal pathogenic *Escherichia coli* is encoded on a genomic island in specific lineages of *Streptococcus agalactiae*. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* 32:44-50.
7. Thompson JD, Gibson TJ, & Higgins DG (2002) Multiple sequence alignment using ClustalW and ClustalX. *Current protocols in bioinformatics / editorial board, Andreas D. Baxevanis ... [et al.]* Chapter 2:Unit 2 3.

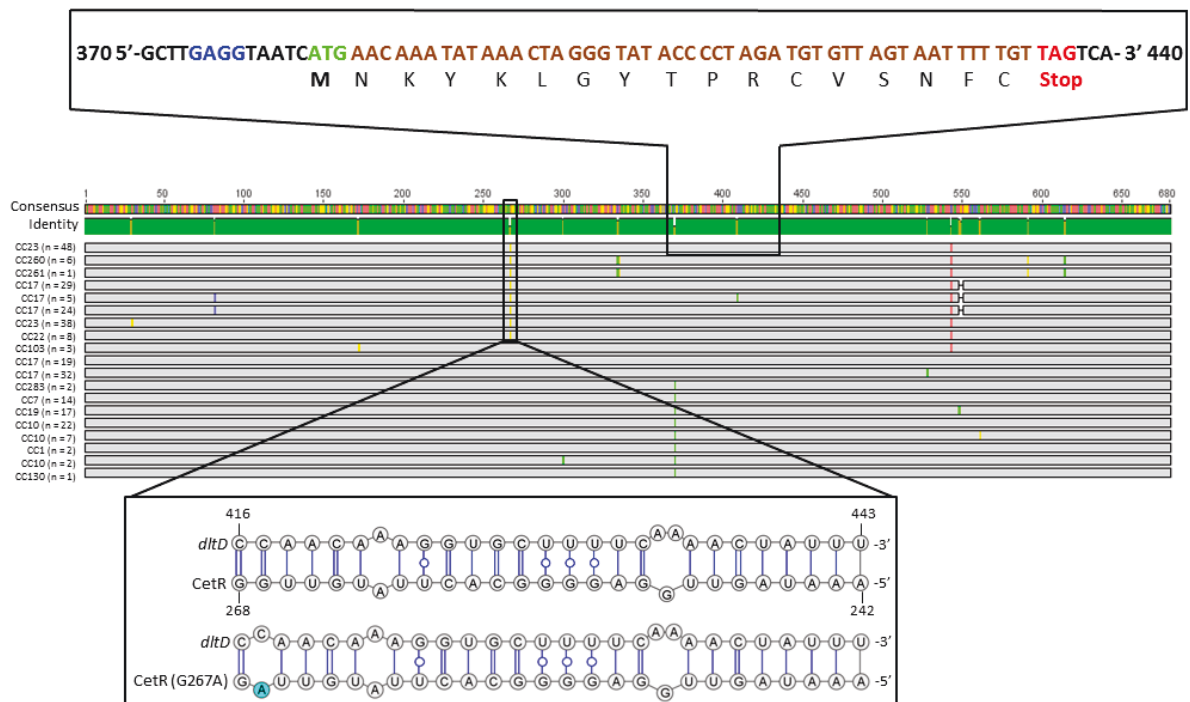


Fig. S1. Genomic context of *cetR* from different *Streptococcus* strains. Multiple sequence alignment of 280 *cetR* sequences distributed in twelve *S. agalactiae* clonal complexes (CCs). The single nucleotide mismatch (G267A, colored in blue) localized in CetR-*dltD* pairing region is squared. The predicted peptide on *cetR* sequence and the stop codon are indicated. Green bar indicated the relative identity. Colored box on the alignment indicate single nucleotide polymorphism (SNP) on the aligned sequences. In brackets is indicated the number of aligned strains for the same CC.

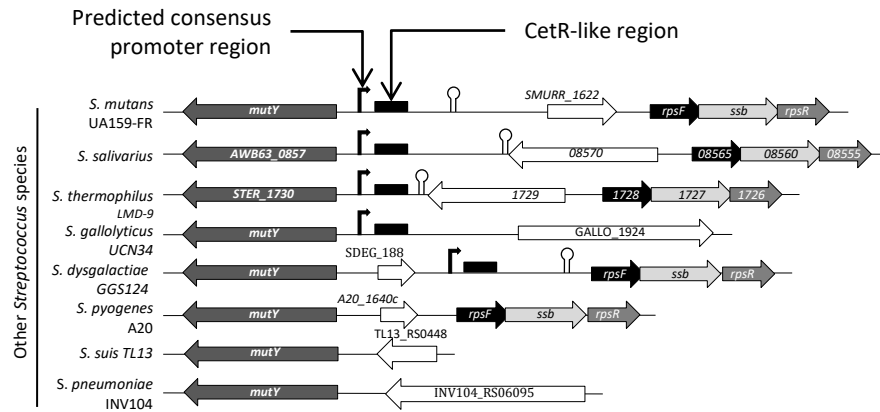


Fig. S2: Multiple alignment of CetR-like RNA sequences. The multiple alignment, using ClustalW program, revealed the presence of a poorly conserved region of *cetR*-like sequence, a 150 nts downstream a potentially conserved promoter region, in related *Streptococcus* species. Homologous genes in different *Streptococcus* species are illustrated by the same colors, whereas unrelated genes are indicated in white.

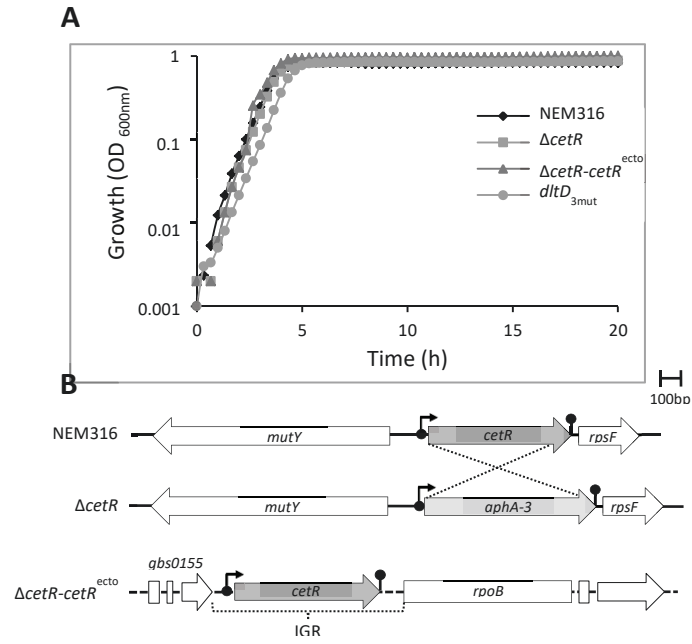


Fig. S3. (A) Growth curves of NEM316, $\Delta cetR$, $\Delta cetR\text{-}cetR^{\text{ecto}}$ and $dltD_{3\text{mut}}$ cells in TH medium as determined by bioscreen. All growth kinetics were repeated at least three times. (B) Schematic representation of the construction of $\Delta cetR$ mutant and its ectopic complemented strain $\Delta cetR\text{-}cetR^{\text{ecto}}$ (see SI Methods for details).

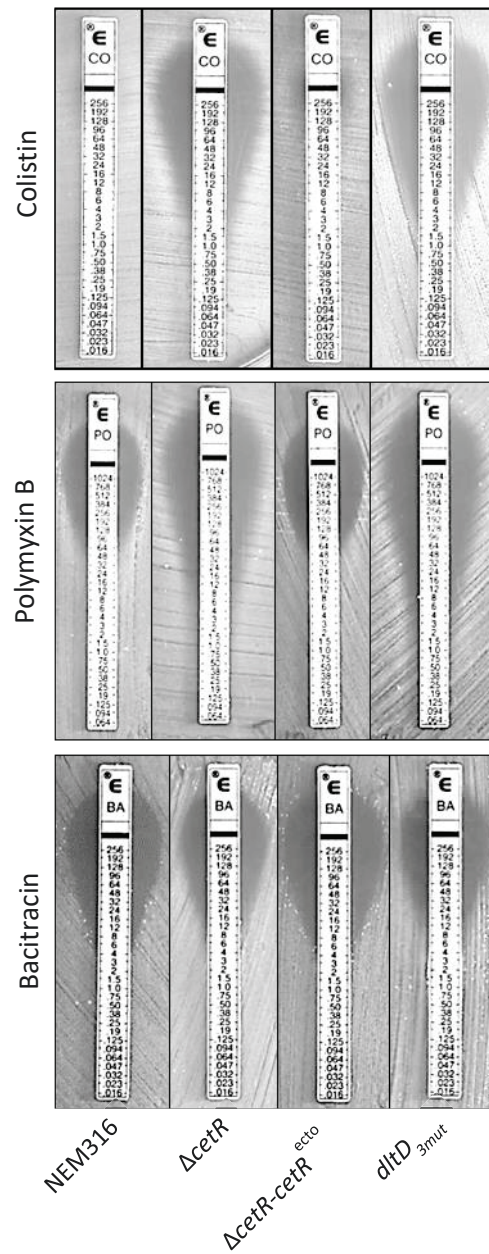


Fig. S4. MICs measured using antibiotic E-test for NEM316, $\Delta cetR$, $\Delta cetR-cetR^{ecto}$ and $dltD_{3mut}$ cells. MIC is determined by the intersection between the E-test and the bottom of the bacterial inhibition zone. Colistin and Polymyxin are hydrophobic AMPs. Bacitracin was used as negative control.

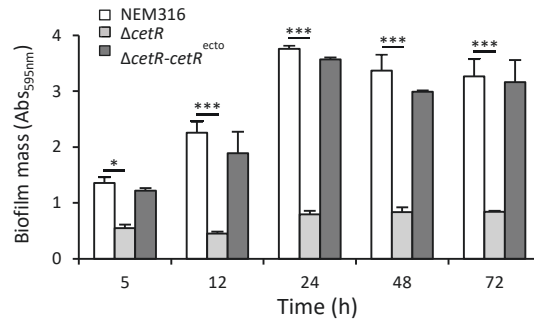


Fig. S5. Biofilm formation was assayed in NEM316, $\Delta cetR$ and $\Delta cetR-cetR^{ecto}$ strains at different time points. Bacterial sessile biomass was measured with crystal violet coloration.

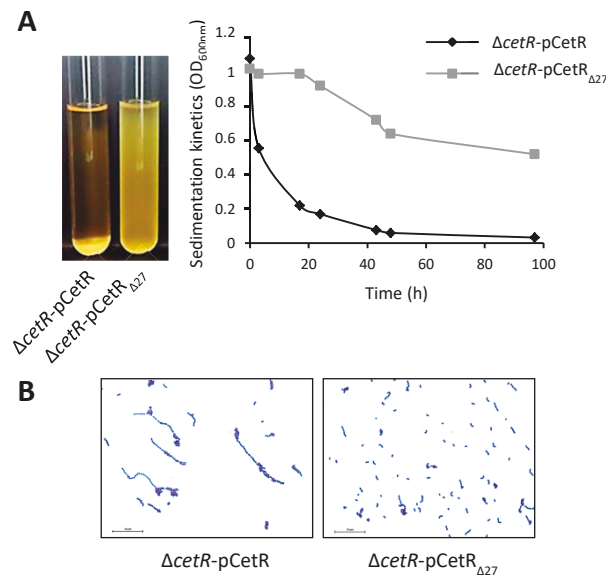


Fig. S6. CetR-*dltD* interaction region is crucial for efficient cell morphology formation. (A) Bacterial sedimentation was observed on standing tubes of liquid cultures (left panel), and kinetics were monitored as described in methods (right panel). $\Delta cetR$ cells were transformed with plasmids carrying wild-type or mutated CetR (pCetR $_{\Delta 27}$, deletion of the entire interaction region with *dltD*). (B) $\Delta cetR$ cells, transformed with mutated RNA pCetR $_{3mut}$, present significantly reduced chain lengths compared to the NEM316 background, as shown with Gram-colored microscopic observations.

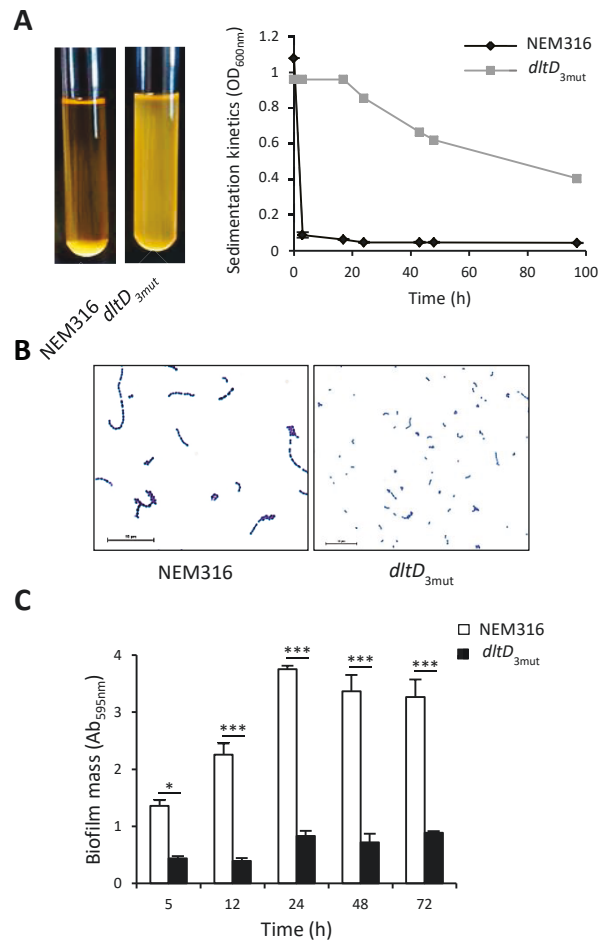


Fig. S7. *dltD*_{3mut} Promotes Cell Morphology Changes. (A) Bacterial sedimentation was observed on standing tubes of liquid cultures (left panel), and kinetics were monitored as described in methods (right panel). (B) *dltD*_{3mut} cells present significantly reduced chain lengths compared to the WT background, as shown with Gram-colored microscopic observations. (C) Biofilm formation was assayed in WT and *dltD*_{3mut} strains at different time points. Bacterial sessile biomass was measured with crystal violet coloration.

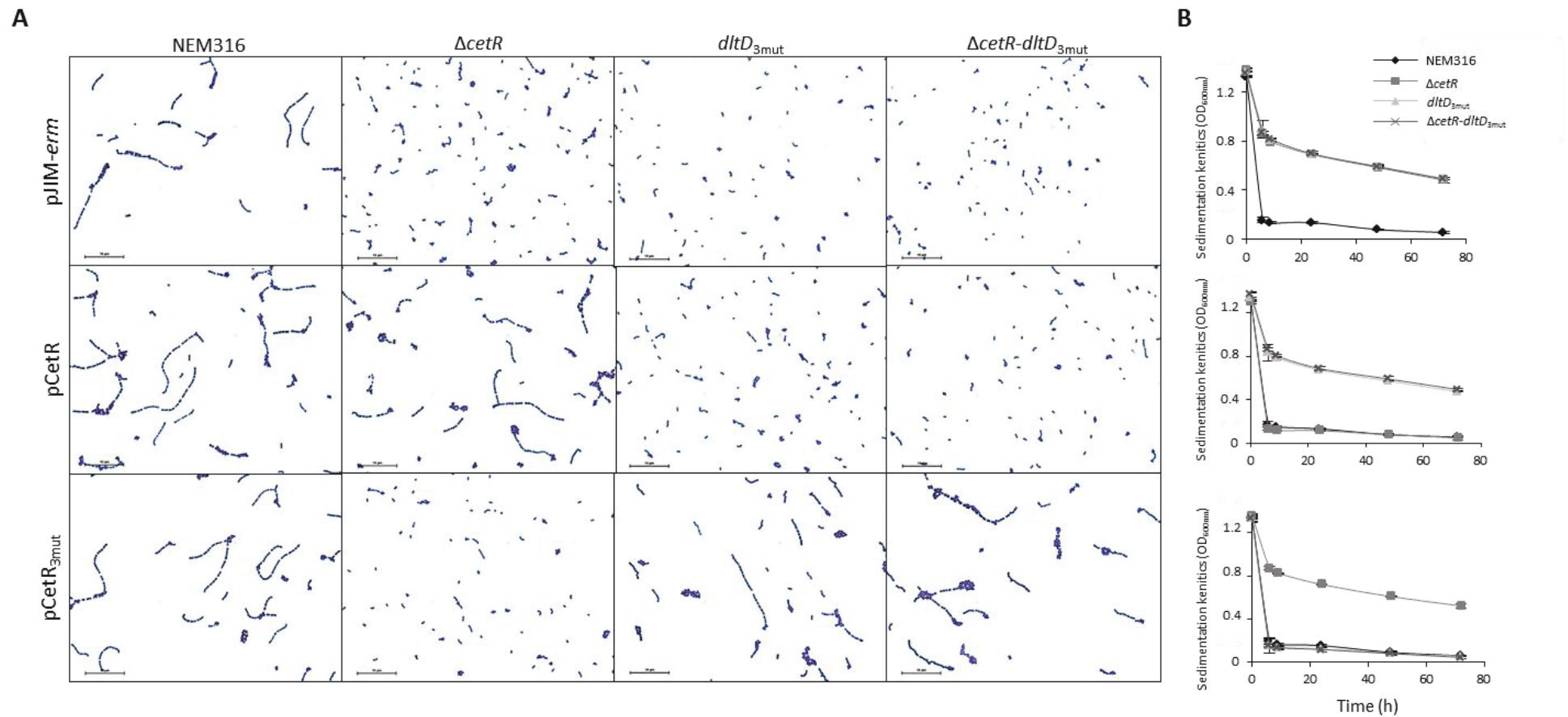


Fig. S8. (A) Gram-colored microscopic observation of NEM316, $\Delta cetR$, $dltD_{3mut}$ and $\Delta cetR-dltD_{3mut}$ cells transformed with empty plasmid (pJIM-*erm*), expressing wild-type RNA (pCetR) or mutated CetR (pCetR_{3mut}). (B) Bacterial sedimentation kinetics monitored in the same conditions.

Table S1: Minimum inhibitory concentrations (MICs)*

Antibiotic target (μg/ml)	WT	Δ <i>cetR</i>	<i>cetR</i> ^{ecto}	<i>dltD</i> _{3mut}
Bacterial envelope				
β-lactams				
Penicillin	0.06	0.06	0.06	0.06
Ampicillin	0.06	0.06	0.06	0.06
Cefotaxime	0.01	0.01	0.01	0.01
Glycopeptides				
Vancomycin	0.50	0.50	0.50	0.50
Teicoplanin	0.06	0.06	0.06	0.06
Nucleic acids synthesis				
Norfloxacin	8.00	8.00	8.00	8.00
Levofloxacin	1.00	1.00	1.00	1.00
Rifampicin	0.01	0.01	0.01	0.01
Proteins synthesis				
Erythromycin	0.25	0.25	0.25	0.25
Lincomycin	0.25	0.25	0.25	0.25
Pristinamycin	0.25	0.25	0.25	0.25
Gentamicin	64.00	64.00	64.00	64.00
Tetracycline	0.25	0.25	0.25	0.25
Cell membrane				
Bacitracin	4.00	4.00	4.00	4.00
Daptomycin	1.00	1.00	1.00	1.00

*tests were repeated three times.

Table S2: Colistin Minimum inhibitory concentrations (MICs)*

	WT [#]	Δ <i>cetR</i> [#]	<i>dltD</i> _{3mut} [#]	Δ <i>cetR-dltD</i> _{3mut} [#]
pJIM-erm [§]	>512	2	2	2
pCetR [§]	>512	>512	2	2
pCetR _{3mut} [§]	>512	2	>512	>512

*tests were repeated at three times

[#] Relevant strain genotypes

[§] plasmids from the corresponding strains.

Table S3: Bacterial strains

Common Name	Collection Name	Genotype	Source or reference
<i>Streptococcus agalactiae</i>			
WT	NEM316	Wild-type (NCBI Acc-no NC_004368) Serotype III isolated from neonatal blood culture	Glaser et al, 2002
$\Delta cetR$	MFL100	NEM316 $\Delta cetR::Kan$, Km ^R	This study
$cetR^{ecto}$	MFL101	NEM316 $\Delta cetR$ IGR ^{<i>gbs0155-rpoB</i>} :: <i>cetR</i> , Km ^R , ectopic complementation of <i>cetR</i> in the intergenic region (<i>gbs0155-rpoB</i>)	This study
<i>dltD</i> _{3mut}	MFL102	NEM316 <i>dltD</i> _{3mut} , compensatory mutations (A423G, T426C, T432C) in the coding sequence of <i>dltD</i> in the interaction region (CetR- <i>dltD</i>)	This study
$\Delta cetR$ - <i>dltD</i> _{3mut}	MFL103	NEM316 $\Delta cetR$ - <i>dltD</i> _{mut} (A423G, T426C, T432C), Km ^R	This study
WT- <i>dltD</i> ::3xFLAG	MFL109	NEM316 <i>DltD</i> ::3xFLAG	This study
$\Delta cetR$ - <i>dtD</i> ::3xFLAG	MFL110	NEM316 $\Delta cetR::Kan$ <i>DltD</i> ::3xFLAG, Km ^R	This study
$\Delta cetR$ - $cetR^{ecto}$ - <i>dltD</i> ::3xFLAG	MFL111	NEM316 $\Delta cetR$ IGR ^{<i>gbs0155-rpoB</i>} :: <i>cetR</i> <i>DltD</i> ::3xFLAG, Km ^R	This study
<i>dltD</i> _{3mut} - <i>dltD</i> ::3xFLAG	MFL112	NEM316 <i>dltD</i> _{3mut} <i>DltD</i> ::3xFLAG	This study
<i>Escherichia coli</i>			
DH5 α	BRMF080	F ⁻ Φ 80lacZ Δ M15 Δ (lacZYA-argF) U169 recA1 endA1 hsdR17 (rK ⁻ , mK ⁺) phoA supE44 λ ⁻ thi-1 gyrA96 relA1	Invitrogen

Table S4: Plasmids

Name	Description	Source or reference
pG+host5	ColE1 replicon, thermosensitive derivative of pGK12; MCS pBluescript, Em ^R	Biswas et al. (1993)
pΔcetR::aphA-3	Em ^R , Kan ^R , derivative of pG+host5 used for <i>cetR</i> deletion by allelic exchange with aphA-3	This study
pBT2	Kan ^R , plasmid carrying aphA-3	Laboratory collection
pCetR ^{ecto}	Em ^R , derivative of pG+host5 used for <i>cetR</i> ectopic complementation	This study
pDltD _{3mut}	Em ^R , derivative of pG+host5 used to generate three point mutations (A423G, T426C and T432C) in <i>dltD</i> chromosomal gene	This study
pDltD3xFlag	Em ^R , derivative of pG+host5 carrying the translational fusion DltD::3xFlag	This study
pDltD _{3mut} 3xFlag	Em ^R , derivative of pG+host5 carrying the translational fusion DltD _{3mut} ::3xFlag	This study
pJIM- <i>erm</i>	Em ^R , derivative of the shuttle vector pJIM2246	This study
pCetR	pJIM- <i>erm-cetR</i> , Em ^R , cloning of <i>cetR</i> in pBRM100 under the control of its native promoter	This study
pCetR _{Δ27}	pJIM- <i>erm-cetR</i> _{Δ27} , Em ^R , cloning of <i>cetR</i> , This study deleted of the entire interaction region (CetR- <i>dltD</i> mRNA, i.e. 27 nts), in pBRM100 under the control of its native promoter	
pCetR _{3mut}	pJIM- <i>erm-cetR</i> _{3mut} , Em ^R , cloning of <i>cetR</i> , This study carrying the mutations (A252G, A258G, T261C) in <i>cetR</i> sequence in the interaction region (CetR- <i>dltD</i> mRNA), in pBRM100 under the control of its native promoter	
pTOOX	Plasmid carrying 3xFLAG	Laboratory collection

Table S5: DNA oligonucleotides

Name	Primers (sequence 5' => 3')	Purpose
<i>cetR</i> deletion oligos		
<i>Srn073</i> -UpF	CGACGGTATCGATAAGCTTGATATCGAATTCCTGCAGCCCCTGCATCTTCAGCTTCAACG	Forward primer for <i>cetR</i> upstream region and insertion in pG+host5
<i>Srn073</i> -UpR	CCATTATTATTTCTTCCTCTTTTCTACAGTATTTAAAGTTTTATTTTAAATTGAATTATAGAA TGGG	Reverse primer for <i>cetR</i> upstream region
<i>Srn073</i> -downF	GAAAATAAAATATTATATTTTACTGGATGAATTGTTTTAGAATTTAATTATAGTGTAACGCTTC CTTA	Forward primer for <i>cetR</i> downstream region
<i>Srn073</i> -downR	TGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAATGCCACCAAAATCAACCAT	Reverse primer for <i>cetR</i> downstream region and insertion in pG+host5
<i>kan</i> -F	CTTTAAATACTGTAGAAAAGAGGAAGG	Forward primer for kanamycin resistance cassette
<i>kan</i> -R	CTAAAACAATTCATCCAGTAAAATA	Reverse primer for kanamycin resistance cassette
pGhost-F	TATCCGCTCACAATTCCACA	Forward primer for pG+host5 amplification
pGhost-R	GGGCTGCAGGAATTCGATA	Reverse primer for pG+host5 amplification
pGhMCS-F	ATGGCGCATTACGGAATAA	Forward primer for pG+host5:: <i>cetR</i> assembly control
pGhMCS-R	TGTGGAATTGTGAGCGGATA	Reverse primer for pG+host5:: <i>cetR</i> assembly control
<i>Dsrn073</i> -F	TGCCTTATATCAAACCTATGTAGCC	Forward primer for <i>cetR</i> deletion control
<i>Dsrn073</i> -R	CAAAGAAAGAGGATTTAACAAAGAGG	Reverse primer for <i>cetR</i> deletion control

***cetR* ectopic chromosomal insertion oligos**

IGR-UpF	CGACGGTATCGATAAGCTTGATATCGAATTCCTGCAGCCCATCACGAAAGTAAACCTGTCCG	Forward primer for IGR upstream region
IGR-UpR	CTTCATCATTAAATTACGATCTTTTTAAACATAGCATACTGGTG	Reverse primer for IGR upstream region
IGR-DownF	AAATGGCTACATAGGTTTGATCATCGTGTTATGCTAAAC	Forward primer for IGR downstream region
IGR-DownR	TGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATATAAACACCAGGAGAACGGAC	Reverse primer for IGR downstream region
<i>cetR</i> -ectoF	CAGTATGCTATGTTTAAAAAGATCGTAATTAATGATGAAGAAAAATG	Forward primer for <i>cetR</i> ectopic insertion
<i>cetR</i> -ectoR	GTTTATAGCATAACACGATGATCAAACCTATGTAGCCATTTTC	Reverse primer for <i>cetR</i> ectopic insertion
IGR-OutF	TGGACTTACAAGATGCTACG	Forward primer for <i>cetR</i> insertion control
IGR-OutR	ACGACGTGTACGGTGTTTC	Reverse primer for <i>cetR</i> insertion control

Insertion of a triple mutations into *dltD* (*dltD*_{3mut}) oligos

<i>dltD</i> _{3mut} -UpF	GATAAGCTTGATATCGAATTCCTGCAGCCCTGGCATATGAGCCTTCTTTCTGG	Forward primer for <i>dltD</i> upstream region
<i>dltD</i> _{3mut} -UpR	AATAGTTTTGGAAAGCGCCCTTGTTGGCTCCCT	Reverse primer inserting a triple mutation to <i>dltD</i>
<i>dltD</i> _{3mut} -DownF	GAGCCAACAAGGGCGCTTCCAAAACATTTTCAGCAAC	Forward primer inserting a triple mutation to <i>dltD</i>
<i>dltD</i> _{3mut} -DownR	TCGTATGTTGTGTGGAATTGTGAGCGGATATGAAACCAAGCCCTAAAAG	Reverse primer for <i>dltD</i> downstream region

pCetR (pBRM101) assembly oligos

pJIM- <i>cetR</i> -F	TAAATAGAATTCGGGCCCCCGATCGTAATTAATGATGAAGAAAAATG	Forward primer for <i>cetR</i> fragment
pJIM- <i>cetR</i> -R	AGAACTAGTGGATCCCCGAAACCTATGTAGCCATTTTCT	Reverse primer for <i>cetR</i> fragment
pJIM-F	CGGGGGATCCACTAGTTCT	Forward primer for pJIM2246 amplification

pJIM-R GGGGGCCCGAATTCTATTTA

Reverse primer for pJIM2246 amplification

pCetR_{Δ27} (pBRM115) assembly oligos

CetR_{Δ27}-upF taaatagaattcgggccccGATCGTAATTAATGATGAAGAAAAATG

Forward primer for *cetR* up fragment

CetR_{Δ27}-upR acatctaaacTAGGAAATGAAGAAAACTTCAAATC

Reverse primer for *cetR* up fragment

CetR_{Δ27}-downF tcatttcctaGTTTAGATGTAGGTCGCTAC

Forward primer for *cetR* down fragment

CetR_{Δ27}-downR tagaactagtggatccccgTCAAACCTATGTAGCCATTTTC

Reverse primer for *cetR* down fragment

pCetR_{3mut} (pBRM102) assembly oligos

CetR_{3mut}-F TTGGGGGGGCGCTCATGTTGGGTTTAGATGTAGGTCGCTACA

Forward primer inserting a triple mutation to *cetR*

CetR_{3mut}-R CAACATGAGCGCCCCCAACTATTTCTAGGAAATGAAGAAAACTTC

Reverse primer inserting a triple mutation to *cetR*

DltD::3xFLAG assembly oligos

gbs1829 F atatcgaattcctgcagcccGGTGCTGTTGACCCCTTAAA

Forward primer for *gbs1829* and insertion in pG+host5

gbs1829 R acaaggatgacgatgacaagTAAATAGTAGAGTCAGATGATTATAGC

Reverse primer for *gbs1829* and link with 3xFLAG

dltD F ccgtcatggtcttttagtcCTTGTAATCACTTGGACGTC

Forward primer for *dltD* and link with 3xFLAG

dltD R tgtggaattgtgagcgataTTTCTCCGCAATGGTTTGT

Reverse primer for *dltD* and insertion in pG+host5

3xFLAG F TTTATCACAGCAGGTATGGG

Forward primer for *3xflag* fragment

3xFLAG R GTGATAACCGCTACTGTAAG

Reverse primer for *3xflag* fragment

qRT-PCR oligos

qRT- <i>cetR</i> -F	GCATGCGCTAATGGTAATGA	Forward primer for <i>cetR</i> fragment
qRT- <i>cetR</i> -R	TGTTGCAAACAGTCACAAGGA	Reverse primer for <i>cetR</i> fragment
qRT- <i>dltD</i> -F	GGCGGCTTCTCTGACACAAT	Forward primer for <i>dltD</i> fragment
qRT- <i>dltD</i> -R	CACCTTTGTTGGCTCCCTTG	Reverse primer for <i>dltD</i> fragment
qRT- <i>gyrA</i> -F	TGCCGCACAGCGTTATACAG	Forward primer for <i>gyrA</i> (<i>gbs0948</i>) fragment
qRT- <i>gyrA</i> -R	CGCGCTGGTAAAACAAGAGG	Reverse primer for <i>gyrA</i> (<i>gbs0948</i>) fragment
qRT- <i>recA</i> -F	TCGCTTTAGGAGCAGGTGGA	Forward primer for <i>recA</i> (<i>gbs2047</i>) fragment
qRT- <i>recA</i> -R	GCAGCATAGGCTGGGTCAAG	Reverse primer for <i>recA</i> (<i>gbs2047</i>) fragment
qRT- <i>rpoB</i> -F	TCAGGTGCAGCTGTGATTGC	Forward primer for <i>rpoB</i> (<i>gbs0156</i>) fragment
qRT- <i>rpoB</i> -R	TTTGAACGGCGGAATTTTG	Reverse primer for <i>rpoB</i> (<i>gbs0156</i>) fragment
qRT- <i>gyrB</i> -F	GCTCACATCAGAACTTTACTTTTAACTC	Forward primer for <i>gyrB</i> fragment
qRT- <i>gyrB</i> -R	TTTAATCTCACTTCCTACTTTGACACC	Reverse primer for <i>gyrB</i> fragment

RESULTATS COMPLEMENTAIRES

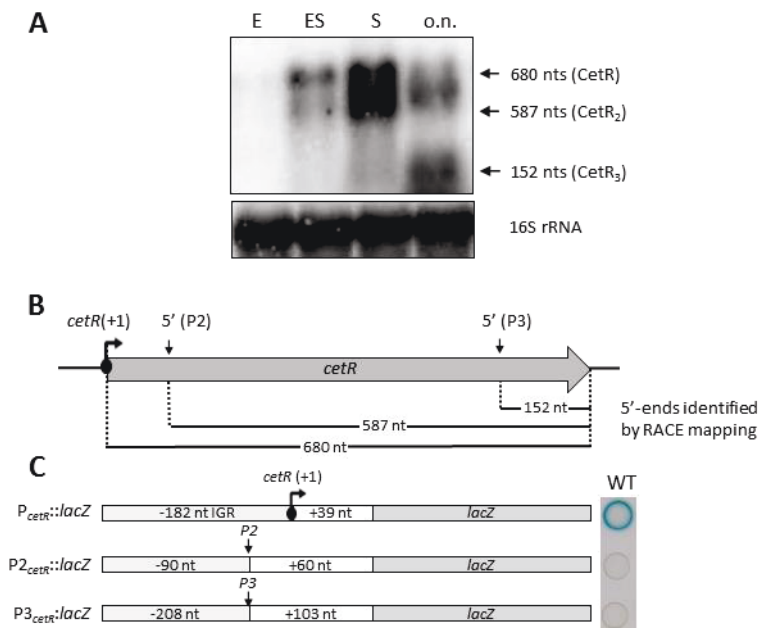


Figure 34 : Profil d'expression de CetR.

(A) Profil northern blot des transcrits de CetR. Les différentes bandes correspondant aux trois transcrits sont indiquées par les flèches. L'ARNr 16S a été utilisé comme contrôle. E, Exponential; ES, Early Stationary; S, Stationary; o.n., over-night. (B) Position et taille des transcrits de CetR. (C) Représentation schématique des fusions transcriptionnelles P_{cetR} -, $P2_{cetR}$ - et $P3_{cetR}::lacZ$ ainsi que l'activité β -galactosidase chez la souche sauvage (WT).

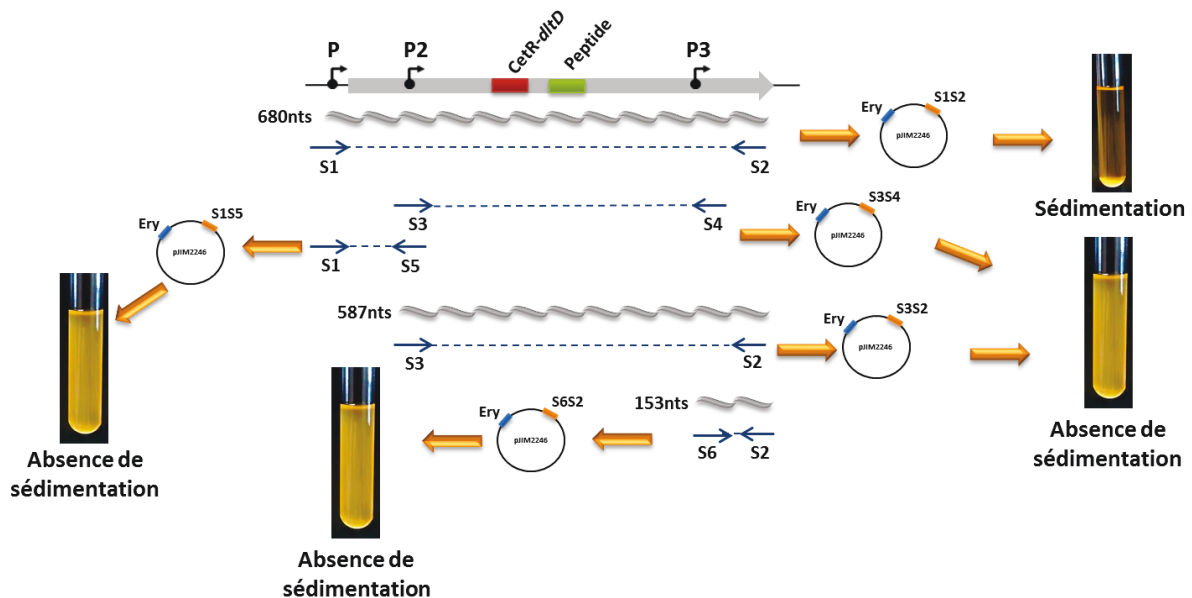


Figure 35 : Caractérisation de la région d'interaction CetR-*dltD*.

Diagramme représentant les différents transcrits de CetR (680, 587 et 153nts). Cinq plasmides, dérivés du plasmide pJIM2246, ont été construits avec les couples d'amorces S1S2, S3S4, S1S5, S3S2 et S6S2 puis transformés dans le mutant *cetR*. L'effet de chaque plasmide a été ensuite évalué par observation de la sédimentation.

E. Résultats complémentaires

Dans cette partie, nous exposerons des résultats complémentaires du travail précédent, qui seront publiés dans un deuxième article.

1. CetR, code-t-il différents transcrits fonctionnels ?

Le profil d'expression de CetR présenté en Figure 34. A corrèle avec celui publié précédemment (Rosinski-Chupin *et al.*, 2015) et montre une expression dépendante de la phase de croissance (Figure 30. C). Cependant, en plus du transcrit de taille 680 nts, deux autres transcrits plus courts ont été détectés (Figure 34. B). Pour tester l'existence d'éventuels promoteurs alternatifs, l'extrémité(s) 5' de l'ARNm *cetR* a été mappé par 5' RACE (5' rapid amplification of cDNA ends, voir annexe 1 pour les méthodes) et les régions de promoteurs putatifs ont été clonées avec le gène rapporteur *lacZ* (Figure 34. C ; voir annexe 1 pour les détails de la construction des fusions transcriptionnelles). Une activité β -galactosidase a été détectée uniquement avec la fusion transcriptionnelle $P_{cetR}::lacZ$, correspondant au transcrit de 680 nts (Figure 34. C). L'absence d'activité des fusions $P2_{cetR}::lacZ$ et $P3_{cetR}::lacZ$ suggère que CetR est un long transcrit unique et que les autres fragments sont des produits de traitement « *processing* » par des RNases.

2. Caractérisation des différents transcrits de CetR

Le profil northern blot de CetR (Figure 34. A) montre que l'ARNrég possède trois différents transcrits ; et l'analyse de l'activité β -galactosidase montre que seul le promoteur du grand transcrit (680 nts) est actif. Cette différence de profil avec le northern blot présenté dans l'article (Rosinski-Chupin *et al.*, 2015) est due à la différence de position des sondes utilisées pour l'hybridation de CetR.

Afin de vérifier si les deux autres transcrits (CetR₂ et CetR₃) jouent un rôle dans les phénotypes observés, nous avons cloné leurs séquences correspondantes sous le contrôle de P_{cetR} dans le plasmide pJIM2246. Les plasmides recombinants ont été ensuite transformés dans la souche $\Delta cetR$ et la sédimentation a été évaluée (Figure 35).

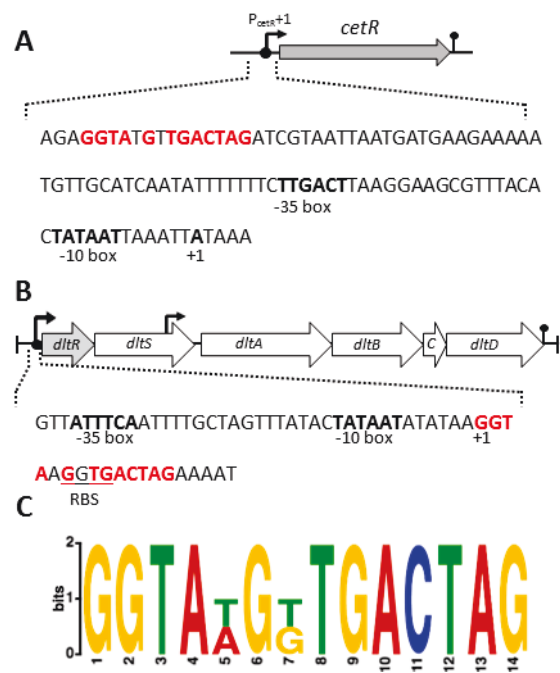


Figure 36 : Région homologue entre les promoteurs de *dltR* et *cetR*.

(A) Région promotrice de *cetR*. (B) Région promotrice de *dltR*. (C) Région homologue entre les promoteurs de *dltR* et *cetR*.

Nos résultats montrent que seul le grand transcrit (680nts) permet de restaurer le phénotype de la souche sauvage (Figure 35). Ils confirment que les transcrits CetR₂ et CetR₃ n'ont pas de rôle dans les phénotypes observés. Avec la même stratégie, nous avons construit des plasmides contenant +/- 50 nts de : 1) la région d'interaction CetR-*dltD*, 2) la séquence codant le peptide putatif et 3) les deux (région d'interaction et peptide). La sédimentation du mutant *cetR* a été ensuite évaluée après transformation avec les différents plasmides. Aucun de ces plasmides n'a permis de restaurer la sédimentation du mutant *cetR* au niveau de la souche sauvage. Ce résultat confirme que : 1) l'intégralité de l'ARNrég CetR est nécessaire pour interagir avec l'ARNm *dltD*, 2) le peptide putatif, même s'il est exprimé, n'influe pas sur la fonction régulatrice de l'ARNrég CetR.

3. DltR, le régulateur du régulateur ?

L'analyse comparative des régions promotrices de *cetR* et du gène codant le régulateur DltR (*dltR*), a montré la présence d'une région homologue entre les deux séquences (Figure 36). Poyart et collaborateurs ont montré que le régulateur DltR régule positivement l'expression de son propre promoteur (Poyart *et al.*, 2001). Cependant, le mécanisme exact impliqué dans cette auto-régulation n'a pas été déterminé. La fixation de DltR sur son propre promoteur pourrait effectivement activer la transcription de l'opéron *dltABCD*. De la même façon, DltR pourrait activer directement la transcription de CetR, en réponse à la présence des PAM, qui activerait à son tour la traduction de DltD. La région d'homologie, entre les régions promotrices des gènes *cetR* et *dltR*, pourrait être impliquée dans cette régulation et sa caractérisation pourrait apporter une meilleure compréhension de la régulation de DltD par CetR.

4. L'expression de CetR est inhibée par le stress acide et activée par le stress salin

Nous avons montré que l'expression de CetR est cruciale pour la résistance aux PAM et la virulence de *S. agalactiae*. Ces résultats suggèrent fortement que CetR joue un rôle dans la réponse au changement dans l'environnement de la bactérie. Effectivement, nous avons montré que l'expression de CetR est légèrement augmentée en présence de 1M NaCl. Nous n'avons pas pu confirmer cette régulation par le dosage de l'activité β -galactosidase

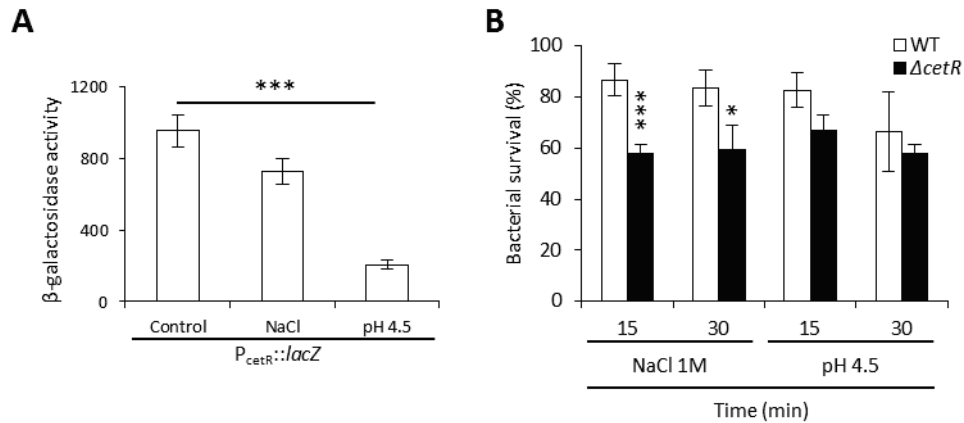


Figure 37 : Effet de stress acide et salin sur l'ARNrég CetR.

(A) Activité β -galactosidase de la fusion transcriptionnelle $P_{cetR}::lacZ$ dans la souche WT cultivée dans le milieu TH en phase exponentielle de croissance pendant 30min de NaCl 1M ou dans un pH 4,5. (B) Survie des souches WT ou $\Delta cetR$ à 15 ou 30 min de stress salin (NaCl 1M) ou acide (pH 4,5).

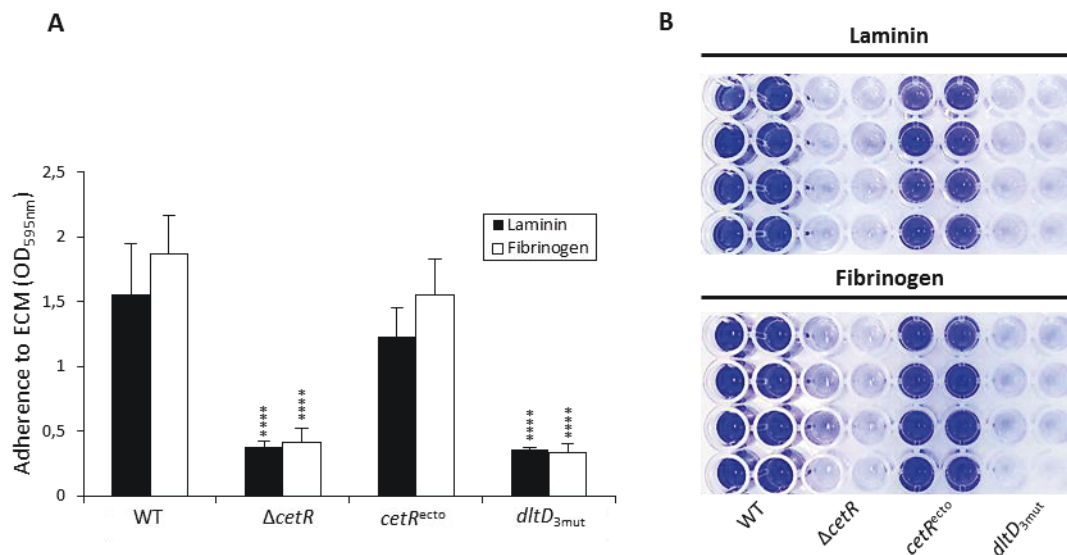


Figure 38 : CetR favorise l'adhérence aux protéines de la matrice extracellulaire.

(A) Adhérence des souches WT, $\Delta cetR$, *cetR*^{ecto} et *dltD*_{3mut} à la laminine et au fibrinogène. (B) La biomasse adhérente est estimée par coloration au Crystal violet, dont la lecture d'absorbance se fait à 595 nm.

dans cette condition ; cependant, nous avons montré que l'activité de la fusion $P_{cetR}::lacZ$ est significativement réprimée en pH acide par rapport au contrôle (Figure 37. A). Ce résultat corrèle avec les résultats du RNA-seq (Rosinski-Chupin *et al.*, 2015). L'inhibition de l'expression de l'ARNrég CetR en pH acide, valeur habituelle du pH des cellules phagocytaires, pourrait être une stratégie de *S. agalactiae* pour échapper au système immunitaire.

L'étude de la survie bactérienne en conditions de stress (salin et acide) a montré que le mutant *cetR* est plus sensible au stress salin que la souche sauvage (Figure 37. B). Ce résultat peut expliquer une baisse d'homéostasie membranaire dans le mutant *cetR*. La membrane cellulaire de la souche $\Delta cetR$ étant fragilisée, cette dernière devient plus sensible aux stress dans lesquels interviennent des molécules importantes pour l'osmose cellulaire, tels que les cations Na^+ ou les ions Cl^- . Aucune différence de survie bactérienne n'a été observée en stress acide (Figure 37. B).

5. CetR favorise l'adhérence aux protéines de la matrice extracellulaire

Afin d'envahir son hôte, *S. agalactiae* doit tout d'abord adhérer aux protéines de la matrice extracellulaire (ECM), qui composent la plupart des cellules eucaryotes. Parmi les composants majeurs de l'ECM, on distingue la laminine et le fibrinogène. Nous avons ainsi évalué la capacité d'adhérence des différentes souches de *S. agalactiae* à ces deux protéines. Afin de réaliser ce test, 200 μ L de la suspension bactérienne sont déposés dans chacun des puits des plaques de microtitration préalablement préparées. Des puits contenant seulement du milieu de culture stérile servent à contrôler l'absence de contamination. Après 3h d'incubation à 37 °C, la suspension bactérienne est retirée et la biomasse adhérente est mesurée par coloration au Crystal violet (*cf.* annexe 1 pour détail des méthodes).

La figure 38 montre clairement que les souches $\Delta cetR$ et $dltD_{3mut}$ adhèrent significativement moins aux deux composants de l'ECM (laminine et fibrinogène) par rapport aux souches WT et *cetR*^{ecto}. Cet effet sur l'adhérence des mutants $\Delta cetR$ et $dltD_{3mut}$ peut expliquer la perte de la capacité à former du biofilm par ces souches (*cf.* figure 33). Par ailleurs, ces expériences d'adhérence effectuées sur des protéines purifiées de l'ECM

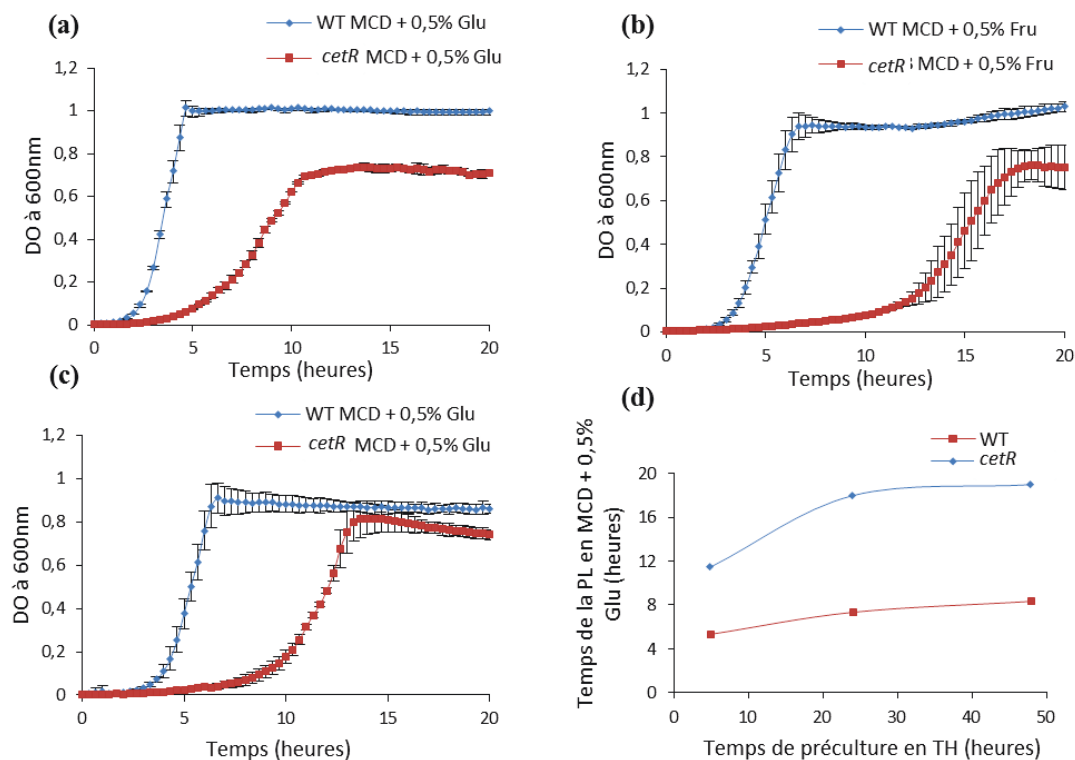


Figure 39 : Rôle de CetR dans la croissance en milieu chimiquement défini.

(a) et (b) Suivi de croissance des souches WT et du mutant *cetR* en MCD supplémenté avec 0,5% de glucose ou de fructose après une préculture en TH, respectivement. (c) Suivi de croissance en MCD supplémenté avec 0,5% de glucose après préculture en MCD supplémenté avec 0,5% de glucose. (d) Corrélation entre le temps de la phase de latence (PL) en MCD et le temps de préculture en TH.

devraient être complétées par l'étude d'autres composants de l'ECM mais aussi par l'étude de la capacité d'adhérence à des cellules épithéliales vaginales, voire intestinales dans le cas de colonisation ou d'infection chez l'Homme.

6. Rôle de CetR dans la croissance en milieu chimiquement défini

La croissance bactérienne des souches sauvages (wilde-type, WT) et du mutant *cetR* a été étudiée en milieu riche (TH), et en milieu chimiquement défini (MCD) supplémenté avec du glucose ou du fructose à différentes concentrations (0,1%, 0,25%, 0,5%, 1% ou 2,5%). Pour l'ensemble des expériences, la croissance bactérienne a été réalisée dans des plaques de 96 puits et suivi en bioanalyser. Dans le milieu TH, les deux souches (WT et *cetR*) présentent le même profil de croissance (*cf*, Figure S1, annexe 2). Cependant, après une préculture en TH, la croissance du mutant *cetR* en MCD, supplémentée avec 0,5 % de glucose, est nettement ralentie par rapport à la souche WT (Figure 39. a). Ce ralentissement de croissance est dû à un rallongement de la phase de latence, indépendamment de la source ou de la concentration du carbone utilisé (Figure 39. b croissance en fru et annexe 2 pour les différentes concentrations en Glu ou Fru). Une préculture préalable en MCD n'a pas permis d'inhiber le ralentissement de la croissance du mutant *cetR* en MCD supplémenté avec 0,5% de glucose (Figure 39. c). Ce résultat suggère que le retard de croissance du mutant n'est pas dû à des mutations spontanées lors de la croissance.

En outre, la corrélation entre le temps de la phase de latence (PL) en MCD et le temps de la préculture en TH montre que, plus le temps de préculture en TH est important, plus le temps de la PL augmente en MCD pour le mutant *cetR* (Figure 39. d). Une hypothèse est que ce changement de milieu (TH puis MCD) provoque un stress pour le mutant, et que CetR joue un rôle dans l'adaptation de *S. agalactiae* au changement de conditions environnementales tel que le milieu de culture. La fragilisation de la paroi cellulaire bactérienne et l'anomalie de formation des chaînettes pourraient aussi avoir un impact sur la croissance du mutant *cetR*.

Chapitre 2 : CetR, un ARNrég ayant un rôle pléiotropique chez *S. agalactiae*

A. Introduction

L'identification des cibles d'un ARNrég est indispensable dans l'étude des mécanismes de régulations modulées par ces biomolécules. Pour cela, il est nécessaire de combiner plusieurs approches *in silico* et *in vivo* afin d'établir une liste de candidats potentiellement régulés par un ARNrég. A l'heure actuelle, aucune approche ne permet d'identifier directement toutes les cibles directes d'un ARNrég. Les approches dites « *in silico* » permettent, *via* des programmes bioinformatiques, de prédire les duplex les plus stables entre l'ARNrég et les ARNm cibles. Néanmoins, ce genre de programme conduit à la détection d'un grand nombre de faux positifs. Il est donc souvent recommandé de comparer les résultats de plusieurs programmes. Les approches dites « *in vivo* » permettent l'identification des ARNm cibles dans des conditions réelles de croissance, mais aussi dans des conditions de stress. Ces approches varient de l'analyse de transcriptome à l'identification du protéome global. La combinaison des approches *in silico* et *in vivo* est indispensable pour l'identification des fonctions régulées par un ARNrég donné.

B. Identification des nouvelles cibles directes ou indirectes de l'ARNrég CetR

1. Identification des cibles directes de CetR par approche *in silico*

Afin d'identifier des nouvelles cibles de CetR, nous avons combiné plusieurs approches *in silico*. Trois programmes bioinformatiques ont été utilisés pour rechercher les appariements entre CetR et les ARNm cibles : 1) le programme **RNApredator** (http://nibiru.tbi.univie.ac.at/cgi-bin/RNApredator/target_search.cgi#) est basé principalement sur le calcul de l'accessibilité des séquences dans les appariements ARNrég/ARNm ; 2) **TargetRNA** (<http://cs.wellesley.edu/~btjaden/TargetRNA2/>) se base sur la recherche des modules conservés dans les appariements ARNrég/ARNm dans des génomes de bactéries proches ; 3) **IntaRNA** (<http://rna.informatik.uni-freiburg.de/IntaRNA/Input.jsp>) fonctionne sur le même principe que RNApredator, avec en plus la possibilité de changer les paramètres dits « par défaut ».

Tableau 6 : Liste des 10 meilleurs ARNm cibles formant potentiellement un duplex avec CetR *in silico*

TargetRNA			RNApredator			IntaRNA		
Cible	ΔG (KJ/mol)	Fonction	Cible	ΔG (KJ/mol)	Fonction	Cible	ΔG (KJ/mol)	Fonction
<i>gbs1502</i>	-14,72	ammonium transporter	<i>gbs1830</i>	-21,62	D-alanine transfer protein	<i>gbs0555</i>	-21,54	serine/alanine adding enzyme
<i>gbs0900</i>	-14,47	uncharacterized protein	<i>gbs0666</i>	-20,28	LysR transcriptional regulator	<i>gbs1172</i>	-19,68	spermidine/putrescine ABC transporter
<i>gbs0108</i>	-14,1	dUTP pyrophosphatase	<i>gbs0682</i>	-19,32	1,2-diacylglycerol-3-alpha-glucose alpha-1,2-glucosyltransferase	<i>gbs1879</i>	-19,34	putative endopeptidase
<i>gbs1216</i>	-13,18	uncharacterized protein	<i>gbs1554</i>	-19,14	formamidopyrimidine-DNA glycosylase	<i>gbs0917</i>	-18,53	GTP-binding protein LepA
<i>gbs1172</i>	-13,06	spermidine/putrescine ABC transporter	<i>gbs2020</i>	-18,73	penicillin-binding protein 2A	<i>gbs0660</i>	-17,54	uncharacterized protein
<i>gbs1096</i>	-12,78	uncharacterized protein	<i>gbs0139</i>	-17,63	cysteine desulfurase / selenocysteine lyase	<i>gbs1556</i>	-17,44	uncharacterized protein
<i>gbs0911</i>	-12,76	CRISPR-associated endonuclease Csn1	<i>gbs0283</i>	-17,12	uncharacterized protein	<i>gbs0666</i>	-17,39	LysR transcriptional regulator
<i>gbs1422</i>	-12,7	uncharacterized protein	<i>gbs1999</i>	-16,88	uncharacterized protein	<i>gbs1534</i>	-16,99	uncharacterized protein
<i>gbs0567</i>	-12,46	uncharacterized protein	<i>gbs1602</i>	-16,85	Ala-tRNA(Pro) deacylase	<i>gbs0207</i>	-16,34	putative transcriptional regulator
<i>gbs0793</i>	-12,39	uncharacterized protein	<i>gbs2011</i>	-16,8	uncharacterized protein	<i>gbs0782</i>	-16,18	uncharacterized protein

En Jaune et bleu, les candidats retrouvés avec deux programmes différents.

Le tableau 6 montre que seulement deux cibles, *gbs1172* et *gbs0666*, sont retrouvées plusieurs fois. La cible *gbs1172*, qui code une spermidine/putrescine ABC transporter, a été identifiée par les trois programmes, dont *RNApredator* en position 148 avec une énergie d'appariement de -12,96 KJ/mol. Les polyamine-like spermidine et putrescine jouent un rôle vital dans la prolifération cellulaire et l'homéostasie des ions [fonction prédite par le serveur web InterPro (<https://www.ebi.ac.uk/interpro/>)]. La cible *gbs0666* est prédite uniquement par *RNApredator* et *IntaRNA* (Tableau 6). La protéine GBS0666 appartient à la famille des régulateurs transcriptionnels LysR. Cette famille de régulateurs est constituée de protéines cytoplasmiques qui reconnaissent des signaux intracellulaires et qui possèdent à la fois un domaine senseur et un domaine effecteur de liaison à l'ADN (Schell, 1993). L'appariement entre CetR/*gbs1172* et CetR/*gbs0666* est localisé sur les régions codantes de ces deux cibles. Nous avons montré précédemment la possibilité de ce mode d'appariement avec la cible *dltD* (*gbs1830*, première cible prédite par *RNApredator*). D'autres cibles intéressantes ont été également prédites par un des trois programmes, dont la *penicillin-binding protein 2A* (*gbs2020*), *CRISPR-associated endonuclease Csn1* (*gbs0911*) et l'*ammonium transporter* (*gbs1502*).

Il faut noter que, même si seules deux cibles sont communes, les critères utilisés entre *RNApredator*, *IntaRNA* et *TargetRNA* sont complémentaires. La conservation de *gbs1172* et *gbs0666* dans les génomes proches, en association avec les régions accessibles à la fois sur l'ARNrég et l'ARNm, en fait de très bons candidats.

2. Identification des cibles de CetR par approche *in vivo*

Nous avons combiné ensuite les résultats des prédictions bioinformatiques avec les résultats de deux approches *in vivo*. La première approche consistait à une analyse comparative du protéome de la souche sauvage et du mutant de délétion $\Delta cetR$ afin d'identifier de nouvelles cibles ou bien de confirmer les cibles déjà identifiées par les autres approches. La deuxième approche consistait à l'identification des ARNm cibles par microarray après la surexpression brève (5 min) de l'ARNrég CetR (Tableau 5).

a- Identification de cibles potentielles de CetR par protéomique comparative

Tableau 7 : Nombre total de protéines identifiées par LC-MS/MS retrouvées dans les différents compartiments cellulaires des deux souches WT et *ΔcetR*

	Cytoplasme	Membrane	Surnageant
Nombre total de protéines identifiées par LC-MS / MS	589	758	73
Nombre total de protéines exprimées significativement de manière différentielle	46	82	23

Afin de compléter les prédictions des cibles de CetR par bioinformatique et d'étudier l'impact de la délétion de *cetR* sur le protéome global de *S. agalactiae*, nous avons utilisé en parallèle une analyse comparative du protéome de la souche sauvage versus le mutant de délétion Δ *cetR* pour identifier et/ou confirmer les cibles régulées par CetR. Nous avons comparé le protéome global (cytoplasme, membrane et surnageant), par LC-MS/MS (voir annexe 3 pour les méthodes), de la souche NEM316 sauvage avec le mutant isogénique Δ *cetR*. Nous avons montré précédemment que la délétion de *cetR* affecte la capacité de *S. agalactiae* à croître en MCD, mais elle n'a pas affecté sa capacité à croître en milieu TH (Figure 39). Le milieu TH a été considéré comme « inapproprié » pour l'analyse LC-MS/MS du protéome global, car il contient des protéines hétérologues à forte concentration qui pourraient interférer avec la détection des protéines de *S. agalactiae*. Nous avons ainsi préparé des extraits protéiques (WT et Δ *cetR*) à partir de cultures en MCD supplémentées avec 0,5% de glucose (voir annexe 3 pour les détails). Comme la croissance du mutant *cetR* est ralentie en MCD par rapport à la souche parentale (Figure 39), nous avons effectué l'extraction des protéines totales à partir du début de la phase stationnaire de croissance. Le dosage des protéines des différents compartiments cellulaires (cytoplasme, membrane et surnageant) a montré une quantité similaire de protéines pour les deux souches (WT et Δ *cetR*). Les extraits protéiques ont été préparés à partir de trois cultures indépendantes.

Afin de vérifier l'absence de contamination (lyse bactérienne, mauvaise séparation des compartiments cellulaires par l'ultracentrifugation,...) durant la préparation des échantillons de protéines, nous avons vérifié l'abondance d'une protéine référence pour chaque compartiment : i) la L-lactate deshydrogénase dans le cytoplasme, ii) la protéine de liaison à la pénicilline 1B dans la membrane et iii) la streptokinase dans le surnageant. Nos résultats indiquent que chacune de ces protéines références est présente uniquement dans le compartiment cellulaire respectif, indiquant ainsi l'absence de contamination. Pour l'analyse des résultats de LC-MS/MS, nous avons considéré qu'une protéine **X** est exprimée différemment, si et seulement si la protéine **X** (moyenne des trois expériences) présente au moins cinq spectre de différence entre les deux souches (WT et Δ *cetR*). Le tableau 7 montre le nombre total de protéines identifiées dans les deux souches et dans chacun des compartiments cellulaires.

Tableau 8 : Protéines exprimées différemment entre le WT et Δ *cetR* identifiées dans le cytoplasme et la membrane

Gene name	Description	Cytoplasm						Membrane					
		Δ <i>cetR</i>		WT		P-value	Regulation ^b	Δ <i>cetR</i>		WT		P-value	Regulation
		Average ^a	SD	Average	SD			Average	SD	Average	SD		
<i>gbs0608 (eno)</i>	Enolase	119,66	11,72	192,33	25,03	1,04E-02	Activator	101,33	3,79	76,67	6,66	5,06E-03	Repressor
<i>gbs0782 (tuf)</i>	Elongation factor Tu	112,66	9,02	83,33	5,86	9,14E-03	Repressor	82,67	8,50	105,33	11,02	4,78E-02	Activator
<i>gbs0268</i>	Transketolase	34,33	1,53	47,33	2,08	9,52E-04	Activator	18,67	2,08	26,33	2,52	1,53E-02	Activator
<i>gbs2096 (rpsD)</i>	30S ribosomal protein S4	33,33	3,06	39,67	1,15	2,83E-02	Activator	38,33	4,04	29,00	1,00	1,78E-02	Repressor
<i>gbs0417 (infB)</i>	Translation initiation factor IF-2	26,67	2,89	21,67	0,58	4,23E-02	Repressor	47,33	0,58	30,33	3,51	1,16E-03	Repressor
<i>gbs1077 (carB)</i>	Carbamoyl-phosphate synthase large chain	30,67	2,52	35,67	0,58	2,85E-02	Activator	37,00	3,61	53,33	4,04	6,41E-03	Activator
<i>gbs1772</i>	Formate C-acetyltransferase	28,67	1,15	34,33	3,06	3,97E-02	Activator	21,67	1,15	40,67	4,04	1,44E-03	Activator
<i>gbs1619</i>	D-3-phosphoglycerate dehydrogenase	30,67	4,73	22,00	0,00	3,37E-02	Repressor	21,00	1,00	13,00	2,00	3,45E-03	Repressor
<i>gbs0493</i>	Uncharacterized protein	24,33	1,15	14,33	0,58	1,79E-04	Repressor	26,67	1,53	19,33	2,08	7,93E-03	Repressor
<i>gbs0780 (ppc)</i>	Phosphoenolpyruvate carboxylase	28,67	5,69	10,00	1,00	4,99E-03	Repressor	24,33	2,08	18,67	1,53	1,91E-02	Repressor
<i>gbs1839</i>	Glycine betaine/proline transport system	12,67	2,08	24,67	1,15	9,48E-04	Activator	26,67	4,73	48,67	2,89	2,34E-03	Activator
<i>gbsp001 (rplA)</i>	50S ribosomal protein L1	12,33	3,06	21,00	2,65	2,06E-02	Activator	20,00	1,73	13,00	3,00	2,49E-02	Repressor
<i>gbs1878</i>	Similar to LD-carboxypeptidase LdcA	18,33	0,58	11,67	0,58	1,45E-04	Repressor	10,00	1,00	5,00	1,00	3,60E-03	Repressor
<i>gbs2026</i>	Uridine phosphorylase	13,67	2,89	8,33	1,15	4,11E-02	Repressor	11,00	1,73	4,33	0,58	3,20E-03	Repressor
<i>gbs1838</i>	Glycine betaine/proline transport system	4,33	0,58	13,33	1,53	6,73E-04	Activator	29,67	2,08	53,67	1,53	8,71E-05	Activator
<i>gbs0497 (asnA)</i>	Aspartate--ammonia ligase	10,33	0,58	2,67	0,58	8,36E-05	Repressor	18,33	1,53	7,00	2,65	3,02E-03	Repressor
<i>gbs0053</i>	Aldehyde-alcohol dehydrogenase	10,33	2,89	1,33	0,58	6,11E-03	Repressor	19,00	1,00	8,00	1,00	1,76E-04	Repressor
<i>gbs0933 (glmS)</i>	Glutamine--fructose-6-phosphate aminotransferase	8,00	1,00	2,67	1,15	3,77E-03	Repressor	7,67	3,06	14,33	0,58	2,06E-02	Activator
<i>gbs0015 (ftsH)</i>	Chaperon, ATP-dependent zinc metalloprotease	5,33	1,53	11,00	1,73	1,32E-02	Activator	40,33	0,58	45,33	1,53	6,07E-03	Activator

SD : Standard deviation. En jaune sont indiquées les protéines dites « *moonlighting* ». ^a moyenne du nombre de spectres. ^b Rôle de CetR (Activator/Repressor)

De manière intéressante, la délétion du gène *cetR* induit des effets multiples sur le protéome de la souche *S. agalactiae* NEM316. Un nombre total de 1420 protéines a été identifié par LC-MS/MS. Ce nombre représente 63,56% du nombre total des protéines prédites pour *S. agalactiae* NEM316 d'après la base de données **Genoscope** (<https://www.genoscope.cns.fr>) (1420 identifiées/2234 prédites et/ou annotées). Le nombre total des protéines exprimées significativement de manière différentielle, entre la souche parentale et le mutant *cetR*, est de 46 protéines dans le cytoplasme, 82 protéines membranaires et 23 protéines identifiées dans le surnageant (Tableau 7). La comparaison des protéines détectées dans les différents compartiments cellulaires, de la souche sauvage et du mutant $\Delta cetR$, a montré que certaines protéines sont **exprimées significativement de manière différentielle** : 1) soit dans deux compartiments (cytoplasme/membrane, cytoplasme/surnageant et membrane/surnageant), 2) soit dans les trois compartiments cellulaires et enfin 3) soit dans un seul compartiment cellulaire. Pour l'ensemble des résultats présentés ci-après, seules les protéines **exprimées significativement de manière différentielle** seront incluses dans l'analyse. Une protéine exprimée significativement de manière différentielle est une protéine dont la valeur de P-value (probabilité statistique calculée avec le test de Student) est inférieure ou égale à 0,01.

Protéines détectées dans le compartiment cytoplasmique et membranaire

Nous avons montré qu'un nombre total de 19 protéines sont exprimées significativement de manière différentielle entre la souche WT et le mutant $\Delta cetR$. Ces protéines ont été identifiées dans les deux compartiments cellulaires cytoplasme et membrane (Tableau 8). L'ensemble de ces protéines peut être regroupé en six classes:

- 1) Les protéines dites « *moonlighting* » (l'enolase, le facteur d'élongation Tu, les protéines S4 et L1 des sous-unités ribosomales 30S et 50S, respectivement, l'aldéhyde-alcool déshydrogénase et la protéine chaperonne FtsH),
- 2) Les protéines participant au métabolisme du pyruvate (la transkétolase, la formate acétyltransferase et la phosphoénolpyruvate carboxylase),
- 3) Celles du métabolisme des pyrimidines (la carbamoyl-phosphate synthase et l'uridine phosphorylase),

4) Celles du métabolisme des acides aminés (la D-3-phosphoglycerate déshydrogénase, l'aspartate-ammonia ligase et la glutamine--fructose-6-phosphate aminotransférase,

5) Les ABC transporteurs (la protéine de liaison à l'ATP et la perméase du système de transport de la glycine et la proline), et

6) Les protéines de fonction inconnue (*gbs0493* et *gbs1878*).

Concernant les protéines dites « *moonlighting* », elles sont multifonctionnelles, c'est-à-dire qu'une protéine unique remplit de multiples fonctions biochimiques qui ne sont pas dues à des fusions de gènes ou de multiples variants de maturation d'ARN (Wang *et al.*, 2014). Plus de 200 protéines ont été identifiées et plusieurs exemples existent, notamment des récepteurs, des enzymes, des facteurs de transcription et des adhésines. Ces protéines sont exprimées par différentes espèces bactériennes, dans différents compartiments cellulaires, fonctionnent dans différentes voies biochimiques et en tant que déterminants de virulence dans des processus infectieux (Henderson and Martin, 2011; Henderson and Martin, 2013).

Parmi ces protéines, nous avons identifié l'enolase, une enzyme clé de la glycolyse. Cette enzyme permet la conversion de 2-phospho-D-glycérate en phosphoénolpyruvate plus une molécule d'H₂O. En effet, cette protéine est plus abondante dans le cytoplasme de la souche sauvage ($192,33 \pm 25,3$ spectre) que dans celui du mutant *cetR* ($119,66 \pm 11,72$ spectre, Tableau 8). Cependant, cette abondance est inversée dans le compartiment membranaire pour la WT et Δ *cetR* (Tableau 8). Ces résultats suggèrent que CetR est activateur de l'enolase dans le cytoplasme alors que l'effet répresseur, observé dans la fraction membranaire, peut être lié à une différence d'exportation de l'enolase vers la membrane. Aucun appariement ARNrég/ARNm n'a été prédit entre CetR et l'ARNm de l'enolase, éliminant ainsi l'hypothèse d'une régulation directe. En plus de sa fonction dans la voie de glycolyse, une étude a montré que l'enolase est présente sur la surface de nombreuses espèces de streptocoque et permet l'adhérence au plasminogène (Pancholi and Fischetti, 1998).

Tableau 9 : Protéines exprimées différemment entre le WT et $\Delta cetR$ identifiées dans le cytoplasme et le surnageant

Gene name	Description	Cytoplasm						Supernatant					
		$\Delta cetR$		WT		Regulation ^b	P-value	$\Delta cetR$		WT		P-value	Regulation*
		Average ^a	SD	Average	SD			Average	SD	Average	SD		
<i>gbs0608 (eno)</i>	Enolase/binds plasminogen	119.67	11.72	192.33	25.03	Activator	1.04E-02	0.00	0.00	7.33	0.58	2.53E-05	Activator
<i>gbs1809 (pgk)</i>	Phosphoglycerate kinase/plasminogen binding	160.00	19.31	108.00	12.17	Repressor	1.69E-02	0.00	0.00	10.33	0.58	6.45E-06	Activator
<i>gbs0931</i>	Pyruvate kinase/binding to invertase	89.00	11.53	128.00	9.54	Activator	1.07E-02	0.00	0.00	5.67	1.15	1.05E-03	Activator
<i>gbs0782 (tuf)</i>	Elongation factor Tu/attachment to human cells and mucins	112.67	9.02	83.33	5.86	Repressor	9.14E-03	0.00	0.00	20.00	3.61	6.56E-04	Activator
<i>gbs0031 (Sip)</i>	Group B streptococcal surface immunogenic protein	11.67	2.08	6.67	1.15	Repressor	2.20E-02	70.67	9.71	42.33	3.06	8.52E-03	Repressor
<i>gbs0580</i>	Zinc transport system substrate-binding protein	0.33	0.58	6.00	1.00	Activator	1.05E-03	0.00	0.00	4.00	1.73	1.61E-02	Activator

Tableau 10 : Protéines exprimées différemment entre le WT et $\Delta cetR$ identifiées dans la membrane et le surnageant

Gene name	Description	Membrane						Supernatant					
		$\Delta cetR$		WT		P-value	Regulation ^b	$\Delta cetR$		WT		P-value	Regulation*
		Average ^a	SD	Average	SD			Average	SD	Average	SD		
<i>gbs0782 (tuf)</i>	Elongation factor Tu/attachment to human cells and mucins	82.67	8.50	105.33	11.02	4.78E-02	Activator	0.00	0.00	20.00	3.61	6.56E-04	Activator
<i>gbs0608 (eno)</i>	Enolase/binds plasminogen	101.33	3.79	76.67	6.66	5.06E-03	Repressor	0.00	0.00	7.33	0.58	2.53E-05	Activator
<i>gbs0942</i>	Basic membrane protein A	54.00	6.56	86.33	2.52	1.34E-03	Activator	1.67	0.58	8.67	1.53	1.76E-03	Activator
<i>gbs1812 (fusA)</i>	Elongation factor G/adhesin	66.67	4.51	79.67	5.13	3.00E-02	Activator	0.00	0.00	10.00	2.65	2.81E-03	Activator
<i>gbs1431</i>	HlyD family secretion protein	23.00	1.00	29.33	1.15	1.99E-03	Activator	1.67	2.08	11.67	1.53	2.57E-03	Activator
<i>gbs1510</i>	Maltooligosaccharide transport system substrate-binding protein	20.33	1.53	6.67	1.15	2.46E-04	Repressor	8.67	3.06	17.67	3.79	3.28E-02	Activator
<i>gbs1042</i>	Iron complex transport system substrate-binding protein	17.00	2.65	5.00	1.73	2.77E-03	Repressor	3.67	1.53	13.67	1.15	8.28E-04	Activator
<i>gbs1688</i>	D-methionine transport system substrate-binding protein	12.67	2.08	6.67	1.53	1.58E-02	Repressor	1.00	1.00	9.67	2.08	2.89E-03	Activator
<i>gbs2008</i>	Lactocepine	0.00	0.00	4.00	2.00	2.57E-02	Activator	24.00	5.00	58.67	5.13	1.11E-03	Activator

SD : Standard deviation. En jaune sont indiquées les protéines dites « moonlighting ». ^a moyenne du nombre de spectres. ^b Rôle de CetR (Activator/Repressor)

Nous avons également identifié d'autres protéines « *moonlighting* » dont le facteur d'élongation (FE) -Tu qui était plus abondant chez le mutant, dans le cytoplasme, par rapport au WT et inversement dans la fraction membranaire (Tableau 8). Cette différence d'expression suggère que CetR régule l'expression de FE-Tu. En effet, nous avons prédit une interaction entre CetR et l'ARNm de FE-Tu (voir ci-après). Ce dernier a été initialement décrit comme une protéine cytoplasmique responsable d'étapes cruciales dans la synthèse des protéines. Cependant, une étude a montré que le FE-Tu joue également un rôle dans l'adhérence aux cellules intestinales et aux mucines humaine chez *Lactobacillus johnsonii* (Granato *et al.*, 2004).

CetR semble également inhiber l'expression de l'aldéhyde-alcool déshydrogénase et activer l'expression de la protéine FtsH, une protéine chaperonne (Tableau 8). Malheureusement, nous n'avons pas réussi à prédire la présence d'un duplex entre CetR et ces deux protéines *moonlighting*.

Protéines détectées dans le compartiment cytoplasmique et le surnageant

En plus de l'enolase et l'FE-Tu, quatre autres protéines (la phosphoglycérate kinase, la pyruvate kinase, la protéine immunogénique de surface et le transporteur ABC du zinc) sont exprimées significativement de manière différentielle dans le cytoplasme et dans le surnageant du WT et du mutant *cetR* (Tableau 9). Tandis que trois protéines semblent être activées par CetR dans le cytoplasme et cinq dans le surnageant, seules trois protéines semblent être inhibées dans le cytoplasme et une dans le surnageant (Tableau 9).

Dans le cytoplasme de *S. agalactiae*, la phosphoglycérate kinase est plus abondante dans le mutant *cetR* ($160 \pm 11,72$ spectre) que dans la WT ($108 \pm 12,17$ spectre). Cette enzyme assure principalement la conversion, dans la voie de glycolyse, du 1,3-bisphosphoglycérate en 3-phosphoglycérate plus une molécule d'ATP. Cependant, dans le surnageant, elle est détectée uniquement chez le WT et permet principalement l'adhérence au plasminogène et à l'actine (Boone and Tyrrell, 2012).

Quant à la pyruvate kinase, elle catalyse la dernière étape de la glycolyse, c'est-à-dire, la conversion du phosphoénolpyruvate (produit de l'enolase) en pyruvate et une

Tableau 11 : Protéines exprimées différemment entre le WT et Δ cetR identifiées uniquement dans le cytoplasme

Gene name	Description	Δ cetR		WT		P-value	Regulation ^b
		Average ^a	SD	Average	SD		
<i>gbs1202</i>	Similar to gls24	45,67	4,04	56,67	4,93	4,04E-02	Activator
<i>gbs1859</i>	Adenylosuccinate synthetase	40,33	4,16	52,67	3,79	1,92E-02	Activator
<i>gbs0837</i>	Ribonucleoside-diphosphate reductase	28,33	3,06	36,33	2,31	2,24E-02	Activator
<i>gbs1265</i>	Acetolactate synthase	36,00	4,36	25,33	0,58	1,37E-02	Repressor
<i>gbs1900 (proS)</i>	Proline--tRNA ligase	23,33	2,08	30,33	2,52	2,06E-02	Activator
<i>gbs2060 (aspS)</i>	Aspartate--tRNA ligase	21,33	2,89	28,00	2,00	3,03E-02	Activator
<i>gbs0198 (pnp)</i>	Polyribonucleotide nucleotidyltransferase	21,33	1,53	27,00	2,65	3,25E-02	Activator
<i>gbs0064 (rpsC)</i>	30S ribosomal protein S3	16,33	1,53	22,33	0,58	3,13E-03	Activator
<i>gbs0047</i>	Adenylosuccinate lyase	14,67	2,89	20,33	1,15	3,43E-02	Activator
<i>gbs1621 (serC)</i>	Phosphoserine aminotransferase	23,67	3,06	10,00	2,00	2,92E-03	Repressor
<i>gbs0106 (pyrG)</i>	CTP synthase	11,00	3,00	20,00	1,00	7,88E-03	Activator
<i>gbs0888 (pheT)</i>	Phenylalanine--tRNA ligase beta subunit	12,33	1,53	18,00	2,00	1,75E-02	Activator
<i>gbs0018</i>	Aromatic amino acid aminotransferase	16,33	3,51	10,00	1,00	3,98E-02	Repressor
<i>gbs1440 (rplU)</i>	50S ribosomal protein L21	7,33	3,79	16,33	3,51	3,92E-02	Activator
<i>gbs1089 (fsh)</i>	Formate--tetrahydrofolate ligase	14,00	4,00	4,33	0,58	1,43E-02	Repressor
<i>gbs0256</i>	N-acetylglucosamine-6-phosphate deacetylase	11,33	1,53	5,33	1,53	8,58E-03	Repressor
<i>gbs0548</i>	DegV family protein	12,00	1,00	3,00	1,00	3,85E-04	Repressor
<i>gbs0334</i>	Malonyl CoA-acyl carrier protein transacylase	12,67	2,52	2,00	0,00	1,83E-03	Repressor
<i>gbs0849 (metK)</i>	S-adenosylmethionine synthase	9,00	1,00	2,67	0,58	6,85E-04	Repressor
<i>gbs2002</i>	Glycerol dehydrogenase	8,33	1,53	3,33	0,58	6,07E-03	Repressor
<i>gbs1667 (Der)</i>	GTPase	7,00	1,73	1,33	0,58	5,78E-03	Repressor
<i>gbs0580</i>	Zinc transport system substrate-binding protein	0,33	0,58	6,00	1,00	1,05E-03	Activator
<i>gbs1073</i>	Phage infection protein	5,67	4,16	24,33	0,58	1,54E-03	Repressor

SD : Standard deviation. ^a moyenne du nombre de spectres. ^b Rôle de CetR (Activator/Repressor)

molécule d'ATP. En effet, cette protéine est plus abondante dans le cytoplasme du WT que dans celui du mutant et absente du surnageant du mutant (Tableau 9). Une étude a montré que cette enzyme cytoplasmique est une protéine « *moonlighting* » qui permet également l'adhérence au mannane (un polysaccharide composé principalement de monomères de mannose) des levures chez *Lactococcus lactis* (Katakura *et al.*, 2010).

Remarquablement, la protéine Sip (*group B streptococcal surface immunogenic protein*) est plus abondante dans le mutant *cetR* que dans la WT, quel que soit le compartiment cellulaire (cytoplasme ou surnageant, voir Tableau 9). Ce résultat suggère que CetR est un régulateur négatif pour la protéine Sip.

Protéines détectées dans le compartiment membranaire et le surnageant

Tandis que la totalité des protéines identifiées dans le surnageant semblent être activées par CetR, seules cinq d'entre elles semblent être activées dans la fraction membranaire (Tableau 10). Parmi ces protéines se trouve le facteur d'élongation G qui catalyse les étapes de translocation des ribosomes lors de la traduction. A la surface de *Streptococcus gordonii*, cette protéine permet en outre l'adhérence au mucine salivaire de la cavité buccale (Kesimer *et al.*, 2009). Nous avons également identifié trois transporteurs ABC qui sont plus abondants dans la fraction membranaire du mutant que dans celle de la WT et inversement dans le surnageant (Tableau 10).

L'analyse de l'ensemble des résultats du protéome a montré que seules l'enolase, et l'FE-Tu, exprimées significativement de manière différentielle entre le WT et le mutant *cetR*, ont été détectées dans les trois compartiments cellulaires.

Protéines détectées dans un seul compartiment cellulaire

La comparaison du protéome global des deux souches WT et mutant *cetR*, par LC-MS/MS, a montré qu'un nombre total de 23 protéines sont exprimées significativement de manière différentielle uniquement dans le cytoplasme (Tableau 11); 56 protéines dans la fraction membranaire (Tableau 12) et 10 seulement dans le surnageant (Tableau 13). Parmi les 23 protéines détectées uniquement dans le cytoplasme, se trouve des protéines impliquées dans le métabolisme des acides aminés, des déshydrogénases, des protéines

Tableau 12 : Protéines exprimées différenciellement entre la WT et $\Delta cetR$ identifiées uniquement dans la fraction membranaire

Gene name	Description	$\Delta cetR$		WT		P-value	Regulation ^b
		Average ^a	SD	Average	SD		
<i>gbs0144</i>	Oligopeptide transport system substrate-binding protein	111,33	9,71	87,67	1,53	1,40E-02	Repressor
<i>gbs0879 (atpA)</i>	ATP synthase subunit alpha	55,00	3,00	65,33	5,51	4,62E-02	Activator
<i>gbs2085 (arcB1)</i>	Ornithine carbamoyltransferase 1, catabolic	52,33	4,51	21,67	4,04	9,31E-04	Repressor
<i>gbs0096 (dnaK)</i>	Hsp70, Chaperone protein	32,67	2,08	52,33	1,15	1,39E-04	Activator
<i>gbs1731 (secA1)</i>	Protein translocase subunit	43,33	2,52	33,00	4,36	2,37E-02	Repressor
<i>gbs2056 (argS)</i>	Arginine--tRNA ligase	37,67	2,52	31,33	2,52	3,69E-02	Repressor
<i>gbs2118 (guaB)</i>	Inosine-5'-monophosphate dehydrogenase	38,67	3,79	27,33	0,58	6,86E-03	Repressor
<i>gbs1477</i>	Interact with sortase family protein	39,67	3,79	24,67	4,93	1,39E-02	Repressor
<i>gbs1662</i>	UPF0755 protein	36,00	2,65	25,00	1,73	3,82E-03	Repressor
<i>gbs1204</i>	Similar to gls24	26,67	1,53	33,67	1,53	4,95E-03	Activator
<i>gbs0061 (rplB)</i>	50S ribosomal protein L2	34,33	3,21	21,00	1,00	2,36E-03	Repressor
<i>gbs1859 (purA)</i>	Adenylosuccinate synthetase	23,33	1,53	32,00	3,00	1,12E-02	Activator
<i>gbs1675</i>	LemA protein	30,33	1,15	19,33	1,15	3,09E-04	Repressor
<i>gbs1478</i>	Interact with manganese-dependent inorganic pyrophosphatase	31,00	2,65	17,67	1,15	1,32E-03	Repressor
<i>gbs0260 (glyS)</i>	Glycine--tRNA ligase beta subunit	21,33	2,31	27,33	1,53	1,99E-02	Activator
<i>gbs0132</i>	Putative lysine transport system permease protein	31,67	4,51	16,33	1,15	4,67E-03	Repressor
<i>gbs1872 (tsf)</i>	Elongation factor Ts	18,00	3,46	29,33	3,51	1,64E-02	Activator
<i>gbs2041</i>	Anaerobic ribonucleoside triphosphate reductase	19,00	4,00	27,67	1,15	2,26E-02	Activator
<i>gbs0058 (rplC)</i>	50S ribosomal protein L3	29,00	2,65	15,33	4,93	1,34E-02	Repressor
<i>gbs0287 (pepC)</i>	Aminopeptidase C	30,33	3,51	13,00	2,65	2,41E-03	Repressor
<i>gbs0155</i>	Penicillin-binding protein 1B	24,00	1,73	17,67	1,15	6,21E-03	Repressor
<i>gbs0842</i>	Polysaccharide deacetylase family protein	9,00	1,00	4,00	0,00	9,78E-04	Repressor
<i>gbs1021</i>	Aminopeptidase N	14,33	1,15	20,00	3,00	3,79E-02	Activator
<i>gbs1864</i>	Interact with Hsp33-like chaperonin	21,67	1,53	11,33	1,53	1,16E-03	Repressor
<i>gbs0002 (dnaN)</i>	DNA polymerase III subunit beta	19,00	1,73	14,00	1,73	2,41E-02	Repressor
<i>gbs2100</i>	DHH family protein	18,33	1,15	12,00	3,46	3,98E-02	Repressor
<i>gbs1407</i>	ABC transporter permease	11,33	1,53	5,67	1,53	1,05E-02	Repressor
<i>gbs0059 (rplD)</i>	50S ribosomal protein L4	16,67	1,15	11,67	1,15	6,07E-03	Repressor
<i>gbs0948 (gyrA)</i>	DNA gyrase subunit A	17,00	0,00	11,00	2,00	6,53E-03	Repressor
<i>gbs1420</i>	Interact with glyceraldehyde-3-phosphate dehydrogenase	19,00	1,00	8,67	3,06	5,10E-03	Repressor
<i>gbs0797 (cshA)</i>	DEAD-box ATP-dependent RNA helicase	9,33	1,53	17,33	2,08	5,82E-03	Activator
<i>gbs2130</i>	ABC transporter ATP-binding protein	10,00	1,00	16,33	3,21	3,11E-02	Activator
<i>gbs1376</i>	ATP-dependent Clp protease ATP-binding subunit ClpL	17,00	1,73	8,33	3,06	1,29E-02	Repressor
<i>gbs1666</i>	Snf2 family protein	16,67	1,53	8,00	2,00	3,97E-03	Repressor
<i>gbs0167 (ackA)</i>	Acetate kinase	7,33	0,58	15,33	3,21	1,32E-02	Activator
<i>gbs1271</i>	dTDP-glucose 4,6-dehydratase	8,33	2,08	13,67	2,08	3,49E-02	Activator
<i>gbs1408</i>	ABC transporter permease	15,33	2,08	5,33	1,15	1,90E-03	Repressor
<i>gbs1405</i>	Glutamate dehydrogenase	14,00	1,73	6,00	1,73	4,81E-03	Repressor
<i>gbs0112 (gltX)</i>	Glutamate--tRNA ligase	7,33	0,58	12,33	1,53	6,07E-03	Activator
<i>gbs0131</i>	Putative lysine transport system ATP-binding protein	13,67	2,52	6,00	2,00	1,45E-02	Repressor

ribosomales ou encore des transporteurs ABC. CetR semble également activer l'expression de 12 protéines et inhiber celle de 11 autres (Tableau 11).

Parmi les 56 protéines, exprimées significativement de manière différentielle, détectées uniquement dans la fraction membranaire, 38 protéines semblent être inhibées par CetR (Tableau 12). Ce résultat suggère que CetR joue le rôle d'un répresseur pour environ 67% de protéines identifiées uniquement dans la fraction membranaire. Trois protéines possèdent une fonction « *moonlighting* », DnaK, la protéine L4 de la sous-unité ribosomale 50S et la glutamate déshydrogénase (Tableau 12). La protéine DnaK (*gbs0096*), appelée également Hsp70 (*heat shock protein 70*), est plus abondante dans la WT ($52,33 \pm 1,15$ spectre) que dans le mutant *cetR* ($32,67 \pm 2,08$ spectre).

Parmi les protéines les plus abondantes du mutant *cetR*, on trouve la protéine de liaison à la pénicilline 1B (*Gbs0155*), appelée également BPB 1B pour « *penicillin-binding protein 1B* » (Tableau 12). Les BPB sont des enzymes impliquées dans les dernières étapes de la synthèse du peptidoglycane (Egan *et al.*, 2015). Elles sont également impliquées dans la résistance aux PAM et jouent un rôle important dans la résistance à l'action des macrophages (Hamilton *et al.*, 2006). Cette protéine représente une véritable cible thérapeutique pour l'industrie biomédicale.

La streptokinase A, synthétisée par plusieurs espèces de streptocoques, a la propriété de se lier au plasminogène humain. Nos résultats de LC-MS/MS ont montré que cette enzyme est environ 4 fois plus abondante dans le surnageant de la souche sauvage que dans celui du mutant *cetR* (Tableau 13). Ce résultat suggère que CetR active l'expression de la streptokinase. Aucun appariement n'a été prédit entre l'ARNm de la streptokinase et CetR. Une étude du mécanisme de cette régulation est nécessaire afin de déterminer la nature de la régulation (transcriptionnelle ou traductionnelle).

Deux protéines impliquées dans la glycolyse ont été détectées uniquement dans le surnageant de la souche sauvage, la glyceraldéhyde-3-phosphate déshydrogénase (GAPDH) et la fructose-bisphosphate aldolase (Fba) (Tableau 13). Chez *S. agalactiae*, la GAPDH catalyse la conversion de glyceraldéhyde-3-phosphate en 1,3-bisphosphoglycérate et agit

Tableau 12 : Suite

<i>gbs1659</i>	Cystine transport system substrate-binding protein	5,67	1,53	13,67	2,52	9,26E-03	Activator
<i>gbs0107</i>	Interact with CTP synthetase	12,00	1,00	6,33	1,15	3,02E-03	Repressor
<i>gbs0133</i>	Interact with adaptor protein	12,33	3,06	5,33	1,15	2,06E-02	Repressor
<i>gbs0579</i>	Dipeptidase	11,00	1,00	5,00	1,00	1,83E-03	Repressor
<i>gbs0843</i>	ATP-dependent RNA helicase, DEAD/DEAH box family	10,67	1,53	5,33	1,53	1,29E-02	Repressor
<i>gbs1494 (rmID)</i>	dTDP-L-rhamnose synthase	5,00	2,00	10,33	1,53	2,14E-02	Activator
<i>gbs1410</i>	ABC transporter ATP-binding protein	12,00	2,65	3,33	1,53	7,97E-03	Repressor
<i>gbs1501</i>	Polar amino acid transport system substrate-binding protein	3,33	0,58	8,33	0,58	4,47E-04	Activator
<i>gbs0055 (thrC)</i>	Threonine synthase	9,33	2,89	2,33	0,58	1,46E-02	Repressor
<i>gbs2066</i>	Uncharacterized protein	8,00	1,00	2,33	0,58	1,05E-03	Repressor
<i>gbs1209</i>	DNA helicase	2,33	1,15	7,67	1,15	4,81E-03	Activator
<i>gbs1928 (relA)</i>	guanosine 3'-diphosphate 5'-diphosphate	1,00	0,00	6,67	0,58	7,02E-05	Activator
<i>gbs1221</i>	Uncharacterized protein	6,00	1,00	0,33	0,58	1,05E-03	Repressor
<i>gbs0327</i>	Aspartokinase	5,67	0,58	0,67	0,58	4,47E-04	Repressor
<i>gbs0690</i>	Putative glutamine transport system substrate-binding protein	5,33	0,58	0,00	0,00	8,92E-05	Repressor
<i>gbs0911</i>	CRISPR-associated endonuclease Cas9	7,00	2,00	1,33	0,58	9,21E-03	Repressor

Tableau 13 : Protéines exprimées différemment entre le WT et Δ cetR identifiées uniquement dans le surnageant

Gene name	Description	Δ cetR		WT		P-value	Regulation ^b
		Average ^a	SD	Average	SD		
<i>gbs0016</i>	PcsB protein	159,33	25,70	93,33	0,58	1,13E-02	Repressor
<i>gbs1195</i>	Streptokinase A	16,33	1,53	62,33	4,51	7,47E-05	Activator
<i>gbs0470</i>	Interact with streptococcal C5a peptidase	36,00	8,72	16,33	4,04	2,39E-02	Repressor
<i>gbs1403</i>	5'-nucleotidase	17,67	2,52	26,67	2,08	8,82E-03	Activator
<i>gbs0280</i>	Polar amino acid transport system substrate-binding protein	5,67	2,52	13,00	3,61	4,46E-02	Activator
<i>gbs1811(GAPDH)</i>	Glyceraldehyde-3-phosphate dehydrogenase	0,67	1,15	14,33	1,15	1,32E-04	Activator
<i>gbs0153</i>	DNA-entry nuclease	2,67	2,52	12,33	3,21	1,48E-02	Activator
<i>gbs2107</i>	LysM domain protein	6,67	2,52	1,33	0,58	2,32E-02	Repressor
<i>gbs0125 (fba)</i>	Fructose-bisphosphate aldolase	0,00	0,00	5,00	1,00	9,78E-04	Activator
<i>gbs2008</i>	Lactocépine	24,00	5,00	58,67	5,13	1,11E-03	Activator

SD : Standard deviation. En jaune sont indiquées les protéines dites « moonlighting ». ^a moyenne du nombre de spectres. ^b Rôle de CetR (Activator/Repressor)

également comme un facteur de virulence avec les lymphocytes B (Madureira *et al.*, 2007). Plusieurs autres fonctions « *moonlighting* » ont été décrites pour cette enzyme cytoplasmique (Henderson and Martin, 2011), détectée principalement dans le surnageant de la souche WT. La fructose-bisphosphate aldolase, quant à elle permet la conversion de fructose-1,6-bisphosphate en dihydroxyacétone phosphate dans la glycolyse.

En résumé, les candidats qui ressortent de l'ensemble de l'analyse protéomique sont :

- Treize protéines « *moonlighting* » (l'enolase, le facteur d'élongation Tu, les protéines S4 et L1/L4 des sous-unités ribosomales 30S et 50S, respectivement, l'aldéhyde-alcool déshydrogénase, la protéine chaperonne FtsH, la phosphoglycérate kinase, la pyruvate kinase, le facteur d'élongation G, la protéine chaperonne Hsp70, la glutamate déshydrogénase et la GAPDH).
- Cinq protéines dans la voie de la glycolyse (la fructose-bisphosphate aldolase, la GAPDH, la phosphoglycérate kinase, l'enolase et la pyruvate kinase) sur dix au total.
- La protéine de liaison à la pénicille 1B (BPB 1B), la protéine Sip (*surface immunogenic protein*) et la streptokinase A.

La protéine DltD a été également identifiée dans cette analyse protéomique comparative, cependant aucune différence significative n'a été observée entre les deux souches (WT et Δ *cetR*) dans les trois compartiments cellulaires étudiés. Ce résultat suggère que la régulation de l'expression de l'ARNm *dltD* par l'ARNrég CetR pourrait être dépendante du milieu. En effet, nous n'avons pas observé de retard de croissance pour la souche *dltD*_{3mut} en MCD. Ce résultat suggère un découplage dans l'action de CetR dépendante de la nature du milieu de culture.

b- Identification des cibles par microarray par qRT-PCR

Nous avons confirmé les cibles identifiées en analyse transcriptomique par une PCR quantitative en temps réel (qRT-PCR). Nous avons choisi un ou deux gènes par opéron pour la validation expérimentale et nous avons utilisé huit gènes de références (*recA*, *gyrA*, *groEL*,

Tableau 14 : Gènes différentiellement exprimés dans le mutant *cetR* induit (I) par rapport au mutant *cetR* non-induit (NI), évalués par microarray et confirmés par qRT-PCR

Opéron	gène	Microarray		qRT-PCR		Fonction prédite (MicroCyc pathways) (http://www.genoscope.cns.fr/microme/microcyc-wsc/)
		Fold Chang (I vs NI)	p-value	Fold Chang (I vs NI)	p-value	
1	<i>gbs0896</i>	1,87	0,005	2,58	0,002	Glycolyse / Gluconéogenèses, métabolisme du pyruvate, biosynthèses des métabolites secondaires
	<i>gbs0897</i>	1,86	0,006	-	-	Cycle citrate (cycle TCA), métabolisme du butanoate, biosynthèses de la valine, leucine et isoleucine
2	<i>gbs1112</i>	0,55	0,007	0,31	0,001	Biosynthèses de la thiamine lipoprotéine
	<i>gbs1113</i>	0,51	0,004	-	-	NADPH-dépendent FMN réductase domain-containing protein
	<i>gbs1114</i>	0,53	0,005	-	-	
3	<i>gbs1202</i>	2,01	0,001	1,91	0,03	Protéine de stress similaire à la famille GlS24
	<i>gbs1203</i>	1,92	0,004	-	-	MF3-like protéine
	<i>gbs1204</i>	1,83	0,007	-	-	Métabolisme de la purine
	<i>gbs1205</i>	1,78	0,013	-	-	Protéine membranaire
	<i>gbs1206</i>	1,8	0,01	-	-	Protéine exportée
	<i>gbs1207</i>	1,84	0,007	1,86	0,01	Transglycosylase
	<i>gbs1208</i>	1,69	0,054	-	-	
4	<i>gbs1504</i>	1,89	0,005	4,8	0,003	Protéine rhodanese-like
5	<i>gbs1786</i>	1,82	0,008	-	-	Phosphorylation oxydative, système à deux composants
	<i>gbs1787</i>	1,87	0,006	-	-	Cytochrome bd-II oxydase, sous-unité I 1.10.3.-
	<i>gbs1788</i>	1,81	0,008	3,01	0,002	Protéine NADH dehydrogenase-like yjID

rpoB, *rpsL*, *glnA*, *adhP* et *pheS*), en collaboration avec E. Jacquet, Institut de Chimie des Substances Naturelles (ICSN), CNRS, UPR 2301, Université Paris-Saclay. Les ARN totaux ont été extraits du mutant *cetR* surexprimant CetR par induction à l'aTc, et du mutant non-induit. Les résultats de la qRT-PCR étaient en accord avec l'analyse transcriptomique (Tableau 14).

De manière surprenante, la comparaison entre les cibles identifiées par microarray et celles prédites *in silico* n'a pas permis de trouver des cibles communes. Seule la cible *gbs1787* a été identifiée par microarray et par *RNApredator* en position 87 et avec une énergie de -13,74 KJ/mol. Le gène *gbs1787* code une cytochrome bd-II oxydase qui permet, chez *E. coli*, la respiration bactérienne en présence de H₂S, souvent consommateur des molécules d'oxygènes (Forte *et al.*, 2016).

3. Synthèse de l'analyse *in silico* et *in vivo*

L'analyse de l'ensemble des résultats de l'approche protéomique montre que la délétion du gène *cetR* induit un effet pléiotropique sur plusieurs voies métaboliques, des voies de biosynthèse, des voies de transport cellulaire ; des protéines impliquées dans la régulation transcriptionnelle ou traductionnelle, dans la signalisation cellulaire et dans virulence. Cette approche globale a montré que l'ARNrég CetR pourrait réguler, directement ou indirectement, l'expression de plusieurs ARNm.

Afin de confirmer les cibles identifiées par LC-MS/MS dans les différents compartiments cellulaires, pour chaque protéine, dont l'expression est dépendante de CetR, nous avons cherché des duplex ARNrég-ARNm prédits par l'approche *in silico*. Nous avons émis l'hypothèse que, la prédiction d'un duplex ARNrég-ARNm stable, pourrait suggérer une régulation directe de la cible d'ARNm par l'ARNrég.

Le tableau 15 montre que, pour l'ensemble des candidats testés, seulement dix protéines ont été identifiées par les deux approches *in silico* et protéomique. Parmi ces candidats, un duplex ARNrég-ARNm a été prédit pour les deux protéines « *moonlighting* », la phosphoglycérate kinase et le FE-Tu, et pour l'endonucléase Cas9. Cependant, l'énergie d'appariement entre CetR et l'ARNm de la phosphoglycérate kinase est faible ($\Delta G = -9.6$). En

Tableau 15 : Cibles d'ARNm prédites par bioinformatique et identifiées par LC-MS/MS

			TargetRNA ^a	RNApredator ^a	IntaRNA ^a	Function
Gene name	Cellular compartment	Regulation ^b				
<i>gbs0911</i>	Membrane	Repressor	-12,76	-	-	CRISPR-associated endonuclease Cas9
<i>gbs1809</i>	Cytoplasm	Repressor	-9,60	-	-	Phosphoglycerate kinase
	Supernatant	Activator				
<i>gbs0782</i>	Cytoplasm	Repressor	-	-	-16,18	Elongation factor Tu
	Membrane	Activator				
<i>gbs0493</i>	Supernatant	Activator	-	-	-15,67	Similar to unknown proteins
	Cytoplasm	Repressor				
<i>gbs1878</i>	Membrane	Repressor	-	-	-13,19	Similar to LD-carboxypeptidase LdcA
	Cytoplasm	Repressor				
<i>gbs0933</i>	Membrane	Repressor	-	-12,62	-	Glyoxylase family protein
	Cytoplasm	Activator				
<i>gbs1900</i>	Cytoplasm	Activator	-	-11,12	-	Phosphatidate cytidyltransferase
<i>gbs2066</i>	Membrane	Repressor	-9,94	-	-	Similar to unknown proteins
<i>gbs0132</i>	Membrane	Repressor	-	-	-13,81	Putative lysine transport system permease protein
<i>gbs1042</i>	Cytoplasm	Repressor	-	-	-12,94	Iron complex transport system substrate-binding protein
	Supernatant	Activator				
<i>gbs1809</i>	Supernatant	Activator	-9,60	-	-	Phosphoglycerate kinase

^a : l'énergie d'appariement est en KJ/mol. ^b : Rôle de CetR (Activator/repressor)

Tableau 16 : Cibles identifiées par LC/MS-MS et par microarray

	Proteomic							
	ΔcetR		WT		P-value	Microarray		Function
	Average	SD	Average	SD		Fold Chang (I vs NI)	P-value	
Cytoplasm								
<i>gbs1202</i>	45.67	4.04	56.67	4.93	4.04E-02	2.01	1.0E-03	Stress protein similar to the Gls24 family
<i>gbs1504</i>	4.67	1.15	0.00	0.00	2.19E-03	1.89	5.0E-03	Rhodanese-like domain protein
Membrane								
<i>gbs1204</i>	26.67	1.53	33.67	1.53	4.95E-03	1.83	7.0E-03	Stress protein similar to the Gls24 family

SD : Standard deviation, I : induit, NI : non-induit

outre, pour chacun de ces trois candidats, le duplex ARNrég-ARNm a été prédit par **TargetRNA** (la phosphoglycérate kinase et l'endonucléase Cas9) ou par **IntaRNA** (FE-Tu). Pour rappel, **RNApredator** prend en compte l'accessibilité des régions d'appariement stable avec un ARNrég. Il est donc probable que **RNApredator** n'ait prédit aucun appariement stable pour ces trois candidats. L'ARNrég CetR semble inhiber l'expression de six des sept candidats identifiés dans le cytoplasme, six des huit prédits et identifiés dans la fraction membranaire et activer l'expression des trois prédits et identifiés dans le surnageant (Tableau 15). De manière surprenante, l'expression du FE-Tu semble être inhibée par CetR dans le cytoplasme et inversement dans la fraction membranaire et dans le surnageant (Tableau 15). La différence dans l'abondance de cette protéine, observée entre les différents compartiments cellulaires, pourrait être liée à d'autres protéines qui vont permettre sa maturation dans le cytoplasme, sa translocation vers la membrane ou son exportation dans le milieu extracellulaire. En effet, ces résultats suggèrent que le FE-Tu est plutôt une cible indirecte de l'ARNrég CetR.

La comparaison entre les candidats identifiés par l'analyse transcriptomique (microarray) et l'analyse du protéome global, montre que seulement trois protéines sont identifiées par les deux approches *in vivo* (Tableau 16). Les résultats de ces deux approches sont concordants sauf pour la cible *gbs1504* (inhibition de l'expression par CetR dans l'approche protéomique et activation de l'expression dans l'analyse transcriptomique). En outre, il est important de noter que les cultures bactériennes utilisées pour ces deux approches sont réalisées dans des milieux différents (THY pour l'analyse transcriptomique et MCD pour l'approche protéomique) et que, dans l'approche protéomique, nous avons comparé la WT avec le mutant *cetR* alors que, dans l'analyse transcriptomique, nous avons étudiés les effets de la surproduction de CetR sur le transcriptome. Ainsi, ces résultats doivent être interprétés avec précaution. Enfin, aucune cible n'a été identifiée par les trois approches.

Les résultats de l'analyse protéomique seront décrits dans une deuxième publication qui est encore de préparation.

Tableau 17 : Répartition des souches de *S. agalactiae* étudiées par PCR en fonction de leur origine, de leur ST et du sérotypage capsulaire

Complexe clonal (nb souches)	Sequence Type (nb souches)	Sérotype capsulaire (nb souches)	Nombre de souches invasives				Nombre de souches de colonisation
			Sang adulte	Sang	LCR NN	Animal	
CC1 (30)	ST1 (20)	V (16) II (2) Ia (1) III (1)	10	-	-	-	10
	ST2 (6)	II (3) NT (2) IV (1)	-	2	-	1	3
	ST196 (4)	IV (3) NT (1)	-	1	1	-	2
CC7 (3)	ST7 (3)	Ia (2) IV (1)	-	-	-	-	3
CC8 (25)	ST8 (10)	Ib (9) NT (1)	6	2	-	-	2
	ST10 (9)	Ib (4) II (2) Ia (1) III (1) V (1)	-	2	-	1	6
	ST12 (6)	Ib (4) II (1) NT (1)	-	3	-	-	3
CC17 (20)	ST17 (20)	III (19) NT (1)	-	-	10	-	10
CC19 (26)	ST19 (21)	III (20) II (1)	-	-	10	1	10
	ST28 (5)	II (4) III (1)	-	-	-	-	5
CC23 (21)	ST23 (21)	Ia (15) III (6)	8	-	4	2	7
Total (125)			24	10	25	5	61

CC : Clonal Complexe ; ST : Sequence Type ; LCR : Liquide céphalorachidien ; NN : nouveau-né

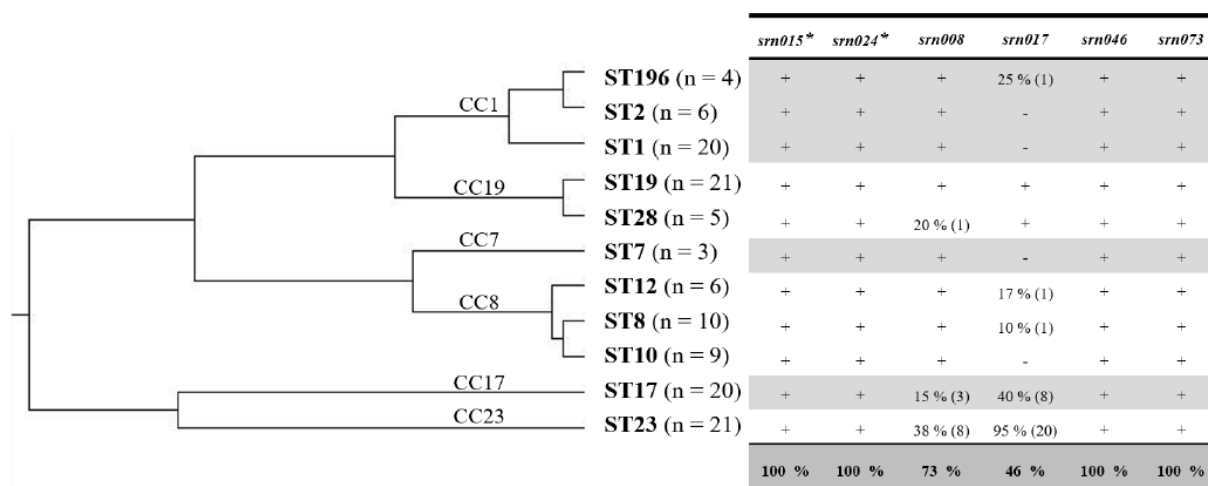


Figure 40 : Relations phylogénétiques entre les 125 souches de *S. agalactiae* typées par *Multi-Locus Sequence Typing* (MLST) et prévalence des six ARNrég étudiée par PCR.

Les dendrogrammes ont été construits selon la méthode UPGMA (*Unweighted Pair Group Method with arithmetic Averages*) avec le logiciel MEGA 6. * : ARNrég régulés par CiaR. (+) : 100 % ; (-) : 0 %

Chapitre 3 : Prévalence des ARNrég chez *S. agalactiae*

A. Introduction

Dans cette partie, nous avons participé à l'étude de la distribution des sept ARNrég, identifiés précédemment (Rosinski-Chupin *et al.*, 2015), en fonction de l'origine écologique des souches et des lignées phylogénétiques auxquelles elles appartiennent. L'étude de la prévalence des ARNrég dans l'espèce *S. agalactiae* a été réalisée par deux approches : 1) une approche *in vitro*, qui a consisté à l'amplification génique (PCR) des gènes codant les ARNrég réalisée sur 125 souches de *S. agalactiae*, représentatives de la diversité génétique de l'espèce et des principaux sites anatomiques où la bactérie est isolée ; 2) une approche *in silico* réalisée sur 249 génomes séquencés de *S. agalactiae* disponibles sur la base de données NCBI (<http://www.ncbi.nlm.nih.gov/>). Cette analyse *in silico* nous a également permis d'étudier la conservation de ces ARNrég à travers l'espèce.

B. Etude de la prévalence des ARNrég par PCR

L'étude de prévalence effectuée par PCR a été réalisée sur une collection de 125 souches de *S. agalactiae*, d'origine humaine et animale (voir annexe 4 pour les méthodes). En plus des critères cités précédemment, les souches ont été également choisies dans le but de représenter les cinq complexes clonaux majeurs retrouvés en pathologie humaine (Tableau 17).

Nos résultats ont montré que quatre des sept gènes codant les ARNrég étudiés (*srn015*, *srn024*, *srn046* et *srn073* (*cetR*)) sont conservés à 100 % dans toutes les souches testées (Figure 40). La prévalence de *srn008* dans l'espèce est de 73 %. Sa présence est égale à 100 % dans la majorité des complexes clonaux : CC1, CC7, CC8 et CC19 en dehors du ST28 et de 15 % et de 38 % pour le CC17 et CC23, respectivement. Quant à *srn017*, il n'est présent que dans 46 % des souches. La prévalence dans le CC19 (n = 26) est de 100 % et inférieure à 3 % dans le CC1. Le gène codant *Srn017* est absent dans le CC7 et présent seulement à 8 % dans le CC8.

Tableau 18 : Prévalence obtenue par analyse *in silico* des sept ARNrég de *S. agalactiae*

Complexe clonal (nb souches)	<i>srn015*</i>	<i>srn024*</i>	<i>srn070*</i>	<i>srn008</i>	<i>srn017</i>	<i>srn046</i>	<i>srn073</i>
CC1 (57)	+	+	+	93 %	5 %	+	+
CC7 (14)	+	+	+	+	-	+	93 %
CC8 (30)	+	97 %	+	+	-	+	+
CC17 (37)	+	97 %	+	8 %	54 %	95 %	+
CC19 (35)	97 %	+	+	89 %	94 %	+	97 %
CC23 (45)	+	98 %	+	31 %	89 %	+	+
CC22 (11)	+	91 %	+	+	-	73 %	+
Autres CC (20)	+	+	+	20 %	-	+	+
Total (249)	> 99 % (248)	> 98 % (245)	100%	64 % (160)	39 % (96)	98 % (244)	> 99 % (247)

* : ARNrég régulés par CiaR. (+) : 100 % ; (-) : 0 %

C. Etude de la prévalence des ARNrég par analyse *in silico*

La prévalence des sept ARNrég a été étudiée par analyse *in silico* parmi les génomes entièrement séquencés mais non assemblés ainsi que les séquences génomiques complètes de 249 souches de *S. agalactiae* disponibles dans la base de données NCBI (voir annexe 4 pour les méthodes). Les résultats obtenus par PCR sont comparables à ceux obtenus par l'analyse *in silico*. En effet, les gènes codant *srn015*, *srn024*, *srn046* et *srn073* sont présents chez plus de 98 % des souches (versus 100 % en PCR). La prévalence de *srn008* est de 64 % (versus 73 % en PCR) et celle de *srn017* est égale à 39 % (versus 46 % en PCR) (Tableau 18). De plus, cette analyse *in silico* confirme une présence supérieure à 98 % des ARNrég régulés par CiaR (100 % en PCR), *srn070* compris. Les souches dans lesquelles un ARNrég est absent, sont souvent des souches dont le génome n'a pas été entièrement séquencé.

D. Synthèse de l'étude de la prévalence

En réunissant l'ensemble des souches analysées par PCR et disponibles sur NCBI, nous disposons de 374 souches de *S. agalactiae* représentatives de la diversité génétique de l'espèce et des principaux sites anatomiques où la bactérie est isolée pour étudier la prévalence des ARNrég dans cette espèce (Tableau 19). Les prévalences de *srn015*, *srn024*, *srn046* et *srn073*, comme déjà observées (Figure 40), sont supérieures à 98 %.

Concernant *srn008*, sa présence est supérieure à 95 % au sein des CC1, CC7, CC8 et CC22. Les différents ST du CC19 sont également caractérisés par une prévalence égale à 100%, exception faite du ST28, à réservoir humain, dont la prévalence est uniquement de 11%. Parmi le CC17, sa prévalence est très faible voire nulle, hormis au sein du ST591, à réservoir strictement bovin, dont la présence est égale à 100% (3 souches/3). La présence de *srn008* parmi les souches du CC23 est faible (33%) et hétérogène, hormis pour le ST88 où elle est de 100%.

Quant à l'ARNrég *Srn017*, il est très peu présent (< 5%) parmi les CC1, CC7, CC8, CC22 et les CC non majeurs. Au sein du CC17, la prévalence de *srn017* est hétérogène. Elle est élevée au sein du ST61 (90%), dont le réservoir est strictement bovin, alors qu'elle est faible (29%)

Tableau 19 : Prévalence des six ARNrég dans une population de 374 souches de *S. agalactiae* étudiées par PCR et analyse *in silico*

CC	ST	Origine	<i>srn015*</i>	<i>srn024*</i>	<i>srn008</i>	<i>srn017</i>	<i>srn046</i>	<i>srn073</i>
CC1 (87)	ST1 (67)	Divers (humain ++)	+	+	95 %	6 %	+	+
	ST2 (10)	Divers (humain ++)	+	+	+	-	+	+
	ST196 (5)	Humain	+	+	+	-	+	+
	Autres (5)	Divers (humain ++)	+	+	80 %	-	+	+
CC7 (17)	ST7 (12)	Divers	+	+	+	-	+	+
	Autres (5)	Divers	+	+	+	-	+	80 %
CC8 (55)	ST8 (21)	Divers (humain ++)	+	+	+	5 %	+	+
	ST10 (13)	Divers (humain ++)	+	+	+	8 %	+	+
	ST12 (17)	Divers (humain ++)	+	94 %	+	-	+	+
	Autres (4)	Bovin, humain	+	+	+	-	+	+
CC17 (57)	ST17 (35)	Humain	+	+	9 %	29 %	+	+
	ST61 (10)	Bovin	+	+	-	90 %	80 %	+
	ST67 (3)	Humain, bovin	+	67 %	-	67 %	+	+
	ST415 (3)	Bovin	+	+	-	67 %	+	+
	ST591 (3)	Bovin	+	+	+	+	+	+
	Autres (3)	Humain, bovin	+	+	-	67 %	+	+
CC19 (61)	ST19 (45)	Divers (humain ++)	+	+	+	98 %	+	+
	ST28 (9)	Humain	+	+	11 %	+	+	+
	Autres (7)	Divers	86 %	+	+	86 %	+	86 %
CC22 (11)	ST22 (11)	Humain	+	91 %	+	-	73 %	+
CC23 (66)	ST23 (52)	Divers (humain ++)	+	98 %	25 %	88 %	+	+
	ST24 (2)	Humain	+	+	-	+	+	+
	ST88 (8)	Divers (humain ++)	+	+	+	+	+	+
	ST452 (2)	Humain	+	+	-	+	+	+
	Autres (2)	Humain	+	+	50 %	+	+	+
Autres CC (20)		Divers	+	+	20 %	-	+	+
Total (374)			> 99 % (373)	> 98 % (370)	67 % (251)	41 % (153)	> 98 % (369)	> 99 % (372)

* : ARNrég régulés par CiaR ; (+) : 100 % ; (-) : 0 %

parmi le ST17 qui comprend uniquement des souches humaines. Elle est égale à 97% et 91 % dans le CC19 et CC23, respectivement.

Les prévalences de *srn008* et *srn017* sont inversées selon les CC, hormis pour le CC19 dont les prévalences sont de 87% et 97%, respectivement. En effet, la présence de *srn008* est élevée (> 95%) parmi les CC1, CC7, CC8 et CC22 alors qu'elle est faible pour *srn017* (< 5%). Lorsque la prévalence de *srn017* est au contraire importante dans un CC, alors celle de *srn008* est moindre (CC17 et CC23).

Enfin, cette étude a permis de mettre en évidence que cinq ARNrég sont présents dans l'ensemble des souches étudiées. De plus, leurs séquences sont bien conservées (> 93 %). Ces observations soutiennent l'hypothèse de leur rôle important dans cette espèce bactérienne. Quant aux gènes *srn017* et *srn008*, ils présentent une prévalence plus faible. Il reste à définir le rôle de ces ARNrég dans la régulation et l'adaptation de *S. agalactiae* aux variations environnementales, ainsi que dans sa virulence.

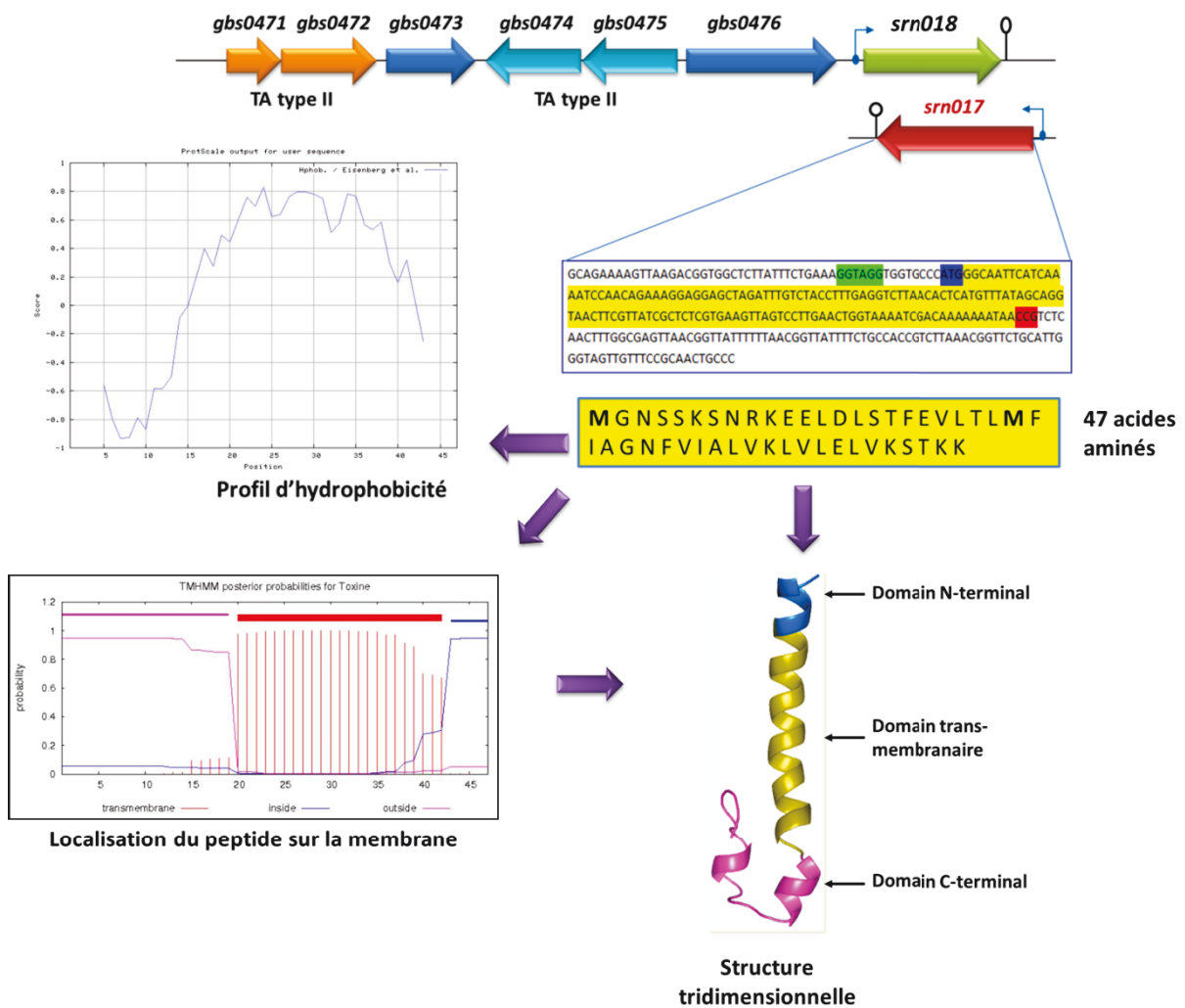


Figure 41 : Diagramme représentant les caractéristiques du duplex Srn017-Srn018.

La séquence du peptide potentiel codé par Srn017 a été prédite par l'outil Translate (<http://web.expasy.org/translate/>). Le profil d'hydrophobicité a été prédit par l'outil ProtScale (<http://web.expasy.org/protscale/>). La localisation du peptide sur la membrane a été prédite par l'outil HMMTP (<http://www.enzim.hu/hmmtop/>). La structure tridimensionnelle a été prédite par le serveur web Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) et visualisée par PyMOL (<https://www.pymol.org/>).

Chapitre 4 : Etude d'autres ARNrég de *S. agalactiae*

A. Introduction

Dans ce chapitre, nous présentons les résultats préliminaires sur la caractérisation des ARNrég Srn008, Srn015, Srn017, Srn024, Srn046 et Srn070. Tout d'abord, nous avons construit les mutants de délétion de Srn008, Srn024, Srn046 et Srn070 ainsi que le double mutant $\Delta srn024\Delta srn015$. Une prédiction bioinformatique a été conduite sur l'ARNrég Srn017, antisens à un autre ARNrég Srn018, potentiel système toxine-antitoxine.

B. Le duplex Srn017-Srn018 est-il un système Toxine-Antitoxine ?

L'expression de Srn017 a été validée par northern blot et est maximum en mi-phase exponentielle et en fin de phase exponentielle de croissance (*cf.* Figure 27, Rosinski-Chupin et al, 2015). De plus, l'expression de Srn017 est induite en réponse aux stress lysozyme et oxydant (*cf.* Figure 28 Rosinski-Chupin et al, 2015). Cet ANRrég est un antisens de Srn018 (Rosinski-Chupin *et al.*, 2015). Le duplex Srn017-Srn018 est transcrit à partir d'une région intergénique avoisinant des gènes codant deux systèmes TA de type II. Une prédiction bioinformatique, *via* le serveur web Expasy (<https://www.expasy.org/>), a montré également que Srn017 code un peptide de 47 acides aminés similaire à certaines porines membranaires (Figure 41). Il est connu que ce type de peptide, souvent hydrophobe (Figure 41), peut jouer le rôle d'une toxine en s'intercalant dans la membrane bactérienne, grâce à sa structure en hélice alpha, afin de déstabiliser sa perméabilité (Kawano *et al.*, 2007). L'action de la toxine est généralement neutralisée par l'antitoxine, souvent un ARN antisens ou un autre peptide.

Ces premières données laissent penser que Srn017 et son antisens (Srn018) font partie d'un système TA. Les résultats préliminaires tels que la disposition génétique de *srn017* par rapport à *srn018*, leur expression induite en condition de stress et la potentielle nature peptidique de Srn017 sont autant d'éléments en faveur d'une implication potentielle dans un système TA. Une validation expérimentale est nécessaire afin de valider cette hypothèse. Ces systèmes TA ont récemment émergé comme centre de la régulation de divers processus cellulaires qui modulent l'activité bactérienne, y compris la réponse générale au stress, la formation de biofilm et la persistance.



Figure 42 : Structures secondaires des csRNA de la souche *S. agalactiae* NEM316.

Les séquences complémentaires à la séquence Shine-Dalgarno (ou site d'initiation de la traduction) sont représentées et soulignées en rouge. Les séquences représentées et soulignées en bleu représentent une séquence conservée de nucléotides retrouvée dans tous les csRNA.

D'après (Marx *et al.*, 2010)

Tableau 20 : ARNm cibles de Srn015 prédites *in silico*

gene	annotation	RNApredator		IntaRNA		TargetRNA		Position de l'interaction
		Position	Energie*	Position	Energie*	Position	Energie*	
<i>gbs1243</i>	glycosyl transferase CpsE	2	-16.36	1	-16.49000	1	-14.46	
<i>gbs0890</i>	exonuclease RexB	87	-11.87	30	-11.63290	26	-10.24	
<i>gbs0101</i>	-	36	-13.22	31	-11.58770	12	-11.31	
<i>gbs0052</i>	acyltransferase	25	-13.55	41	-10.96620	21	-10.66	
<i>gbs0538</i>	similar to amino acid ABC transporter (ATP-binding protein)	5	-15.24	6	-15.06620	-	-	
<i>gbs0100</i>	phosphomethylpyrimidine kinase	40	-13.05	7	-14.66640	-	-	
<i>gbs0241</i>	weakly similar to DNA translocase	18	-13.79	17	-13.12840	-	-	

En Jaune : cible prédite par les trois programmes ; En orange : cible prédite par deux programmes

C. Les ARNrég Srn015, Srn024 et Srn070 ont-ils le même mode d'action ?

Afin de caractériser les ARNrég associés au système CiaRH (Srn015, Srn024 et Srn070 ; Figure 42), nous avons construit les mutants de délétion de Srn024 et Srn070 et le double $\Delta srn024\Delta srn015$. Le mutant $\Delta srn015$ était déjà disponible au laboratoire. Aucun phénotype n'a été observé après la délétion de ces ARNrég, comme pour les deux autres mutants $\Delta srn008$ et $\Delta srn046$. Afin de caractériser la fonction de ces ARNrég, une étude phénotypique complémentaire, tels que des tests de survie en stress et/ou de compétition, est nécessaire.

Nous avons également prédit *in silico* les cibles ARNm de ces csRNA. La première cible prédite pour Srn015 est le gène *gbs1243*, codant une glycosyl transférase CpsE impliquée dans la synthèse de la capsule (Tableau 20). Cette cible a été prédite par les trois logiciels (RNApredator, AntaRNA et TargetRNA). D'autres cibles ont été également prédites pour Srn015, dont la zone d'interaction inclut la région d'initiation de la traduction (Tableau 20). Un résultat identique concernant la zone d'interaction a été observé pour les Srn024 et Srn070 (voir annexe 5 et 6). La comparaison des résultats de prédiction des cibles de Srn015 et Srn070 a montré que les gènes *gbs0101* et *gbs0100* sont des cibles communes. Les gènes *gbs0101* et *gbs0100* codent un transporteur de riboflavine et une phosphométhylpyrimidine kinase, respectivement. L'étude de l'expression des cibles d'ARNm prédites *in silico* pourrait confirmer la présence d'une potentielle régulation directe par les ARNrég Srn015, Srn024 et Srn070.

En outre, l'alignement des séquences des ARNrég Srn015, Srn024 et Srn070 a montré la conservation de plusieurs motifs, notamment le motif de liaison de régulateur CiaR (voir annexe 7). Cette conservation suggère que ces ARNrég pourraient avoir des fonctions similaires.

Les différents mutants construits durant cette thèse ainsi que les premiers résultats de prédictions seront indispensables pour étudier le rôle de ces ARNrég dans la régulation chez *S. agalactiae*.

DISCUSSION ET PERSPECTIVES

Discussion et perspectives

L'analyse du transcriptome par séquençage massif a montré que le génome de *S. agalactiae* héberge un nombre important d'ARNrég (Rosinski-Chupin *et al.*, 2015). Ces biomolécules apparaissent désormais comme une classe majeure de régulateurs de l'expression de gènes dans de nombreux micro-organismes (Waters and Storz, 2009), mais également dans des organismes supérieurs (Morris and Mattick, 2014). Chez l'Homme, de nombreuses études ont montré que les micro-ARN (miRNA) sont impliqués dans diverses maladies tels que le cancer (Brower *et al.*, 2014), la maladie d'Alzheimer (Lukiw *et al.*, 2012) ou encore le diabète (Witkowski *et al.*, 2016). La caractérisation de la fonction de leurs homologues procaryotes, les ARNrég, nécessite la mise en place de différentes approches *in silico*, *in vitro* et *in vivo*.

A. Distribution des ARN régulateurs chez *S. agalactiae*

Nous avons effectué une analyse de prévalence de sept ARNrég chez *S. agalactiae*. L'objectif de cette étude était de déterminer s'il existe une distribution particulière des ARNrég en fonction de l'origine écologique des souches ou des lignées phylogénétiques auxquelles elles appartiennent. Notre démarche consistait à rechercher la présence des gènes des ARNrég d'une part par amplification génique au sein d'une population de 125 souches caractérisées sur le plan phylogénétique, et d'autre part par analyse *in silico* sur 249 génomes séquencés de *S. agalactiae* disponibles dans la base de données NCBI.

Les résultats de l'analyse *in silico* confirment ceux obtenus par PCR, c'est-à-dire une prévalence de 100 % dans l'ensemble des souches étudiées des gènes codant les ARNrég suivants : *srn015*, *srn024*, *srn046*, *srn070* et *srn073*. Ce résultat suggère que ces cinq gènes, présents dans l'ensemble des 374 souches étudiées, appartiennent au « core » génome de *S. agalactiae*. A propos de *srn017*, sa prévalence est la plus faible parmi les sept ARNrég étudiés, et sa répartition est hétérogène au sein des différents complexes clonaux. A l'inverse de tous les autres ARNrég étudiés, *srn017* et *Srn008* ont été trouvés dans le génome d'autres espèces de streptocoques, tels que *S. dysgalactiae* et *S. pyogenes*. Ils ne sont donc pas spécifiques de *S. agalactiae*.

Enfin, nous n'avons pu mettre en évidence aucune relation entre l'origine écologique (humaine, bovine, ichthyique), le caractère invasif des souches (colonisation vaginale *versus* méningite ou bactériémie néonatale) et la prévalence des gènes des ARNrég. Ces derniers ne semblent donc pas spécifiques d'une niche écologique, ni associé au caractère invasif des souches. Néanmoins, la conservation de certains de ces ARNrég dans l'espèce *S. agalactiae* suggère un rôle dans la physiologie et/ou la virulence de cette bactérie.

Au cours de notre recherche d'homologues de CetR, nous avons observé que la séquence du gène codant cet ARNrég est exclusive à l'espèce *S. agalactiae*. Remarquablement, la recherche bioinformatique des homologues de *dltD* a montré qu'il a un profil de conservation similaire à son régulateur, CetR. En effet, les séquences de *dltD* sont hautement conservées au sein de tous les génomes de *S. agalactiae* disponibles dans la base de données NCBI. En revanche, elles ont moins de 50% d'identité de séquence nucléotidique avec d'autres streptocoques. Ces observations suggèrent que des espèces bactériennes phylogénétiquement proches ont développé des réseaux de régulation distincts qui reposent sur des interactions spécifiques entre les ARNrég et leurs ARNm cibles. Ce type d'interaction spécifique pourrait être crucial pour l'adaptation des bactéries à un environnement particulier, ou pour favoriser l'infection de leurs hôtes. Il est connu que les ARNrég se lient à des petites régions de leurs ARNm cibles (Salvail *et al.*, 2010; Rice *et al.*, 2012). Nous pouvons supposer que des mismatch nucléotidiques dans ces régions pourraient altérer la fonction de l'ARNrég ou affecter la régulation des ARNm cibles. L'alignement multiple de 280 séquences de gènes codant CetR a révélé la présence de mismatch nucléotidiques répartis sur toute la longueur de l'ARNrég dans douze complexes clonaux de *S. agalactiae* (Figure S1A, publication 1). De manière intéressante, un seul mismatch nucléotidique (G267A) est localisé dans la région d'interaction CetR-*dltD*. Il sera intéressant d'étudier l'impact de ce mismatch nucléotidique sur la structure secondaire de l'ARNrég ou sur l'interaction avec la cible de l'ARNm.

Une recherche des séquences CetR-like chez *S. mutans*, *S. salivarius*, les streptococques du groupe *bovis* (*infantarius*, *lutetiensis*, *pasteurianus*, *macedonicus* et *gallolyticus*), les streptococques du groupe *pyogenes* à l'exception de *S. pyogenes* (*S. dysgalactiae*, *S. equi*, *S. parauberis*), on trouve en amont du gène *mutY* une région de 500 nt

à 1100 nt non-codante, dans laquelle on peut trouver une séquence promoteur et une séquence terminateur qui prédisent donc l'existence d'un ARNrég potentiel.

Ces séquences promotrices s'alignent bien avec celles de *cetR* et ceci malgré le fait que toutes ces espèces sont éloignées phylogénétiquement. La conservation de cette séquence est forte à l'intérieur du groupe *bovis* par exemple, c'est-à-dire quand les espèces sont relativement proches. Ces résultats suggèrent, dans un certain nombre de groupes de streptocoques, l'existence d'un ARNrég localisé au même endroit que CetR et présentant une identité de séquence avec CetR. Cette identité est faible et partielle suggérant que ces ARNrég sont des orthologues de CetR. Il est donc possible que CetR possède des homologues chez d'autres *Streptococci* mais qu'il a probablement évolué afin de remplir une fonction spécifique chez *S. agalactiae* en lien avec la régulation de *dltD* et la voie de D-alanylation.

B. CetR, un long ARN régulateur qui a un rôle pléiotrope

La majeure partie de ma thèse porte sur la caractérisation de l'ARNrég CetR chez notre souche modèle, *S. agalactiae* NEM316. CetR est un des rares ARNrég de grande taille caractérisé à ce jour. Le plus souvent, l'étude de la fonction des gènes bactériens est intimement associée à l'identification des phénotypes qui résultent de l'inactivation de ces mêmes gènes. L'inactivation du gène codant CetR induit de nombreux phénotypes associés au métabolisme de la membrane cellulaire, au métabolisme du carbone ou encore à l'expression des facteurs de virulence, tels que la streptokinase, la protéine de liaison à la pénicilline ou la protéine Sip. En effet, d'autres ARNrég pléiotropes ont été identifiés et caractérisés. L'ARNrég RNAIII est l'un des régulateurs pléiotropes les mieux caractérisés chez *S. aureus*. Il contrôle, au cours de la croissance bactérienne, les niveaux d'expression de deux facteurs de transcription aux effets antagonistes que sont AgrA et Rot (*repressor of toxins*) (Boisset *et al.*, 2007). RNAIII agit également directement, par un mécanisme antisens, sur les ARNm de plusieurs protéines impliquées dans la virulence, telles que *spa* et *coa*, codant respectivement une protéine d'adhérence aux cellules humaines et une coagulase du plasma sanguin. De plus, le gène RNAIII promeut l'expression de l'hémolysine A en se liant à la région 3' de l'ARNm et en empêchant la formation d'une structure secondaire inhibitrice au

sein de l'ARNm (Morfeldt *et al.*, 1995). Un autre ARN pléiotrope est l'ARNrég DsrA d'*E. coli* qui régule l'expression de plusieurs gènes, tels que *rpoS*, codant le facteur σ^S de la phase stationnaire, *hns*, un régulateur transcriptionnel global (Majdalani *et al.*, 1998), et plus récemment l'ARNm *mreB*, régulateur de la morphologie cellulaire (Cayrol *et al.*, 2015). La surexpression de DsrA induit des phénotypes pléiotropiques incluant des colonies de morphologie mucoïde (Sledjeski and Gottesman, 1995).

Nous avons démontré que l'expression de CetR est cruciale pour la modification de la membrane cellulaire, la formation du biofilm, la résistance aux PAM et la virulence à travers l'unique activation post-transcriptionnelle de l'expression de l'ARNm *dltD*. Le gène *dltD* appartient à l'ARNm polycistronique *dltABCD*, qui forme la principale composante dans la biosynthèse de D-alanyl-ALT. Chez *S. aureus*, l'incorporation des résidus des D-Ala est importante pour la virulence dans un modèle animal (Peschel *et al.*, 1999; Collins *et al.*, 2002). Chez *S. agalactiae*, le dysfonctionnement dans la voie de biosynthèse de D-alanyl-ALT se traduit par la sensibilité aux PAM et une virulence atténuée de cette bactérie (Poyart *et al.*, 2003). Pendant plus d'une décennie, le TCS DltR/DltS était considéré comme l'unique régulateur de l'opéron *dlt*. L'expression de l'ARNm polycistronique est activée par DltR/DltS lorsque la quantité de l'ester D-Ala diminue sur la membrane cellulaire (Poyart *et al.*, 2001).

Dans ce travail de thèse, nous avons montré que la régulation post-transcriptionnelle par CetR est tout aussi importante. Bien que le mécanisme exact par lequel CetR active la traduction de DltD ne soit pas démontré, notre étude souligne l'importance de cette régulation dans la résistance aux PAM.

C. CetR, un régulateur global de virulence ?

Les réseaux de régulation des bactéries sont dynamiques, ils intègrent une multitude de signaux aboutissant à la régulation fine de nombreux processus cellulaires, notamment ceux impliquant l'expression de facteurs de virulence. Nous avons suivi la survie des bactéries lors de l'infection d'un modèle animal pour étudier le rôle de CetR dans la pathogénicité de *S. agalactiae in vivo*. La délétion du gène codant CetR induit une mortalité bactérienne pendant l'invasion du foie et du cerveau, ou la dissémination dans le sang de souris infectées (publication 1). Des résultats similaires ont été obtenus avec la souche

*dltD*_{3mut} (publication 1). Ce résultat renforce notre conclusion dans laquelle CetR active l'expression de *dltD* pour promouvoir la virulence de *S. agalactiae*. Nous émettons l'hypothèse que l'élimination efficace des mutants Δ *cetR* et *dltD*_{3mut} chez la souris est une conséquence de leur sensibilité aux PAM.

La réponse de l'hôte à l'infection par *S. agalactiae* comprend également l'activation des macrophages et d'autres composants du système immunitaire inné (Maisey *et al.*, 2008). En effet, nous avons montré que les cellules Δ *cetR* et *dltD*_{3mut} sont plus sensibles aux macrophages par rapport aux souches sauvage et complémentée (publication 1). Il est connu que les souris produisent également des peptides antimicrobiens (CRAMPs, *cathelin-related antimicrobial peptides*) (Gallo *et al.*, 1997). Ainsi, nous suggérons que la survie de *S. agalactiae* dans les souris est contrôlée par DltD, d'une manière CetR-dépendante, à travers la résistance aux PAM produits par les macrophages.

La compréhension du réseau de régulation liée à l'expression de CetR est une des perspectives majeures de ce travail de thèse. Il serait intéressant d'étudier ce réseau de régulation dans des cellules uniques en utilisant des outils de microfluidique, ou du « *single cell* ». En effet, il sera alors possible d'évaluer l'expression des gènes à l'échelle d'une cellule unique. En régulant diverses voies métaboliques et en modulant l'expression de différents facteurs de virulence, CetR induit des effets pléiotropes qui pourraient affecter la viabilité des cellules. Si certaines cellules se lysent plus facilement, l'ADN exogène ainsi généré pourrait servir comme matrice à la formation du biofilm. Cependant, nos résultats indiquent que le mutant *cetR* forme significativement moins de biofilm par rapport à la souche parentale (publication 1). On peut alors supposer que, en condition de stress, les cellules du mutant *cetR* lysent très rapidement et ne laissent pas même le temps à la bactérie d'adhérer et donc d'initier la formation du biofilm. En effet, le mutant *cetR* adhère significativement moins aux protéines de la matrice extracellulaire (résultats complémentaire, chapitre 1).

Le rôle de CetR dans la viabilité cellulaire avant et après stress pourrait être évalué *via* le kit *LIVE/DEAD BacLight bacterail viability*. Le principe est de marquer simultanément les membranes des bactéries avec une sonde fluorescente et l'ADN génomique avec l'iodure

de propidium. Si la bactérie lyse lors d'un stress, le marquage de l'ADN génomique sera d'avantage visible que celui des membranes.

D. CetR, régulateur des protéines *moonlighting* chez *S. agalactiae* ?

Nous avons montré également par une approche protéomique que l'ARNrég régule significativement l'expression d'environ 6,8 % (151 protéines) des protéines codées par le génome de la souche *S. agalactiae* NEM316 (Chapitre 2). Ces protéines sont réparties dans les trois compartiments cellulaires (cytoplasme, membrane et surnageant). Une étude comparative effectuée uniquement sur le surnageant de la souche *S. agalactiae* NEM316 et le mutant du TCS CovRS a permis d'identifier 53 protéines (Papaserigi *et al.*, 2013). Des protéines telles que la GAPDH, la Sip, le domaine de protéine LysM ou la 5-nucléotidase ont été également identifiées dans notre étude. La spectrométrie de masse est devenue l'une des techniques les plus utilisées en protéomique en raison de son forte efficacité dans l'identification des protéines trouvées dans différents types de cellules, organelles et complexes multiprotéique. Ces informations sur la localisation des protéines peuvent fournir des indices précieux pour déduire la fonction d'une protéine.

En effet, la spectrométrie de masse est de plus en plus utilisée pour l'identification des protéines dites « *moonlighting* » (Jeffery, 2005). Dans notre étude, nous avons identifié treize protéines *moonlighting*, dont des facteurs de virulence tels que l'enolase et la GAPDH, réparties sur les différents compartiments cellulaires (Chapitre 2). Nous avons également identifié un nombre important de protéines dont la fonction *moonlighting* n'est pas encore caractérisée chez *S. agalactiae* (l'aldéhyde-alcool déshydrogénase, la protéine FtsH, la protéine DnaK, La glutamate déshydrogénase et les protéines S4, L4 et L1 des sous-unités ribosomales 30S et 50S).

En plus de sa fonction dans la voie de glycolyse, l'enolase permet l'adhérence au plasminogène. L'adhérence au plasminogène permet au streptocoque de pénétrer la membrane basale et d'envahir les cellules eucaryotes (Eberhard *et al.*, 1999). L'inactivation du gène codant l'enolase chez *S. pneumoniae* a conduit à la mort des cellules, montrant ainsi que ce gène est essentiel pour cette bactérie (Bergmann *et al.*, 2001). D'autres rôles ont été décrits récemment pour l'enolase de *S. pneumoniae*, dont l'adhérence à l'inhibiteur du

complément humain C4b, l'échappement à l'action du complément (Agarwal *et al.*, 2012), l'adhérence aux neutrophiles (Mori *et al.*, 2012) et l'adhérence aux cellules épithéliales et endothéliales (Bergmann *et al.*, 2013).

L'aldéhyde-alcool déshydrogénase est une enzyme de la voie métabolique ayant une fonction « *moonlighting* ». Elle interagit avec une autre protéine ayant une fonction « *moonlighting* » (Jagadeesan *et al.*, 2010). Cette enzyme a été initialement décrite comme la principale adhésine de *L. monocytogenes* (LAP, *Listeria* adhesine protein). La LAP permet l'adhérence aux cellules épithéliales intestinales (Santiago *et al.*, 1999), la liaison aux protéines chaperonnes Cpn60 (*chapronin* 60) et Hsp60 (*heat shock protein* 60) de la cellule hôte (Wampler *et al.*, 2004). La protéine Hsp60 est considérée également comme une protéine « *moonlighting* ». En effet, l'interaction entre ces deux protéines « *moonlighting* », l'aldéhyde-alcool déshydrogénase et l'Hsp60 identifiées dans un site topologiquement inhabituel, permet à ce pathogène de coloniser les cellules humaines.

Dans le cytoplasme, la protéine FtsH joue le rôle d'une protéine chaperonne (Suzuki *et al.*, 1997). Cependant, si cette protéine est exportée vers la membrane, elle joue le rôle d'une métallopeptidase ou d'une hydrolase qui permet la dégradation du facteur transcriptionnel σ^{32} chez *E. coli* (Tomoyasu *et al.*, 1995).

La protéine DnaK, une protéine chaperonne, a été identifiée précédemment dans la fraction membranaire de *S. agalactiae* (Hughes *et al.*, 2002). Cependant, plusieurs autres fonctions ont été décrites pour DnaK chez *Mycobacterium tuberculosis*, notamment la stimulation de la production des chimiokines des lymphocytes CD8 (Lehner *et al.*, 2000) et la synthèse des chimiokines des monocytes CD40 (Wang *et al.*, 2001). Cette protéine se lie au récepteur CCR5 du VIH (Floto *et al.*, 2006) et permet l'adhérence au plasminogène humain (Xolalpa *et al.*, 2007). Quant à la glutamate déshydrogénase (RocG), elle catalyse la désamination du glutamate pour former l' α -kétooglutarate. Cette enzyme métabolique se lie également au facteur de transcription GltC, qui régule la production de glutamate à partir du l' α -kétooglutarate (Commichau *et al.*, 2007).

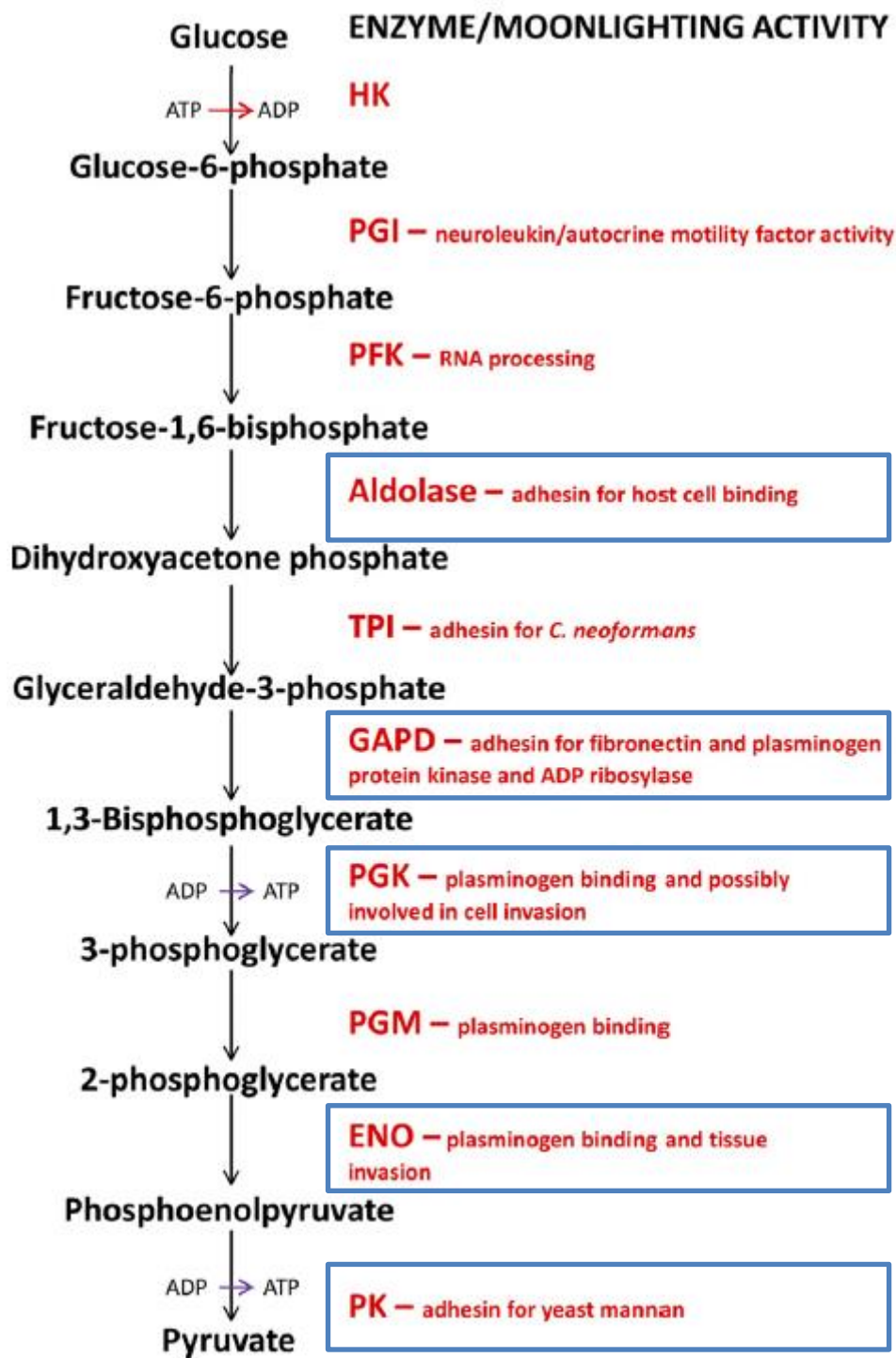


Figure 43 : La voie de glycolyse et l'activité *moonlighting* des protéines glycolytiques bactériennes.

Les protéines encadrées sont les protéines identifiées dans notre étude. HK, hexokinase; PGI, phosphoglucose isomérase; PFK, phosphofructokinase; TPI, triose phosphate isomérase; GAPDH, glycéraldéhyde 3-phosphate déshydrogénase; PGK, phosphoglycérate kinase; PGM, phosphoglycérate mutase; ENO, enolase; PK, pyruvate kinase.

Enfin, les protéines S4 et L1 des sous-unités ribosomales 30S et 50S, respectivement, jouent un rôle d'inhibiteur traductionnel chez *E. coli* (Yates *et al.*, 1980). Pour sa fonction « *moonlighting* », la protéine L4 de la sous-unité ribosomale 50S agit comme un répresseur transcriptionnel, causant la terminaison prématurée de la transcription de l'opéron S10 chez *E. coli* (Zengel and Lindahl, 1996).

La présence d'une protéine dans un compartiment cellulaire inattendu, ou l'observation des résultats inattendus, peut suggérer que la protéine a une autre fonction. Cependant, cette conclusion n'est pas toujours vraie. Certaines protéines à fonction unique se trouvent dans plusieurs compartiments parce qu'elles catalysent la même fonction dans les différents compartiments cellulaires (Par exemple, une activité kinase dans le cytoplasme et au niveau de la membrane). Ainsi, une protéine ayant la même fonction dans deux compartiments différents n'est pas une protéine *moonlighting*. Idéalement, pour clarifier la (les) fonction(s) des nombreuses protéines étudiées dans des projets de protéomique, il serait utile de combiner les résultats de plusieurs méthodes, telles que les données de spectrométrie de masse avec l'inactivation des gènes et/ou des systèmes à double hybride.

Nous avons montré que CetR semble inhiber l'expression de la protéine Sip. Cette protéine semble être régulée négativement par un autre ARNrég, SQ18, identifié par Pichon et ses collaborateurs (Pichon *et al.*, 2012). L'immunisation avec une protéine Sip recombinante a permis de protéger des souris CD-1 contre l'infection avec six souches de *S. agalactiae* de sérotypes Ia/c, Ib, II/R, III, V, et VI (Brodeur *et al.*, 2000). Les auteurs de cette étude suggèrent que cette protéine hautement conservée chez *S. agalactiae* est un potentiel candidat pour la conception d'un vaccin universel pour lutter contre les infections de ce pathogène opportuniste.

Nous avons montré également que cinq protéines de la voie de la glycolyse (la fructose-bisphosphate aldolase, la GAPDH, la phosphoglycérate kinase, l'énolase et la pyruvate kinase) sont régulées par l'ARNrég CetR (Figure 43), dans un ou plusieurs compartiments cellulaires (Chapitre 2). Ce résultat peut expliquer le ralentissement de croissance observé pour le mutant *cetR* dans le milieu chimiquement défini (MCD) supplémenté avec 0,5% de glucose (Chapitre 2). La voie de la glycolyse étant complètement

perturbée, le mutant est forcé d'allonger sa phase de latence afin d'activer d'autres voies et d'autres enzymes, telles que de nombreuses enzymes impliquées dans le métabolisme des acides aminés ou encore la protéine *moonlighting*, l'aldéhyde-alcool déshydrogénase. La mise en place de telles stratégies par la bactérie afin de lutter contre le manque d'apport énergétique explique l'activation de l'expression d'un arsenal d'enzymes dont des senseurs, des transporteurs ABC ou encore des protéines de réponse au stress.

E. CetR, interagit-il avec des protéines ?

Nous avons montré que CetR se lie à la région codante de l'ARNm *dltD* pour activer la traduction. Ceci est loin des mécanismes de régulation ARNrég-dépendants classiques. L'activation de la traduction d'ARNm cible est habituellement réalisée par deux mécanismes distincts : (i) l'inhibition de la formation de structures secondaires qui séquestrent les RBS à l'extrémité 5' (Frohlich and Vogel, 2009) ; (ii) ou l'augmentation de la stabilité de l'ARNm à l'extrémité 3' (Opdyke *et al.*, 2004). Cette situation est peu susceptible d'être le cas de CetR, parce que le site d'interaction avec l'ARNm *dltD* est loin en aval du RBS. Outre l'appariement de bases avec des régions d'initiation de traduction, les ARNrég bactériens peuvent également favoriser l'inactivation de gènes en ciblant les séquences codantes (Rice *et al.*, 2012). De manière remarquable, en utilisant le programme **IntaRNA**, nous avons prédit une interaction potentielle entre *dltB* et *dltD* sur l'ARNm polycistronique *dltABCD*. Étant donné que cette région d'interaction est située en aval du RBS de l'ARNm *dltD*, nous supposons que ce repliement de l'ARNm est suffisant pour inhiber l'initiation ou l'élongation de la traduction. Ainsi, en empêchant la formation d'une telle structure inhibitrice, la liaison de CetR, en aval du site d'interaction *dltB-dltD*, pourrait permettre le chargement des ribosomes et par conséquent la traduction de DltD. Ainsi, il pourrait être intéressant d'étudier l'effet de la délétion de CetR sur la traduction de DltB.

Existe-t-il une protéine capable de se lier à CetR pour l'assister dans la régulation de l'expression des gènes cibles ? La protéine Hfq ne semble pas impliquée dans la régulation ARNrég-dépendante chez *S. agalactiae*, car aucun homologue à cette protéine chaperonne n'a été identifié à ce jour. Alors, quelles seraient les techniques permettant l'identification

des potentielles protéines impliquées dans la stabilisation du duplex ARNrég-ARNm chez *S. agalactiae* ?

Une des approches pour identifier les potentielles protéines chaperonnes de CetR est basée sur l'utilisation d'un gradient de glycérol pour séparer les protéines qui co-purifient avec les ARNrég d'intérêt. Cette technique permet d'étudier plusieurs ARN dans le même gradient. Si une ou plusieurs protéines sont identifiées, le site de fixation de cette dernière sur l'ARN pourrait être caractérisé en utilisant la technique de HITS-CLIP (*Hight-throughput sequencing of RNA isolated by Crosslinking and Immunoprecipitation*). Le principe de cette technique consiste à lier de manière covalente l'ensemble des ARN interagissant avec une protéine d'intérêt *via* un pontage ultraviolet à 254 nm. Après purification de la protéine d'intérêt par chromatographie d'affinité, le séquençage à haut débit permet d'identifier la liste des ARN liés à cette protéine. Récemment, une approche très sophistiquée a été inventée par l'équipe de Jörg Vogel pour l'identification des protéines chaperonnes et de leurs ARNrég partenaires à partir de la même culture. Cette approche appelée Grad-seq, pour « *Gradient-sequencing* », consiste à séparer les complexes ARNrég-protéine selon leur poids moléculaire dans un gradient de glycérol. Ces complexes sont ensuite séparés dans 20 fractions, dont chaque fraction est analysée à la fois par RNA-seq, pour l'identification de partenaire ARNrég, et par LC-MS/MS pour l'identification de la protéine chaperonne d'intérêt (Smirnov *et al.*, 2016).

Enfin, il serait intéressant d'étudier la régulation de l'ARNrég CetR. Nous avons montré que l'expression de CetR est influencée par la phase de croissance. En outre, nous avons identifié une région homologue entre les régions promotrices de CetR et le régulateur DltR du TCS DltRS. Nous pouvons, par mutagenèse dirigée, étudier le rôle de cette région dans la régulation de CetR par DltR. Un criblage, en condition de stress par exemple, pour identifier d'autres TCS ou des facteurs de transcription serait intéressant afin de comprendre le mécanisme de régulation de l'ARNrég CetR.

F. CetR et la résistance aux antibiotiques : nouvelle cible thérapeutique ?

Plusieurs ARNrég sont connus pour avoir un rôle capital dans la virulence (Boisset *et al.*, 2007; Gong *et al.*, 2011; Chabelskaya *et al.*, 2014; Michaux *et al.*, 2014; Porcheron and

Dozois, 2015). Cependant, peu d'entre eux interfèrent avec la résistance bactérienne aux antibiotiques (Eyraud *et al.*, 2014; Lalaouna *et al.*, 2014; Khan *et al.*, 2016). Nous avons montré, pour la première fois chez les streptocoques, que la délétion d'un ARN rég induit une diminution très importante de la résistance aux PAM chez *S. agalactiae*. La concentration minimale inhibitrice (CMI) de la colistine est supérieure à 512 µg/mL pour la souche sauvage et la souche complémentée alors qu'elle est égale à 2 µg/mL pour le mutant *cetR* (cf, publication 1). A cause de sa toxicité, la colistine est peu utilisée en thérapeutique (voir ci-après). Néanmoins, selon les recommandations de CASFM-EUCAST 2016 (Comité de l'Antibiogramme De La Société Française de Microbiologie- *European Committee on Antimicrobial Susceptibility Testing*), une souche bactérienne est considérée comme sensible à la colistine si la CMI est ≤ 2 µg/mL.

L'émergence de bactéries multi-résistantes qui peuvent être à l'origine d'infections nosocomiales constitue un problème croissant dans le monde entier. Les alternatives thérapeutiques limitées ont entraîné une augmentation de l'utilisation clinique de la colistine, un antibiotique de la famille des polymyxines développé il y a plus de 50 ans et qui a conservé son activité contre un certain nombre de bactéries à Gram négatif multirésistantes. Les polymyxines sont un groupe d'antibiotiques naturels produits par la bactérie *Paenibacillus polymyxa*. Seule la polymyxine E (dénommée colistine) est approuvée pour une utilisation clinique dans l'Union Européenne. Deux formes de la colistine sont utilisées à des fins cliniques : le sulfate de colistine et sa prodrogue microbiologiquement inactive, le colistiméthate de sodium (CMS).

Une des principales voies de résistance à ces antibiotiques est la modification de la charge globale de la membrane cellulaire bactérienne, et plus exactement des ALT (cf revue de la littérature). Ainsi, inhiber l'activité de(s) enzyme(s) impliquée(s) dans cette voie serait une potentielle alternative thérapeutique. Une étude a montré que l'inhibition de la liaison de phosphatidylglycérol à la protéine LtaS (LTA synthase, cf, données bibliographiques) par la molécule 1771 [2-oxo-2-(5-phenyl-1,3,4-oxadiazol-2-ylamino) ethyl 2-naphtho[2,1-b]furan-1-ylacetate] inhibe la biosynthèse des ALT chez *S. aureus* (Richter *et al.*, 2013). La molécule 1771 mime la chaîne polaire du phosphatidylglycérol. En suivant la même stratégie, nous pourrions imaginer la synthèse d'une molécule qui inhibe la liaison de CetR à sa cible *dltD* et

inhiber ainsi la D-alanylation des ALT chez *S. agalactiae* ce qui provoquera la diminution de sa résistance aux PAM.

D'autres stratégies sont en développement pour l'élaboration de nouvelles molécules thérapeutiques, notamment celles visant la substitution des antibiotiques à large spectre longtemps privilégiés par les compagnies pharmaceutiques. Une de ces stratégies consiste à cibler les gènes essentiels à la survie de la bactérie. De nos jours, divers antibiotiques ciblent l'ARN ribosomique bactérien, tels que les aminoglycosides ou les macrolides. De nouvelles approches se basent actuellement sur la synthèse des analogues au métabolite reconnu par les *riboswitch*. La synthèse des analogues à la guanine capables de se lier au *riboswitch* reconnaissant la purine a permis d'inhiber la croissance de *B. subtilis* (Kim *et al.*, 2009). Chez *S. aureus*, un dérivé des pyrimidines, nommé PC1, a été mis au point pour se lier au *riboswitch* senseur des purines. Cet analogue possède une activité antimicrobienne contre un nombre restreint de bactéries (Mulhbachter *et al.*, 2010).

D'autres équipes se sont lancées dans la conception de molécules thérapeutiques capables de perturber les fonctions de la bactérie, tels que les systèmes de sécrétion, les protéines d'adhérence ou encore les facteurs de virulence. Récemment, une molécule, le fluoro-phenyl-styrene-sulfonamide (FPSS), inhibant le facteur de transcription σ^B a été mise en évidence chez *L. monocytogenes* et *B. subtilis*. Le FPSS régule négativement la majorité des gènes (208 gènes) appartenant au régulon σ^B et empêche ainsi l'invasion des cellules épithéliales intestinales (Palmer *et al.*, 2011). Récemment, une classe d'ARNrég artificiels agissant en *trans* (atsRNA) a été développée pour inhiber l'expression de certaines cibles chez *E. coli* (Man *et al.*, 2011). Les atsRNA sont composés de trois domaines distincts : i) une région de liaison à l'ARNm, ii) un site de liaison de la protéine Hfq et iii) un terminateur rho-indépendant (Man *et al.*, 2011). A l'heure actuelle, ces ARNrég artificiels ont montré leur efficacité uniquement chez *E. coli* (Sharma *et al.*, 2012; Park *et al.*, 2013).

Le traitement d'une infection est basé sur l'administration du médicament jusqu'au site de l'infection dans lequel se trouve la cellule bactérienne. La biomolécule d'ARN est sensible à la dégradation par les RNases avant de franchir la membrane bactérienne. Le développement des oligonucléotides synthétiques, tels que les acides nucléiques

peptidiques (PNA) ou les mimétiques d'ADN, appelés aussi phosphorodiamidate morpholino oligomers (PMO), a permis de contourner la dégradation par les RNases mais pas le franchissement de la membrane bactérienne (Dinan and Loftus, 2013). L'efficacité d'absorption des oligonucléotides synthétiques peut être augmentée par la conjugaison des PNA et PMO à de courts peptides cationiques. Ces molécules chargées positivement fonctionnent probablement en co-localisant les oligonucléotides synthétiques avec la membrane externe des bactéries chargée négativement (Dinan and Loftus, 2013).

L'ensemble de ces approches visent à générer des molécules stables, non-toxiques et surtout avec une activité antimicrobienne contre diverses espèces bactériennes. Il est clair que l'utilisation des ARNrég dans l'industrie pharmaceutique devra attendre encore quelques années de recherche et de développement. Cependant, de telles biomolécules peuvent être des agents thérapeutiques utiles pour lutter contre des maladies infectieuses humaines causées par des bactéries à Gram positif multi-résistantes.

CONCLUSION GENERALE

Conclusion générale

La bactérie *S. agalactiae* est reconnue depuis plus de 30 ans comme le principal micro-organisme responsable des infections materno-fœtales. L'infection du nouveau-né par cette bactérie peut être mortelle ou entraîner des séquelles neurologiques dans un nombre non négligeable de cas (1-2% des nouveau-nés infectés) (Verani *et al.*, 2010). Au cours de l'infection, la bactérie doit faire face à de nombreux stress tels que l'acidité de la cavité vaginale ou la réponse des cellules immunitaires. La survie de la bactérie dans cet environnement passe par l'expression et la régulation de nombreux gènes d'adaptation et de virulence. Les mécanismes qui contrôlent l'adaptation de *S. agalactiae* à son environnement sont encore mal connus.

Afin de caractériser les ARNrég de *S. agalactiae*, différentes approches ont été utilisées allant de l'analyse *in silico* à l'analyse transcriptomique et protéomique, en passant par la génétique classique. L'ensemble des résultats obtenus ont permis d'aborder sous un nouvel angle la compréhension de la régulation de l'expression des gènes chez *S. agalactiae*, que ce soit au niveau transcriptionnel ou post-transcriptionnel.

La majorité des résultats de ce travail de thèse présente un grand intérêt pour une meilleure compréhension des réseaux de régulations chez cette bactérie. Nous avons montré que l'ARNrég CetR pourrait être le nouveau régulateur global de virulence chez *S. agalactiae*. Il est ainsi nécessaire de compléter certains résultats obtenus par d'autres études afin d'avoir une vision précise sur les mécanismes de régulations.

Enfin, les résultats obtenus lors de cette thèse sont une base solide pour la recherche de nouvelles molécules antibactériennes à partir de la caractérisation des ARNrég impliqués dans plusieurs fonctions vitales pour *S. agalactiae*. L'ARNrég CetR pourrait faire l'objet d'une première étude de conception de biomolécules thérapeutiques. Il serait également intéressant de croiser les résultats de ce travail avec les autres thématiques de l'équipe, notamment celle portant sur l'étude du métabolisme du carbone et celle sur la caractérisation des éléments génétiques mobiles. En effet, l'ensemble des thématiques développées par l'équipe B.R.M.F (Bactéries et Risque Materno-Fœtales) a pour finalité de comprendre comment *S. agalactiae* s'adapte à son environnement et ainsi passe d'une bactérie commensale à un pathogène opportuniste.

REFERENCES BIBLIOGRAPHIQUES

- Acebo, P., A. J. Martin-Galiano, S. Navarro, A. Zaballos and M. Amblar (2012). "Identification of 88 regulatory small RNAs in the TIGR4 strain of the human pathogen *Streptococcus pneumoniae*." RNA **18**(3): 530-546.
- Adderson, E. E., S. Takahashi, Y. Wang, J. Armstrong, D. V. Miller and J. F. Bohnsack (2003). "Subtractive hybridization identifies a novel predicted protein mediating epithelial cell invasion by virulent serotype III group B *Streptococcus agalactiae*." Infection and immunity **71**(12): 6857-6863.
- Agarwal, V., S. Hammerschmidt, S. Malm, S. Bergmann, K. Riesbeck and A. M. Blom (2012). "Enolase of *Streptococcus pneumoniae* binds human complement inhibitor C4b-binding protein and contributes to complement evasion." Journal of immunology **189**(7): 3575-3584.
- Argaman, L., R. Hershberg, J. Vogel, G. Bejerano, E. G. Wagner, H. Margalit and S. Altuvia (2001). "Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*." Current biology : CB **11**(12): 941-950.
- Arraiano, C. M., J. M. Andrade, S. Domingues, I. B. Guinote, M. Malecki, R. G. Matos, R. N. Moreira, V. Pobre, F. P. Reis, M. Saramago, I. J. Silva and S. C. Viegas (2010). "The critical role of RNA processing and degradation in the control of gene expression." FEMS microbiology reviews **34**(5): 883-923.
- Atilano, M. L., J. Yates, M. Glittenberg, S. R. Filipe and P. Ligoxygakis (2011). "Wall teichoic acids of *Staphylococcus aureus* limit recognition by the drosophila peptidoglycan recognition protein-SA to promote pathogenicity." PLoS pathogens **7**(12): e1002421.
- Babitzke, P. and T. Romeo (2007). "CsrB sRNA family: sequestration of RNA-binding regulatory proteins." Current opinion in microbiology **10**(2): 156-163.
- Baker, C. and M. Edwards (1990). "Group B streptococcal infections, In « Infectious diseases of the foetus and the newborn »." Remington JS, Klein JO (eds), Saunders **3ème**: 742-811.
- Barquist, L. and J. Vogel (2015). "Accelerating Discovery and Functional Analysis of Small RNAs with New Technologies." Annual review of genetics **49**: 367-394.
- Beckmann, C., J. D. Waggoner, T. O. Harris, G. S. Tamura and C. E. Rubens (2002). "Identification of novel adhesins from Group B *streptococci* by use of phage display reveals that C5a peptidase mediates fibronectin binding." Infection and immunity **70**(6): 2869-2876.
- Bergmann, S., M. Rohde, G. S. Chhatwal and S. Hammerschmidt (2001). "alpha-Enolase of *Streptococcus pneumoniae* is a plasmin(ogen)-binding protein displayed on the bacterial cell surface." Molecular microbiology **40**(6): 1273-1287.
- Bergmann, S., H. Schoenen and S. Hammerschmidt (2013). "The interaction between bacterial enolase and plasminogen promotes adherence of *Streptococcus pneumoniae* to epithelial and endothelial cells." International journal of medical microbiology : IJMM **303**(8): 452-462.
- Bishop, C. J., D. M. Aanensen, G. E. Jordan, M. Kilian, W. P. Hanage and B. G. Spratt (2009). "Assigning strains to bacterial species via the internet." BMC biology **7**: 3.
- Biswas, I., A. Gruss, S. D. Ehrlich and E. Maguin (1993). "High-efficiency gene inactivation and replacement system for gram-positive bacteria." J Bacteriol **175**(11): 3628-3635.
- Boisset, S., T. Geissmann, E. Huntzinger, P. Fechter, N. Bendridi, M. Possedko, C. Chevalier, A. C. Helfer, Y. Benito, A. Jacquier, C. Gaspin, F. Vandenesch and P. Romby (2007). "*Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism." Genes & development **21**(11): 1353-1366.

- Bolduc, G. R., M. J. Baron, C. Gravekamp, C. S. Lachenauer and L. C. Madoff (2002). "The alpha C protein mediates internalization of group B *Streptococcus* within human cervical epithelial cells." Cellular microbiology **4**(11): 751-758.
- Boone, T. J. and G. J. Tyrrell (2012). "Identification of the actin and plasminogen binding regions of group B streptococcal phosphoglycerate kinase." The Journal of biological chemistry **287**(34): 29035-29044.
- Bordeau, V., A. Cady, M. Revest, O. Rostan, M. Sassi, P. Tattevin, P. Y. Donnio and B. Felden (2016). "*Staphylococcus aureus* Regulatory RNAs as Potential Biomarkers for Bloodstream Infections." Emerging infectious diseases **22**(9): 1570-1578.
- Braff, M. H., A. L. Jones, S. J. Skerrett and C. E. Rubens (2007). "*Staphylococcus aureus* exploits cathelicidin antimicrobial peptides produced during early pneumonia to promote staphylokinase-dependent fibrinolysis." The Journal of infectious diseases **195**(9): 1365-1372.
- Breaker, R. R. (2012). "Riboswitches and the RNA world." Cold Spring Harbor perspectives in biology **4**(2).
- Brodeur, B. R., M. Boyer, I. Charlebois, J. Hamel, F. Couture, C. R. Rioux and D. Martin (2000). "Identification of group B streptococcal Sip protein, which elicits cross-protective immunity." Infection and immunity **68**(10): 5610-5618.
- Brower, J. V., P. A. Clark, W. Lyon and J. S. Kuo (2014). "MicroRNAs in cancer: glioblastoma and glioblastoma cancer stem cells." Neurochemistry international **77**: 68-77.
- Brown, S., J. P. Santa Maria, Jr. and S. Walker (2013). "Wall teichoic acids of gram-positive bacteria." Annual review of microbiology **67**: 313-336.
- Brown, W. V. and E. B. Collins (1977). "End products and fermentation balances for lactic *streptococci* grown aerobically on low concentrations of glucose." Applied and environmental microbiology **33**(1): 38-42.
- Bryan, J. D., R. Liles, U. Cvek, M. Trutschl and D. Shelper (2008). "Global transcriptional profiling reveals *Streptococcus agalactiae* genes controlled by the MtaR transcription factor." BMC genomics **9**: 607.
- Brzychczy-Wloch, M., T. Gosiewski, M. Bodaszewska, W. Pabian, M. Bulanda, P. Kochan, M. Strus and P. B. Heczko (2010). "Genetic characterization and diversity of *Streptococcus agalactiae* isolates with macrolide resistance." Journal of medical microbiology **59**(Pt 7): 780-786.
- Burnham, C. A., S. E. Shokoples and G. J. Tyrrell (2007). "Invasion of HeLa cells by group B *streptococcus* requires the phosphoinositide-3-kinase signalling pathway and modulates phosphorylation of host-cell Akt and glycogen synthase kinase-3." Microbiology **153**(Pt 12): 4240-4252.
- Bury-Mone, S., Y. Nomane, N. Reymond, R. Barbet, E. Jacquet, S. Imbeaud, A. Jacq and P. Boulloc (2009). "Global analysis of extracytoplasmic stress signaling in *Escherichia coli*." PLoS genetics **5**(9): e1000651.
- Campbell, J., A. K. Singh, J. P. Santa Maria, Jr., Y. Kim, S. Brown, J. G. Swoboda, E. Mylonakis, B. J. Wilkinson and S. Walker (2011). "Synthetic lethal compound combinations reveal a fundamental connection between wall teichoic acid and peptidoglycan biosyntheses in *Staphylococcus aureus*." ACS chemical biology **6**(1): 106-116.
- Campos, M. A., M. A. Vargas, V. Regueiro, C. M. Llompарт, S. Alberti and J. A. Bengoechea (2004). "Capsule polysaccharide mediates bacterial resistance to antimicrobial peptides." Infection and immunity **72**(12): 7107-7114.
- Cayrol, B., E. Fortas, C. Martret, G. Cech, A. Kloska, S. Caulet, M. Barbet, S. Trepout, S. Marco, A. Taghbalout, F. Busi, G. Wegrzyn and V. Arluison (2015). "Riboregulation

- of the bacterial actin-homolog MreB by DsrA small noncoding RNA." Integrative biology : quantitative biosciences from nano to macro **7**(1): 128-141.
- Chabelskaya, S., V. Bordeau and B. Felden (2014). "Dual RNA regulatory control of a *Staphylococcus aureus* virulence factor." Nucleic acids research **42**(8): 4847-4858.
- Chabelskaya, S., O. Gaillot and B. Felden (2010). "A *Staphylococcus aureus* small RNA is required for bacterial virulence and regulates the expression of an immune-evasion molecule." PLoS pathogens **6**(6): e1000927.
- Chapot-Chartier, M. P. and S. Kulakauskas (2014). "Cell wall structure and function in lactic acid bacteria." Microbial cell factories **13 Suppl 1**: S9.
- Cheng, Q., D. Stafslie, S. S. Purushothaman and P. Cleary (2002). "The group B streptococcal C5a peptidase is both a specific protease and an invasin." Infection and immunity **70**(5): 2408-2413.
- Christie, R., N. E. Atkins and E. Munch-Petersen (1944). "A note on a lytic phenomenon shown by group B *streptococci*." Aust. J. Exp. Biol. Med. Sci **22**: 197-200.
- Collins, L. V., S. A. Kristian, C. Weidenmaier, M. Faigle, K. P. Van Kessel, J. A. Van Strijp, F. Gotz, B. Neumeister and A. Peschel (2002). "*Staphylococcus aureus* strains lacking D-alanine modifications of teichoic acids are highly susceptible to human neutrophil killing and are virulence attenuated in mice." The Journal of infectious diseases **186**(2): 214-219.
- Commichau, F. M., C. Herzberg, P. Tripal, O. Valerius and J. Stulke (2007). "A regulatory protein-protein interaction governs glutamate biosynthesis in *Bacillus subtilis*: the glutamate dehydrogenase RocG moonlights in controlling the transcription factor GltC." Molecular microbiology **65**(3): 642-654.
- Cowgill, K., T. H. Taylor, Jr., A. Schuchat and S. Schrag (2003). "Report from the CDC. Awareness of perinatal group B streptococcal infection among women of childbearing age in the United States, 1999 and 2002." Journal of women's health **12**(6): 527-532.
- Da Cunha, V., M. R. Davies, P. E. Douarre, I. Rosinski-Chupin, I. Margarit, S. Spinali, T. Perkins, P. Lechat, N. Dmytruk, E. Sauvage, L. Ma, B. Romi, M. Tichit, M. J. Lopez-Sanchez, S. Descorps-Declere, E. Souche, C. Buchrieser, P. Trieu-Cuot, I. Moszer, D. Clermont, D. Maione, C. Bouchier, D. J. McMillan, J. Parkhill, J. L. Telford, G. Dougan, M. J. Walker, M. T. Holden, C. Poyart and P. Glaser (2014). "*Streptococcus agalactiae* clones infecting humans were selected and fixed through the extensive use of tetracycline." Nature communications **5**: 4544.
- Dathe, M., H. Nikolenko, J. Meyer, M. Beyermann and M. Bienert (2001). "Optimization of the antimicrobial activity of magainin peptides by modification of charge." FEBS letters **501**(2-3): 146-150.
- Debabov, D. V., M. P. Heaton, Q. Zhang, K. D. Stewart, R. H. Lambalot and F. C. Neuhaus (1996). "The D-Alanyl carrier protein in *Lactobacillus casei*: cloning, sequencing, and expression of dltC." Journal of bacteriology **178**(13): 3869-3876.
- Diamond, G., N. Beckloff, A. Weinberg and K. O. Kisich (2009). "The roles of antimicrobial peptides in innate host defense." Current pharmaceutical design **15**(21): 2377-2392.
- Dinan, A. M. and B. J. Loftus (2013). "(Non-)translational medicine: targeting bacterial RNA." Frontiers in genetics **4**: 230.
- Djabri, B., N. Bareille, F. Beaudeau and H. Seegers (2002). "Quarter milk somatic cell count in infected dairy cows: a meta-analysis." Veterinary research **33**(4): 335-357.
- Durand, S., A. Tomasini, F. Braun, C. Condon and P. Romby (2015). "sRNA and mRNA turnover in Gram-positive bacteria." FEMS microbiology reviews **39**(3): 316-330.
- Eberhard, T., G. Kronvall and M. Ullberg (1999). "Surface bound plasmin promotes migration of *Streptococcus pneumoniae* through reconstituted basement membranes." Microbial pathogenesis **26**(3): 175-181.

- Edwards, M. S. and C. J. Baker (2005). "Group B streptococcal infections in elderly adults." Clinical infectious diseases : an official publication of the Infectious Diseases Society of America **41**(6): 839-847.
- Egan, A. J., J. Biboy, I. van't Veer, E. Breukink and W. Vollmer (2015). "Activities and regulation of peptidoglycan synthases." Philosophical transactions of the Royal Society of London. Series B, Biological sciences **370**(1679).
- Eickhoff, T. C., J. O. Klein, A. K. Daly, D. Ingall and M. Finland (1964). "Neonatal Sepsis and Other Infections Due to Group B Beta-Hemolytic *Streptococci*." The New England journal of medicine **271**: 1221-1228.
- Elliott, J. A., R. R. Facklam and C. B. Richter (1990). "Whole-cell protein patterns of nonhemolytic group B, type Ib, *streptococci* isolated from humans, mice, cattle, frogs, and fish." Journal of clinical microbiology **28**(3): 628-630.
- Evans, J. J., D. J. Pasnik, P. H. Klesius and S. Al-Ablani (2006). "First report of *Streptococcus agalactiae* and *Lactococcus garvieae* from a wild bottlenose dolphin (*Tursiops truncatus*)." Journal of wildlife diseases **42**(3): 561-569.
- Eyraud, A., P. Tattevin, S. Chabelskaya and B. Felden (2014). "A small RNA controls a protein regulator involved in antibiotic resistance in *Staphylococcus aureus*." Nucleic acids research **42**(8): 4892-4905.
- Felden, B., F. Vandenesch, P. Boulouc and P. Romby (2011). "The *Staphylococcus aureus* RNome and its commitment to virulence." PLoS pathogens **7**(3): e1002006.
- Fernie-King, B. A., D. J. Seilly, A. Davies and P. J. Lachmann (2002). "Streptococcal inhibitor of complement inhibits two additional components of the mucosal innate immune system: secretory leukocyte proteinase inhibitor and lysozyme." Infection and immunity **70**(9): 4908-4916.
- Fischer, W. (1994). "Lipoteichoic acid and lipids in the membrane of *Staphylococcus aureus*." Medical microbiology and immunology **183**(2): 61-76.
- Floto, R. A., P. A. MacAry, J. M. Boname, T. S. Mien, B. Kampmann, J. R. Hair, O. S. Huey, E. N. Houben, J. Pieters, C. Day, W. Oehlmann, M. Singh, K. G. Smith and P. J. Lehner (2006). "Dendritic cell stimulation by mycobacterial Hsp70 is mediated through CCR5." Science **314**(5798): 454-458.
- Fogg, G. C., C. M. Gibson and M. G. Caparon (1994). "The identification of *rofA*, a positive-acting regulatory component of *prtF* expression: use of an m gamma delta-based shuttle mutagenesis strategy in *Streptococcus pyogenes*." Molecular microbiology **11**(4): 671-684.
- Forte, E., V. B. Borisov, M. Falabella, H. G. Colaco, M. Tinajero-Trejo, R. K. Poole, J. B. Vicente, P. Sarti and A. Giuffrè (2016). "The Terminal Oxidase Cytochrome bd Promotes Sulfide-resistant Bacterial Respiration and Growth." Scientific reports **6**: 23788.
- Fozo, E. M., M. R. Hemm and G. Storz (2008). "Small toxic proteins and the antisense RNAs that repress them." Microbiology and molecular biology reviews : MMBR **72**(4): 579-589, Table of Contents.
- Francis, M. (2013). "Position phylogénétique des souches de *Streptococcus agalactiae* d'origine pulmonaire." Thèse - Université de Tours.
- Froehlich, B. J., C. Bates and J. R. Scott (2009). "*Streptococcus pyogenes* CovRS mediates growth in iron starvation and in the presence of the human cationic antimicrobial peptide LL-37." Journal of bacteriology **191**(2): 673-677.
- Froehlich, K. S. and J. Vogel (2009). "Activation of gene expression by small RNA." Current opinion in microbiology **12**(6): 674-682.
- Gallo, R. L., K. J. Kim, M. Bernfield, C. A. Kozak, M. Zanetti, L. Merluzzi and R. Gennaro (1997). "Identification of CRAMP, a cathelin-related antimicrobial peptide expressed

- in the embryonic and adult mouse." The Journal of biological chemistry **272**(20): 13088-13093.
- Gennaro, R. and M. Zanetti (2000). "Structural features and biological activities of the cathelicidin-derived antimicrobial peptides." Biopolymers **55**(1): 31-49.
- Giammarinaro, P., M. Sicard and A. M. Gasc (1999). "Genetic and physiological studies of the CiaH-CiaR two-component signal-transducing system involved in cefotaxime resistance and competence of *Streptococcus pneumoniae*." Microbiology **145** (Pt 8): 1859-1869.
- Giard, J. C., A. Rince, H. Capiiaux, Y. Auffray and A. Hartke (2000). "Inactivation of the stress- and starvation-inducible *gls24* operon has a pleiotrophic effect on cell morphology, stress sensitivity, and gene expression in *Enterococcus faecalis*." Journal of bacteriology **182**(16): 4512-4520.
- Gibson, D. G., L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison, 3rd and H. O. Smith (2009). "Enzymatic assembly of DNA molecules up to several hundred kilobases." Nat Methods **6**(5): 343-345.
- Gibson, R. L., V. Nizet and C. E. Rubens (1999). "Group B streptococcal beta-hemolysin promotes injury of lung microvascular endothelial cells." Pediatric research **45**(5 Pt 1): 626-634.
- Gillrie, M. R., L. Zbytnuik, E. McAvoy, R. Kapadia, K. Lee, C. C. Waterhouse, S. P. Davis, D. A. Muruve, P. Kubes and M. Ho (2010). "Divergent roles of Toll-like receptor 2 in response to lipoteichoic acid and *Staphylococcus aureus* in vivo." European journal of immunology **40**(6): 1639-1650.
- Glaser, P., C. Rusniok, C. Buchrieser, F. Chevalier, L. Frangeul, T. Msadek, M. Zouine, E. Couve, L. Lalioui, C. Poyart, P. Trieu-Cuot and F. Kunst (2002). "Genome sequence of *Streptococcus agalactiae*, a pathogen causing invasive neonatal disease." Mol Microbiol **45**(6): 1499-1513.
- Gogol, E. B., V. A. Rhodius, K. Papenfort, J. Vogel and C. A. Gross (2011). "Small RNAs endow a transcriptional activator with essential repressor functions for single-tier control of a global stress regulon." Proceedings of the National Academy of Sciences of the United States of America **108**(31): 12875-12880.
- Gong, H., G. P. Vu, Y. Bai, E. Chan, R. Wu, E. Yang, F. Liu and S. Lu (2011). "A *Salmonella* small non-coding RNA facilitates bacterial invasion and intracellular replication by modulating the expression of virulence factors." PLoS pathogens **7**(9): e1002120.
- Gossringer, M. and R. K. Hartmann (2012). "3'-UTRs as a source of regulatory RNAs in bacteria." The EMBO journal **31**(20): 3958-3960.
- Gottesman, S. (2005). "Micros for microbes: non-coding regulatory RNAs in bacteria." Trends in genetics : TIG **21**(7): 399-404.
- Gottesman, S. and G. Storz (2011). "Bacterial small RNA regulators: versatile roles and rapidly evolving variations." Cold Spring Harbor perspectives in biology **3**(12).
- Graham, M. R., L. M. Smoot, C. A. Migliaccio, K. Virtaneva, D. E. Sturdevant, S. F. Porcella, M. J. Federle, G. J. Adams, J. R. Scott and J. M. Musser (2002). "Virulence control in group A *Streptococcus* by a two-component gene regulatory system: global expression profiling and in vivo infection modeling." Proceedings of the National Academy of Sciences of the United States of America **99**(21): 13855-13860.
- Granato, D., G. E. Bergonzelli, R. D. Pridmore, L. Marvin, M. Rouvet and I. E. Corthesy-Theulaz (2004). "Cell surface-associated elongation factor Tu mediates the attachment of *Lactobacillus johnsonii* NCC533 (La1) to human intestinal cells and mucins." Infection and immunity **72**(4): 2160-2169.
- Green, N. J., F. J. Grundy and T. M. Henkin (2010). "The T box mechanism: tRNA as a regulatory molecule." FEBS letters **584**(2): 318-324.

- Grundling, A. and O. Schneewind (2007). "Synthesis of glycerol phosphate lipoteichoic acid in *Staphylococcus aureus*." Proceedings of the National Academy of Sciences of the United States of America **104**(20): 8478-8483.
- Guenzi, E., A. M. Gasc, M. A. Sicard and R. Hakenbeck (1994). "A two-component signal-transducing system is involved in competence and penicillin susceptibility in laboratory mutants of *Streptococcus pneumoniae*." Molecular microbiology **12**(3): 505-515.
- Guerrier-Takada, C., K. Gardiner, T. Marsh, N. Pace and S. Altman (1983). "The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme." Cell **35**(3 Pt 2): 849-857.
- Gutekunst, H., B. J. Eikmanns and D. J. Reinscheid (2003). "Analysis of RogB-controlled virulence mechanisms and gene repression in *Streptococcus agalactiae*." Infection and immunity **71**(9): 5056-5064.
- Gutekunst, H., B. J. Eikmanns and D. J. Reinscheid (2004). "The novel fibrinogen-binding protein FbsB promotes *Streptococcus agalactiae* invasion into epithelial cells." Infection and immunity **72**(6): 3495-3504.
- Hamilton, A., D. L. Popham, D. J. Carl, X. Lauth, V. Nizet and A. L. Jones (2006). "Penicillin-binding protein 1a promotes resistance of group B *streptococcus* to antimicrobial peptides." Infection and immunity **74**(11): 6179-6187.
- Hauser, A. R., D. L. Stevens, E. L. Kaplan and P. M. Schlievert (1991). "Molecular analysis of pyrogenic exotoxins from *Streptococcus pyogenes* isolates associated with toxic shock-like syndrome." Journal of clinical microbiology **29**(8): 1562-1567.
- Henderson, B. and A. Martin (2011). "Bacterial virulence in the moonlight: multitasking bacterial moonlighting proteins are virulence determinants in infectious disease." Infection and immunity **79**(9): 3476-3491.
- Henderson, B. and A. Martin (2013). "Bacterial moonlighting proteins and bacterial virulence." Current topics in microbiology and immunology **358**: 155-213.
- Hillerton, J. E. and E. A. Berry (2003). "The management and treatment of environmental *Streptococcal mastitis*." The Veterinary clinics of North America. Food animal practice **19**(1): 157-169.
- Hindley, J. (1967). "Fractionation of 32P-labelled ribonucleic acids on polyacrylamide gels and their characterization by fingerprinting." Journal of molecular biology **30**(1): 125-136.
- Hughes, M. J., J. C. Moore, J. D. Lane, R. Wilson, P. K. Pribul, Z. N. Younes, R. J. Dobson, P. Everest, A. J. Reason, J. M. Redfern, F. M. Greer, T. Paxton, M. Panico, H. R. Morris, R. G. Feldman and J. D. Santangelo (2002). "Identification of major outer surface proteins of *Streptococcus agalactiae*." Infection and immunity **70**(3): 1254-1259.
- Innocenti, N., M. Golumbeanu, A. Fouquier d'Herouel, C. Lacoux, R. A. Bonnin, S. P. Kennedy, F. Wessner, P. Serror, P. Boulloc, F. Repoila and E. Aurell (2015). "Whole-genome mapping of 5' RNA ends in bacteria by tagged sequencing: a comprehensive view in *Enterococcus faecalis*." RNA **21**(5): 1018-1030.
- Jagadeesan, B., O. K. Koo, K. P. Kim, K. M. Burkholder, K. K. Mishra, A. Aroonnuan and A. K. Bhunia (2010). "LAP, an alcohol acetaldehyde dehydrogenase enzyme in *Listeria*, promotes bacterial adhesion to enterocyte-like Caco-2 cells only in pathogenic species." Microbiology **156**(Pt 9): 2782-2795.
- Jeffery, C. J. (2005). "Mass spectrometry and the search for moonlighting proteins." Mass spectrometry reviews **24**(6): 772-782.
- Jerlstrom, P. G., S. R. Talay, P. Valentin-Weigand, K. N. Timmis and G. S. Chhatwal (1996). "Identification of an immunoglobulin A binding motif located in the beta-antigen of

- the c protein complex of group B *streptococci*." Infection and immunity **64**(7): 2787-2793.
- Jester, B. C., P. Romby and E. Lioliou (2012). "When ribonucleases come into play in pathogens: a survey of gram-positive bacteria." International journal of microbiology **2012**: 592196.
- Jiang, S. M., M. J. Cieslewicz, D. L. Kasper and M. R. Wessels (2005). "Regulation of virulence by a two-component system in group B *streptococcus*." Journal of bacteriology **187**(3): 1105-1113.
- Jiang, S. M., N. Ishmael, J. Dunning Hotopp, M. Puliti, L. Tissi, N. Kumar, M. J. Cieslewicz, H. Tettelin and M. R. Wessels (2008). "Variation in the group B *Streptococcus* CsrRS regulon and effects on pathogenicity." Journal of bacteriology **190**(6): 1956-1965.
- Jin, T., M. Bokarewa, T. Foster, J. Mitchell, J. Higgins and A. Tarkowski (2004). "*Staphylococcus aureus* resists human defensins by production of staphylokinase, a novel bacterial evasion mechanism." Journal of immunology **172**(2): 1169-1176.
- Johansson, J., P. Mandin, A. Renzoni, C. Chiaruttini, M. Springer and P. Cossart (2002). "An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*." Cell **110**(5): 551-561.
- Jones, N., J. F. Bohnsack, S. Takahashi, K. A. Oliver, M. S. Chan, F. Kunst, P. Glaser, C. Rusniok, D. W. Crook, R. M. Harding, N. Bisharat and B. G. Spratt (2003). "Multilocus sequence typing system for group B *streptococcus*." Journal of clinical microbiology **41**(6): 2530-2536.
- Joubrel, C., A. Tazi, A. Six, N. Dmytruk, G. Touak, P. Bidet, J. Raymond, P. Trieu Cuot, A. Fouet, S. Kerneis and C. Poyart (2015). "Group B *streptococcus* neonatal invasive infections, France 2007-2012." Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases **21**(10): 910-916.
- Kallman, J., J. Schollin, C. Schalen, A. Erlandsson and E. Kihlstrom (1998). "Impaired phagocytosis and opsonisation towards group B *streptococci* in preterm neonates." Archives of disease in childhood. Fetal and neonatal edition **78**(1): F46-50.
- Karvelis, T., G. Gasiunas, A. Miksys, R. Barrangou, P. Horvath and V. Siksnys (2013). "crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*." RNA biology **10**(5): 841-851.
- Kasahara, K., A. J. Baltus, S. H. Lee, M. A. Edelstein and P. H. Edelstein (2010). "Prevalence of non-penicillin-susceptible group B *streptococcus* in Philadelphia and specificity of penicillin resistance screening methods." Journal of clinical microbiology **48**(4): 1468-1469.
- Katakura, Y., R. Sano, T. Hashimoto, K. Ninomiya and S. Shioya (2010). "Lactic acid bacteria display on the cell surface cytosolic proteins that recognize yeast mannan." Applied microbiology and biotechnology **86**(1): 319-326.
- Kawamoto, H., Y. Koide, T. Morita and H. Aiba (2006). "Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq." Molecular microbiology **61**(4): 1013-1022.
- Kawano, M., L. Aravind and G. Storz (2007). "An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin." Molecular microbiology **64**(3): 738-754.
- Keefe, G. P. (1997). "*Streptococcus agalactiae* mastitis: a review." The Canadian veterinary journal. La revue veterinaire canadienne **38**(7): 429-437.
- Kesimer, M., N. Kilic, R. Mehrotra, D. J. Thornton and J. K. Sheehan (2009). "Identification of salivary mucin MUC7 binding proteins from *Streptococcus gordonii*." BMC microbiology **9**: 163.

- Khan, M. A., Y. Gopel, S. Milewski and B. Gorke (2016). "Two Small RNAs Conserved in *Enterobacteriaceae* Provide Intrinsic Resistance to Antibiotics Targeting the Cell Wall Biosynthesis Enzyme Glucosamine-6-Phosphate Synthase." Frontiers in microbiology **7**: 908.
- Kilian, M., K. Poulsen, T. Blomqvist, L. S. Havarstein, M. Bek-Thomsen, H. Tettelin and U. B. Sorensen (2008). "Evolution of *Streptococcus pneumoniae* and its close commensal relatives." PloS one **3**(7): e2683.
- Kim, J. N., K. F. Blount, I. Puskarz, J. Lim, K. H. Link and R. R. Breaker (2009). "Design and antimicrobial action of purine analogues that bind Guanine riboswitches." ACS chemical biology **4**(11): 915-927.
- Konto-Ghiorghi, Y., E. Mairey, A. Mallet, G. Dumenil, E. Caliot, P. Trieu-Cuot and S. Dramsi (2009). "Dual role for pilus in adherence to epithelial cells and biofilm formation in *Streptococcus agalactiae*." PLoS pathogens **5**(5): e1000422.
- Kreikemeyer, B., M. D. Boyle, B. A. Buttaro, M. Heinemann and A. Podbielski (2001). "Group A streptococcal growth phase-associated virulence factor regulation by a novel operon (Fas) with homologies to two-component-type regulators requires a small RNA molecule." Molecular microbiology **39**(2): 392-406.
- Kreikemeyer, B., K. S. McIver and A. Podbielski (2003). "Virulence factor regulation and regulatory networks in *Streptococcus pyogenes* and their impact on pathogen-host interactions." Trends in microbiology **11**(5): 224-232.
- Krohn, M. A., S. L. Hillier and C. J. Baker (1999). "Maternal peripartum complications associated with vaginal group B *streptococci* colonization." The Journal of infectious diseases **179**(6): 1410-1415.
- Kumar, R., P. Shah, E. Swiatlo, S. C. Burgess, M. L. Lawrence and B. Nanduri (2010). "Identification of novel non-coding small RNAs from *Streptococcus pneumoniae* TIGR4 using high-resolution genome tiling arrays." BMC genomics **11**: 350.
- Lalaouna, D., A. Eyraud, S. Chabelskaya, B. Felden and E. Masse (2014). "Regulatory RNAs involved in bacterial antibiotic resistance." PLoS Pathog **10**(8): e1004299.
- Lamy, M. C., M. Zouine, J. Fert, M. Vergassola, E. Couve, E. Pellegrini, P. Glaser, F. Kunst, T. Msadek, P. Trieu-Cuot and C. Poyart (2004). "CovS/CovR of group B *streptococcus*: a two-component global regulatory system involved in virulence." Mol Microbiol **54**(5): 1250-1268.
- Lamy, M. C., M. Zouine, J. Fert, M. Vergassola, E. Couve, E. Pellegrini, P. Glaser, F. Kunst, T. Msadek, P. Trieu-Cuot and C. Poyart (2004). "CovS/CovR of group B *streptococcus*: a two-component global regulatory system involved in virulence." Molecular microbiology **54**(5): 1250-1268.
- Lancefield, R. C. (1933). "A Serological Differentiation of Human and Other Groups of Hemolytic *Streptococci*." The Journal of experimental medicine **57**(4): 571-595.
- Lancefield, R. C. and E. H. Freimer (1966). "Type-specific polysaccharide antigens of group B *streptococci*." The Journal of hygiene **64**(2): 191-203.
- Lang, S. and M. Palmer (2003). "Characterization of *Streptococcus agalactiae* CAMP factor as a pore-forming toxin." The Journal of biological chemistry **278**(40): 38167-38173.
- Lartigue, M. F. and P. Boulloc (2014). "A tetracycline-inducible expression vector for *Streptococcus agalactiae* allowing controllable gene expression." Journal of microbiological methods **96**: 16-18.
- Lauth, X., M. von Kockritz-Blickwede, C. W. McNamara, S. Myskowski, A. S. Zinkernagel, B. Beall, P. Ghosh, R. L. Gallo and V. Nizet (2009). "M1 protein allows Group A streptococcal survival in phagocyte extracellular traps through cathelicidin inhibition." Journal of innate immunity **1**(3): 202-214.

- Lays, C., C. Romilly, A. Tomasini, I. Caldelari, Y. Benito, P. Hammann, T. Geissmann, S. Boisset, P. Romby and F. Vandenesch (2014). "[RsaA non-coding RNA promotes bacterial persistence and attenuates virulence in *Staphylococcus aureus*]." Medecine sciences : M/S **30**(10): 839-841.
- Le Doare, K. and P. T. Heath (2013). "An overview of global GBS epidemiology." Vaccine **31 Suppl 4**: D7-12.
- Le Rhun, A. and E. Charpentier (2012). "Small RNAs in *streptococci*." RNA biology **9**(4): 414-426.
- Lehner, T., L. A. Bergmeier, Y. Wang, L. Tao, M. Sing, R. Spallek and R. van der Zee (2000). "Heat shock proteins generate beta-chemokines which function as innate adjuvants enhancing adaptive immunity." European journal of immunology **30**(2): 594-603.
- Levin, J. C. and M. R. Wessels (1998). "Identification of *csrR/csrS*, a genetic locus that regulates hyaluronic acid capsule synthesis in group A *Streptococcus*." Molecular microbiology **30**(1): 209-219.
- Levy, S. B. and B. Marshall (2004). "Antibacterial resistance worldwide: causes, challenges and responses." Nature medicine **10**(12 Suppl): S122-129.
- Lindahl, G., M. Stalhammar-Carlemalm and T. Areschoug (2005). "Surface proteins of *Streptococcus agalactiae* and related proteins in other bacterial pathogens." Clinical microbiology reviews **18**(1): 102-127.
- Liu, G. Y., K. S. Doran, T. Lawrence, N. Turkson, M. Puliti, L. Tissi and V. Nizet (2004). "Sword and shield: linked group B streptococcal beta-hemolysin/cytolysin and carotenoid pigment function to subvert host phagocyte defense." Proceedings of the National Academy of Sciences of the United States of America **101**(40): 14491-14496.
- Liu, M. Y., G. Gui, B. Wei, J. F. Preston, 3rd, L. Oakford, U. Yuksel, D. P. Giedroc and T. Romeo (1997). "The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in *Escherichia coli*." The Journal of biological chemistry **272**(28): 17502-17510.
- Liu, Z., J. Trevino, E. Ramirez-Pena and P. Sumbly (2012). "The small regulatory RNA FasX controls pilus expression and adherence in the human bacterial pathogen group A *Streptococcus*." Molecular microbiology **86**(1): 140-154.
- Livny, J., H. Teonadi, M. Livny and M. K. Waldor (2008). "High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs." PloS one **3**(9): e3197.
- Llobet, E., J. M. Tomas and J. A. Bengoechea (2008). "Capsule polysaccharide is a bacterial decoy for antimicrobial peptides." Microbiology **154**(Pt 12): 3877-3886.
- Loh, E., O. Dussurget, J. Gripenland, K. Vaitkevicius, T. Tiensuu, P. Mandin, F. Repoila, C. Buchrieser, P. Cossart and J. Johansson (2009). "A trans-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*." Cell **139**(4): 770-779.
- Lukiw, W. J., P. N. Alexandrov, Y. Zhao, J. M. Hill and S. Bhattacharjee (2012). "Spreading of Alzheimer's disease inflammatory signaling through soluble micro-RNA." Neuroreport **23**(10): 621-626.
- Madureira, P., M. Baptista, M. Vieira, V. Magalhaes, A. Camelo, L. Oliveira, A. Ribeiro, D. Tavares, P. Trieu-Cuot, M. Vilanova and P. Ferreira (2007). "*Streptococcus agalactiae* GAPDH is a virulence-associated immunomodulatory protein." Journal of immunology **178**(3): 1379-1387.
- Maiden, M. C., J. A. Bygraves, E. Feil, G. Morelli, J. E. Russell, R. Urwin, Q. Zhang, J. Zhou, K. Zurth, D. A. Caugant, I. M. Feavers, M. Achtman and B. G. Spratt (1998). "Multilocus sequence typing: a portable approach to the identification of clones within

- populations of pathogenic microorganisms." Proceedings of the National Academy of Sciences of the United States of America **95**(6): 3140-3145.
- Maisey, H. C., K. S. Doran and V. Nizet (2008). "Recent advances in understanding the molecular basis of group B *Streptococcus* virulence." Expert Rev Mol Med **10**: e27.
- Maisey, H. C., D. Quach, M. E. Hensler, G. Y. Liu, R. L. Gallo, V. Nizet and K. S. Doran (2008). "A group B streptococcal pilus protein promotes phagocyte resistance and systemic virulence." FASEB journal : official publication of the Federation of American Societies for Experimental Biology **22**(6): 1715-1724.
- Majdalani, N., C. Cuning, D. Sledjeski, T. Elliott and S. Gottesman (1998). "DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription." Proceedings of the National Academy of Sciences of the United States of America **95**(21): 12462-12467.
- Man, S., R. Cheng, C. Miao, Q. Gong, Y. Gu, X. Lu, F. Han and W. Yu (2011). "Artificial trans-encoded small non-coding RNAs specifically silence the selected gene expression in bacteria." Nucleic acids research **39**(8): e50.
- Mandal, M., M. Lee, J. E. Barrick, Z. Weinberg, G. M. Emilsson, W. L. Ruzzo and R. R. Breaker (2004). "A glycine-dependent riboswitch that uses cooperative binding to control gene expression." Science **306**(5694): 275-279.
- Mangold, M., M. Siller, B. Roppenser, B. J. Vlamincx, T. A. Penfound, R. Klein, R. Novak, R. P. Novick and E. Charpentier (2004). "Synthesis of group A streptococcal virulence factors is controlled by a regulatory RNA molecule." Molecular microbiology **53**(5): 1515-1527.
- Mann, B., T. van Opijnen, J. Wang, C. Obert, Y. D. Wang, R. Carter, D. J. McGoldrick, G. Ridout, A. Camilli, E. I. Tuomanen and J. W. Rosch (2012). "Control of virulence by small RNAs in *Streptococcus pneumoniae*." PLoS pathogens **8**(7): e1002788.
- Marx, P., M. Nuhn, M. Kovacs, R. Hakenbeck and R. Bruckner (2010). "Identification of genes for small non-coding RNAs that belong to the regulon of the two-component regulatory system CiaRH in *Streptococcus*." BMC genomics **11**: 661.
- Masse, E., C. K. Vanderpool and S. Gottesman (2005). "Effect of RyhB small RNA on global iron use in *Escherichia coli*." Journal of bacteriology **187**(20): 6962-6971.
- Mauri, M. and S. Klumpp (2014). "A model for sigma factor competition in bacterial cells." PLoS computational biology **10**(10): e1003845.
- Meneveau, N., F. Schiele, A. Vuilleminot, B. Valette, G. Grollier, Y. Bernard and J. P. Bassand (1997). "Streptokinase vs alteplase in massive pulmonary embolism. A randomized trial assessing right heart haemodynamics and pulmonary vascular obstruction." European heart journal **18**(7): 1141-1148.
- Michaux, C., A. Hartke, C. Martini, S. Reiss, D. Albrecht, A. Budin-Verneuil, M. Sanguinetti, S. Engelmann, T. Hain, N. Verneuil and J. C. Giard (2014). "Involvement of *Enterococcus faecalis* small RNAs in stress response and virulence." Infection and immunity **82**(9): 3599-3611.
- Miller, J. H. (1972). "Experiments in molecular genetics " Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Mizuno, T., M. Y. Chou and M. Inouye (1984). "A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA)." Proceedings of the National Academy of Sciences of the United States of America **81**(7): 1966-1970.
- Moll, I., T. Afonyushkin, O. Vytvytska, V. R. Kaberdin and U. Blasi (2003). "Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs." RNA **9**(11): 1308-1314.

- Moller, T., T. Franch, P. Hojrup, D. R. Keene, H. P. Bachinger, R. G. Brennan and P. Valentin-Hansen (2002). "Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction." Molecular cell **9**(1): 23-30.
- Morello, E., A. Mallet, Y. Konto-Ghiorgi, T. Chaze, M. Y. Mistou, G. Oliva, L. Oliveira, A. M. Di Guilmi, P. Trieu-Cuot and S. Dramsi (2015). "Evidence for the Sialylation of PilA, the PI-2a Pilus-Associated Adhesin of *Streptococcus agalactiae* Strain NEM316." PloS one **10**(9): e0138103.
- Morfeldt, E., D. Taylor, A. von Gabain and S. Arvidson (1995). "Activation of alpha-toxin translation in *Staphylococcus aureus* by the trans-encoded antisense RNA, RNAIII." The EMBO journal **14**(18): 4569-4577.
- Mori, Y., M. Yamaguchi, Y. Terao, S. Hamada, T. Ooshima and S. Kawabata (2012). "alpha-Enolase of *Streptococcus pneumoniae* induces formation of neutrophil extracellular traps." The Journal of biological chemistry **287**(13): 10472-10481.
- Morita, T., K. Maki and H. Aiba (2005). "RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs." Genes & development **19**(18): 2176-2186.
- Morris, K. V. and J. S. Mattick (2014). "The rise of regulatory RNA." Nature reviews. Genetics **15**(6): 423-437.
- Mulhbachter, J., E. Brouillette, M. Allard, L. C. Fortier, F. Malouin and D. A. Lafontaine (2010). "Novel riboswitch ligand analogs as selective inhibitors of guanine-related metabolic pathways." PLoS pathogens **6**(4): e1000865.
- Nawrocki, K. L., E. K. Crispell and S. M. McBride (2014). "Antimicrobial Peptide Resistance Mechanisms of Gram-Positive Bacteria." Antibiotics **3**(4): 461-492.
- Nealon, T. J. and S. J. Mattingly (1984). "Role of cellular lipoteichoic acids in mediating adherence of serotype III strains of group B *streptococci* to human embryonic, fetal, and adult epithelial cells." Infection and immunity **43**(2): 523-530.
- Neuhaus, F. C. and J. Baddiley (2003). "A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in gram-positive bacteria." Microbiology and molecular biology reviews : MMBR **67**(4): 686-723.
- Nobbs, A. H., R. J. Lamont and H. F. Jenkinson (2009). "*Streptococcus* adherence and colonization." Microbiology and molecular biology reviews : MMBR **73**(3): 407-450, Table of Contents.
- Nocard, E. and H. Mollereau (1887). "Sur une mammite contagieuse des vaches laitieres." Ann Inst Pasteur(1): 109.
- Nordstrom, K. (2006). "Plasmid R1--replication and its control." Plasmid **55**(1): 1-26.
- O'Brien, K. L., L. J. Wolfson, J. P. Watt, E. Henkle, M. Deloria-Knoll, N. McCall, E. Lee, K. Mulholland, O. S. Levine and T. Cherian (2009). "Burden of disease caused by *Streptococcus pneumoniae* in children younger than 5 years: global estimates." Lancet **374**(9693): 893-902.
- Obana, N., Y. Shirahama, K. Abe and K. Nakamura (2010). "Stabilization of *Clostridium perfringens* collagenase mRNA by VR-RNA-dependent cleavage in 5' leader sequence." Molecular microbiology **77**(6): 1416-1428.
- Oku, Y., K. Kurokawa, M. Matsuo, S. Yamada, B. L. Lee and K. Sekimizu (2009). "Pleiotropic roles of polyglycerolphosphate synthase of lipoteichoic acid in growth of *Staphylococcus aureus* cells." Journal of bacteriology **191**(1): 141-151.
- Opdyke, J. A., J. G. Kang and G. Storz (2004). "GadY, a small-RNA regulator of acid response genes in *Escherichia coli*." Journal of bacteriology **186**(20): 6698-6705.
- Palmer, M. E., S. Chaturongakul, M. Wiedmann and K. J. Boor (2011). "The *Listeria monocytogenes* sigmaB regulon and its virulence-associated functions are inhibited by a small molecule." mBio **2**(6).

- Pancholi, V. and V. A. Fischetti (1998). "alpha-enolase, a novel strong plasmin(ogen) binding protein on the surface of pathogenic *streptococci*." The Journal of biological chemistry **273**(23): 14503-14515.
- Papasergi, S., S. Brega, M. Y. Mistou, A. Firon, V. Oxaran, R. Dover, G. Teti, Y. Shai, P. Trieu-Cuot and S. Dramsi (2011). "The GBS PI-2a pilus is required for virulence in mice neonates." PloS one **6**(4): e18747.
- Papasergi, S., R. Galbo, V. Lanza-Cariccio, M. Domina, G. Signorino, C. Biondo, I. Pernice, C. Poyart, P. Trieu-Cuot, G. Teti and C. Beninati (2013). "Analysis of the *Streptococcus agalactiae* exoproteome." Journal of proteomics **89**: 154-164.
- Papenfort, K. and J. Vogel (2010). "Regulatory RNA in bacterial pathogens." Cell host & microbe **8**(1): 116-127.
- Park, H., G. Bak, S. C. Kim and Y. Lee (2013). "Exploring sRNA-mediated gene silencing mechanisms using artificial small RNAs derived from a natural RNA scaffold in *Escherichia coli*." Nucleic acids research **41**(6): 3787-3804.
- Pasquina, L. W., J. P. Santa Maria and S. Walker (2013). "Teichoic acid biosynthesis as an antibiotic target." Current opinion in microbiology **16**(5): 531-537.
- Patenge, N., A. Billion, P. Raasch, J. Normann, A. Wisniewska-Kucper, J. Retey, V. Boisguerin, T. Hartsch, T. Hain and B. Kreikemeyer (2012). "Identification of novel growth phase- and media-dependent small non-coding RNAs in *Streptococcus pyogenes* M49 using intergenic tiling arrays." BMC genomics **13**: 550.
- Patron, K., P. Gilot, E. Camiade and L. Mereghetti (2015). "An homolog of the Frz Phosphoenolpyruvate:carbohydrate phosphoTransferase System of extraintestinal pathogenic *Escherichia coli* is encoded on a genomic island in specific lineages of *Streptococcus agalactiae*." Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases **32**: 44-50.
- Perego, M., P. Glaser, A. Minutello, M. A. Strauch, K. Leopold and W. Fischer (1995). "Incorporation of D-alanine into lipoteichoic acid and wall teichoic acid in *Bacillus subtilis*. Identification of genes and regulation." The Journal of biological chemistry **270**(26): 15598-15606.
- Pereira, U. P., G. F. Mian, I. C. Oliveira, L. C. Benchetrit, G. M. Costa and H. C. Figueiredo (2010). "Genotyping of *Streptococcus agalactiae* strains isolated from fish, human and cattle and their virulence potential in Nile tilapia." Veterinary microbiology **140**(1-2): 186-192.
- Perez-Pascual, D., P. Gaudu, B. Fleuchot, C. Besset, I. Rosinski-Chupin, A. Guillot, V. Monnet and R. Gardan (2015). "RovS and its associated signaling peptide form a cell-to-cell communication system required for *Streptococcus agalactiae* pathogenesis." mBio **6**(1).
- Perez, N., J. Trevino, Z. Liu, S. C. Ho, P. Babitzke and P. Sumby (2009). "A genome-wide analysis of small regulatory RNAs in the human pathogen group A *Streptococcus*." PloS one **4**(11): e7668.
- Peschel, A., M. Otto, R. W. Jack, H. Kalbacher, G. Jung and F. Gotz (1999). "Inactivation of the *dlt* operon in *Staphylococcus aureus* confers sensitivity to defensins, protegrins, and other antimicrobial peptides." The Journal of biological chemistry **274**(13): 8405-8410.
- Phares, C. R., R. Lynfield, M. M. Farley, J. Mohle-Boetani, L. H. Harrison, S. Petit, A. S. Craig, W. Schaffner, S. M. Zansky, K. Gershman, K. R. Stefonek, B. A. Albanese, E. R. Zell, A. Schuchat and S. J. Schrag (2008). "Epidemiology of invasive group B streptococcal disease in the United States, 1999-2005." JAMA **299**(17): 2056-2065.
- Pichon, C., L. du Merle, M. E. Caliot, P. Trieu-Cuot and C. Le Bouguenec (2012). "An in silico model for identification of small RNAs in whole bacterial genomes:

- characterization of antisense RNAs in pathogenic *Escherichia coli* and *Streptococcus agalactiae* strains." Nucleic acids research **40**(7): 2846-2861.
- Porcheron, G. and C. M. Dozois (2015). "Interplay between iron homeostasis and virulence: Fur and RyhB as major regulators of bacterial pathogenicity." Veterinary microbiology **179**(1-2): 2-14.
- Poyart, C., M. C. Lamy, C. Boumaila, F. Fiedler and P. Trieu-Cuot (2001). "Regulation of D-alanyl-lipoteichoic acid biosynthesis in *Streptococcus agalactiae* involves a novel two-component regulatory system." Journal of bacteriology **183**(21): 6324-6334.
- Poyart, C., E. Pellegrini, M. Marceau, M. Baptista, F. Jaubert, M. C. Lamy and P. Trieu-Cuot (2003). "Attenuated virulence of *Streptococcus agalactiae* deficient in D-alanyl-lipoteichoic acid is due to an increased susceptibility to defensins and phagocytic cells." Molecular microbiology **49**(6): 1615-1625.
- Prescott, Harley., Klein., Wiley., Sherwood. and Woolverton. (2010). "Microbiologie." DE BOECK SUPÉRIEUR.
- Quach, D., N. M. van Sorge, S. A. Kristian, J. D. Bryan, D. W. Shelver and K. S. Doran (2009). "The CiaR response regulator in group B *Streptococcus* promotes intracellular survival and resistance to innate immune defenses." Journal of bacteriology **191**(7): 2023-2032.
- Quentin, R. (2002). "Prise en charge des infections à *Streptococcus agalactiae* en obstétrique." Journal de Gynécologie Obstétrique et Biologie de la Reproduction **31**(6): 65-73.
- Rajagopal, L. (2009). "Understanding the regulation of Group B Streptococcal virulence factors." Future microbiology **4**(2): 201-221.
- Rajagopal, L., A. Clancy and C. E. Rubens (2003). "A eukaryotic type serine/threonine kinase and phosphatase in *Streptococcus agalactiae* reversibly phosphorylate an inorganic pyrophosphatase and affect growth, cell segregation, and virulence." The Journal of biological chemistry **278**(16): 14429-14441.
- Rajagopal, L., A. Vo, A. Silvestroni and C. E. Rubens (2006). "Regulation of cytotoxin expression by converging eukaryotic-type and two-component signalling mechanisms in *Streptococcus agalactiae*." Molecular microbiology **62**(4): 941-957.
- Ramirez-Pena, E., J. Trevino, Z. Liu, N. Perez and P. Sumby (2010). "The group A *Streptococcus* small regulatory RNA FasX enhances streptokinase activity by increasing the stability of the ska mRNA transcript." Molecular microbiology **78**(6): 1332-1347.
- Reichmann, N. T., C. P. Cassona and A. Grundling (2013). "Revised mechanism of D-alanine incorporation into cell wall polymers in Gram-positive bacteria." Microbiology **159**(Pt 9): 1868-1877.
- Reichmann, N. T. and A. Grundling (2011). "Location, synthesis and function of glycolipids and polyglycerolphosphate lipoteichoic acid in Gram-positive bacteria of the phylum Firmicutes." FEMS microbiology letters **319**(2): 97-105.
- Reignier, J. (2010). "Etude de petits ARN régulateurs chez *Helicobacter pylori*." Thèse de Doctorat de l'Université Bordeaux II – Victor Segalen: 53.
- Renault, P., G. Corthier, N. Goupil, C. Delorme and S. D. Ehrlich (1996). "Plasmid vectors for gram-positive bacteria switching from high to low copy number." Gene **183**(1-2): 175-182.
- Repoila, F. and F. Darfeuille (2009). "Small regulatory non-coding RNAs in bacteria: physiology and mechanistic aspects." Biology of the cell / under the auspices of the European Cell Biology Organization **101**(2): 117-131.
- Riani, C., K. Standar, S. Srimuang, C. Lembke, B. Kreikemeyer and A. Podbielski (2007). "Transcriptome analyses extend understanding of *Streptococcus pyogenes* regulatory

- mechanisms and behavior toward immunomodulatory substances." International journal of medical microbiology : IJMM **297**(7-8): 513-523.
- Rice, J. B., D. Balasubramanian and C. K. Vanderpool (2012). "Small RNA binding-site multiplicity involved in translational regulation of a polycistronic mRNA." Proceedings of the National Academy of Sciences of the United States of America **109**(40): E2691-2698.
- Richter, S. G., D. Elli, H. K. Kim, A. P. Hendrickx, J. A. Sorg, O. Schneewind and D. Missiakas (2013). "Small molecule inhibitor of lipoteichoic acid synthesis is an antibiotic for Gram-positive bacteria." Proceedings of the National Academy of Sciences of the United States of America **110**(9): 3531-3536.
- Rivas, E., R. J. Klein, T. A. Jones and S. R. Eddy (2001). "Computational identification of noncoding RNAs in *E. coli* by comparative genomics." Current biology : CB **11**(17): 1369-1373.
- Roberts, S. A. and J. R. Scott (2007). "RivR and the small RNA RivX: the missing links between the CovR regulatory cascade and the Mga regulon." Molecular microbiology **66**(6): 1506-1522.
- Rosenau, A., K. Martins, S. Amor, F. Gannier, P. Lanotte, N. van der Mee-Marquet, L. Mereghetti and R. Quentin (2007). "Evaluation of the ability of *Streptococcus agalactiae* strains isolated from genital and neonatal specimens to bind to human fibrinogen and correlation with characteristics of the fbsA and fbsB genes." Infection and immunity **75**(3): 1310-1317.
- Rosinski-Chupin, I., E. Sauvage, O. Sismeiro, A. Villain, V. Da Cunha, M. E. Caliot, M. A. Dillies, P. Trieu-Cuot, P. Bouloc, M. F. Lartigue and P. Glaser (2015). "Single nucleotide resolution RNA-seq uncovers new regulatory mechanisms in the opportunistic pathogen *Streptococcus agalactiae*." BMC genomics **16**: 419.
- Ruiz de los Mozos, I., M. Vergara-Irigaray, V. Segura, M. Villanueva, N. Bitarte, M. Saramago, S. Domingues, C. M. Arraiano, P. Fechter, P. Romby, J. Valle, C. Solano, I. Lasa and A. Toledo-Arana (2013). "Base pairing interaction between 5'- and 3'-UTRs controls icaR mRNA translation in *Staphylococcus aureus*." PLoS genetics **9**(12): e1004001.
- Saar-Dover, R., A. Bitler, R. Nezer, L. Shmuel-Galia, A. Firon, E. Shimoni, P. Trieu-Cuot and Y. Shai (2012). "D-alanylation of lipoteichoic acids confers resistance to cationic peptides in group B *streptococcus* by increasing the cell wall density." PLoS pathogens **8**(9): e1002891.
- Sabat, A., K. Kosowska, K. Poulsen, A. Kasproicz, A. Sekowska, B. van Den Burg, J. Travis and J. Potempa (2000). "Two allelic forms of the aureolysin gene (aur) within *Staphylococcus aureus*." Infection and immunity **68**(2): 973-976.
- Salloum, M., N. van der Mee-Marquet, A. S. Domelier, L. Arnault and R. Quentin (2010). "Molecular characterization and prophage DNA contents of *Streptococcus agalactiae* strains isolated from adult skin and osteoarticular infections." Journal of clinical microbiology **48**(4): 1261-1269.
- Salvail, H., P. Lanthier-Bourbonnais, J. M. Sobota, M. Caza, J. A. Benjamin, M. E. Mendieta, F. Lepine, C. M. Dozois, J. Imlay and E. Masse (2010). "A small RNA promotes siderophore production through transcriptional and metabolic remodeling." Proceedings of the National Academy of Sciences of the United States of America **107**(34): 15223-15228.
- Samen, U. M., B. J. Eikmanns and D. J. Reinscheid (2006). "The transcriptional regulator RovS controls the attachment of *Streptococcus agalactiae* to human epithelial cells and the expression of virulence genes." Infection and immunity **74**(10): 5625-5635.

- Santi, I., R. Grifantini, S. M. Jiang, C. Brettoni, G. Grandi, M. R. Wessels and M. Soriani (2009). "CsrRS regulates group B *Streptococcus* virulence gene expression in response to environmental pH: a new perspective on vaccine development." Journal of bacteriology **191**(17): 5387-5397.
- Santi, I., D. Maione, C. L. Galeotti, G. Grandi, J. L. Telford and M. Soriani (2009). "BibA induces opsonizing antibodies conferring in vivo protection against group B *Streptococcus*." The Journal of infectious diseases **200**(4): 564-570.
- Santiago, N. I., A. Zipf and A. K. Bhunia (1999). "Influence of temperature and growth phase on expression of a 104-kilodalton *Listeria* adhesion protein in *Listeria monocytogenes*." Applied and environmental microbiology **65**(6): 2765-2769.
- Schell, M. A. (1993). "Molecular biology of the LysR family of transcriptional regulators." Annual review of microbiology **47**: 597-626.
- Schmidtchen, A., I. M. Frick, E. Andersson, H. Tapper and L. Bjorck (2002). "Proteinases of common pathogenic bacteria degrade and inactivate the antibacterial peptide LL-37." Molecular microbiology **46**(1): 157-168.
- Schneewind, O. and D. Missiakas (2014). "Lipoteichoic acids, phosphate-containing polymers in the envelope of gram-positive bacteria." Journal of bacteriology **196**(6): 1133-1142.
- Schuchat, A. (1995). "Group B streptococcal disease in newborns: a global perspective on prevention." Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie **49**(1): 19-25.
- Serganov, A. and E. Nudler (2013). "A decade of riboswitches." Cell **152**(1-2): 17-24.
- Sharma, V., A. Yamamura and Y. Yokobayashi (2012). "Engineering artificial small RNAs for conditional gene silencing in *Escherichia coli*." ACS synthetic biology **1**(1): 6-13.
- Shelburne, S. A., 3rd, D. Keith, N. Horstmann, P. Sumby, M. T. Davenport, E. A. Graviss, R. G. Brennan and J. M. Musser (2008). "A direct link between carbohydrate utilization and virulence in the major human pathogen group A *Streptococcus*." Proceedings of the National Academy of Sciences of the United States of America **105**(5): 1698-1703.
- Shelver, D., L. Rajagopal, T. O. Harris and C. E. Rubens (2003). "MtaR, a regulator of methionine transport, is critical for survival of group B *streptococcus in vivo*." Journal of bacteriology **185**(22): 6592-6599.
- Shimizu, T., H. Yaguchi, K. Ohtani, S. Banu and H. Hayashi (2002). "Clostridial VirR/VirS regulon involves a regulatory RNA molecule for expression of toxins." Molecular microbiology **43**(1): 257-265.
- Siezen, R. J., G. Wilson and T. Todt (2010). "Prokaryotic whole-transcriptome analysis: deep sequencing and tiling arrays." Microbial biotechnology **3**(2): 125-130.
- Silhavy, T. J., D. Kahne and S. Walker (2010). "The bacterial cell envelope." Cold Spring Harbor perspectives in biology **2**(5): a000414.
- Silvaggi, J. M., J. B. Perkins and R. Losick (2005). "Small untranslated RNA antitoxin in *Bacillus subtilis*." Journal of bacteriology **187**(19): 6641-6650.
- Simonsen, K. A., A. L. Anderson-Berry, S. F. Delair and H. D. Davies (2014). "Early-onset neonatal sepsis." Clinical microbiology reviews **27**(1): 21-47.
- Six, A., C. Joubrel, A. Tazi and C. Poyart (2014). "Maternal and perinatal infections to *Streptococcus agalactiae*." Presse medicale **43**(6 Pt 1): 706-714.
- Sledjeski, D. and S. Gottesman (1995). "A small RNA acts as an antisilencer of the H-NS-silenced rcsA gene of *Escherichia coli*." Proceedings of the National Academy of Sciences of the United States of America **92**(6): 2003-2007.
- Smirnov, A., K. U. Forstner, E. Holmqvist, A. Otto, R. Gunster, D. Becher, R. Reinhardt and J. Vogel (2016). "Grad-seq guides the discovery of ProQ as a major small RNA-binding protein." Proceedings of the National Academy of Sciences of the United States of America.

- Smith, E. J., R. M. Corrigan, T. van der Sluis, A. Grundling, P. Speziale, J. A. Geoghegan and T. J. Foster (2012). "The immune evasion protein Sbi of *Staphylococcus aureus* occurs both extracellularly and anchored to the cell envelope by binding lipoteichoic acid." Molecular microbiology **83**(4): 789-804.
- Soutourina, O. A., M. Monot, P. Boudry, L. Saujet, C. Pichon, O. Sismeiro, E. Semenova, K. Severinov, C. Le Bouguenec, J. Y. Coppee, B. Dupuy and I. Martin-Verstraete (2013). "Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile*." PLoS genetics **9**(5): e1003493.
- Spellerberg, B., E. Rozdzinski, S. Martin, J. Weber-Heynemann and R. Luttkicken (2002). "rgf encodes a novel two-component signal transduction system of *Streptococcus agalactiae*." Infection and immunity **70**(5): 2434-2440.
- Stock, A. M., V. L. Robinson and P. N. Goudreau (2000). "Two-component signal transduction." Annual review of biochemistry **69**: 183-215.
- Stork, M., M. Di Lorenzo, T. J. Welch and J. H. Crosa (2007). "Transcription termination within the iron transport-biosynthesis operon of *Vibrio anguillarum* requires an antisense RNA." Journal of bacteriology **189**(9): 3479-3488.
- Storz, G., J. Vogel and K. M. Wassarman (2011). "Regulation by small RNAs in bacteria: expanding frontiers." Molecular cell **43**(6): 880-891.
- Stougaard, P., S. Molin and K. Nordstrom (1981). "RNAs involved in copy-number control and incompatibility of plasmid R1." Proceedings of the National Academy of Sciences of the United States of America **78**(10): 6008-6012.
- Sulavik, M. C., G. Tardif and D. B. Clewell (1992). "Identification of a gene, rgg, which regulates expression of glucosyltransferase and influences the Spp phenotype of *Streptococcus gordonii* Challis." Journal of bacteriology **174**(11): 3577-3586.
- Suzuki, C. K., M. Rep, J. M. van Dijl, K. Suda, L. A. Grivell and G. Schatz (1997). "ATP-dependent proteases that also chaperone protein biogenesis." Trends in biochemical sciences **22**(4): 118-123.
- Tazi, A., O. Disson, S. Bellais, A. Bouaboud, N. Dmytruk, S. Dramsi, M. Y. Mistou, H. Khun, C. Mechler, I. Tardieux, P. Trieu-Cuot, M. Lecuit and C. Poyart (2010). "The surface protein HvgA mediates group B *streptococcus* hypervirulence and meningeal tropism in neonates." The Journal of experimental medicine **207**(11): 2313-2322.
- Tazi, A., C. Joubrel, A. Six, N. Dmytruk, G. Touak and C. Poyart (2015). "Infections néonatales à *Streptococcus agalactiae* : épidémiologie, physiopathologie et diagnostic biologique." BACTÉRIOLOGIE Infections néonatales à *S. agalactiae* feuillets de Biologie(N°323).
- Tesorero, R. A., N. Yu, J. O. Wright, J. P. Svencionis, Q. Cheng, J. H. Kim and K. H. Cho (2013). "Novel regulatory small RNAs in *Streptococcus pyogenes*." PloS one **8**(6): e64021.
- Teti, G., F. Tomasello, M. S. Chiofalo, G. Orefici and P. Mastroeni (1987). "Adherence of group B *streptococci* to adult and neonatal epithelial cells mediated by lipoteichoic acid." Infection and immunity **55**(12): 3057-3064.
- Tettelin, H., V. Masignani, M. J. Cieslewicz, C. Donati, D. Medini, N. L. Ward, S. V. Angiuoli, J. Crabtree, A. L. Jones, A. S. Durkin, R. T. Deboy, T. M. Davidsen, M. Mora, M. Scarselli, I. Margarit y Ros, J. D. Peterson, C. R. Hauser, J. P. Sundaram, W. C. Nelson, R. Madupu, L. M. Brinkac, R. J. Dodson, M. J. Rosovitz, S. A. Sullivan, S. C. Daugherty, D. H. Haft, J. Selengut, M. L. Gwinn, L. Zhou, N. Zafar, H. Khouri, D. Radune, G. Dimitrov, K. Watkins, K. J. O'Connor, S. Smith, T. R. Utterback, O. White, C. E. Rubens, G. Grandi, L. C. Madoff, D. L. Kasper, J. L. Telford, M. R. Wessels, R. Rappuoli and C. M. Fraser (2005). "Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the

- microbial "pan-genome". Proceedings of the National Academy of Sciences of the United States of America **102**(39): 13950-13955.
- Thomason, M. K. and G. Storz (2010). "Bacterial antisense RNAs: how many are there, and what are they doing?" Annual review of genetics **44**: 167-188.
- Thompson, J. D., T. J. Gibson and D. G. Higgins (2002). "Multiple sequence alignment using ClustalW and ClustalX." Current protocols in bioinformatics / editorial board, Andreas D. Baxevanis ... [et al.] Chapter 2: Unit 2 3.
- Tomizawa, J. and T. Itoh (1981). "Plasmid ColE1 incompatibility determined by interaction of RNA I with primer transcript." Proceedings of the National Academy of Sciences of the United States of America **78**(10): 6096-6100.
- Tomoyasu, T., J. Gamer, B. Bukau, M. Kanemori, H. Mori, A. J. Rutman, A. B. Oppenheim, T. Yura, K. Yamanaka, H. Niki and et al. (1995). "*Escherichia coli* FtsH is a membrane-bound, ATP-dependent protease which degrades the heat-shock transcription factor sigma 32." The EMBO journal **14**(11): 2551-2560.
- Trevino, J., N. Perez, E. Ramirez-Pena, Z. Liu, S. A. Shelburne, 3rd, J. M. Musser and P. Sumbly (2009). "CovS simultaneously activates and inhibits the CovR-mediated repression of distinct subsets of group A *Streptococcus* virulence factor-encoding genes." Infection and immunity **77**(8): 3141-3149.
- Tsui, H. C., D. Mukherjee, V. A. Ray, L. T. Sham, A. L. Feig and M. E. Winkler (2010). "Identification and characterization of noncoding small RNAs in *Streptococcus pneumoniae* serotype 2 strain D39." Journal of bacteriology **192**(1): 264-279.
- Ulrich, L. E., E. V. Koonin and I. B. Zhulin (2005). "One-component systems dominate signal transduction in prokaryotes." Trends in microbiology **13**(2): 52-56.
- Ungureanu, V. (2010). "[Macrolides, lincosamides, streptogramins (MLS): mechanisms of action and resistance]." Bacteriologia, virusologia, parazitologia, epidemiologia **55**(2): 131-138.
- van der Mee-Marquet, N., L. Fourny, L. Arnault, A. S. Domelier, M. Salloum, M. F. Lartigue and R. Quentin (2008). "Molecular characterization of human-colonizing *Streptococcus agalactiae* strains isolated from throat, skin, anal margin, and genital body sites." Journal of clinical microbiology **46**(9): 2906-2911.
- Verani, J. R., L. McGee and S. J. Schrag (2010). "Prevention of perinatal group B streptococcal disease--revised guidelines from CDC, 2010." MMWR. Recommendations and reports : Morbidity and mortality weekly report. Recommendations and reports / Centers for Disease Control **59**(RR-10): 1-36.
- Vogel, J. and B. F. Luisi (2011). "Hfq and its constellation of RNA." Nature reviews. Microbiology **9**(8): 578-589.
- Vogel, J. and K. Papenfert (2006). "Small non-coding RNAs and the bacterial outer membrane." Current opinion in microbiology **9**(6): 605-611.
- Wagner, E. G., S. Altuvia and P. Romby (2002). "Antisense RNAs in bacteria and their genetic elements." Advances in genetics **46**: 361-398.
- Wampler, J. L., K. P. Kim, Z. Jaradat and A. K. Bhunia (2004). "Heat shock protein 60 acts as a receptor for the *Listeria* adhesion protein in Caco-2 cells." Infection and immunity **72**(2): 931-936.
- Wang, G., Y. Xia, J. Cui, Z. Gu, Y. Song, Y. Q. Chen, H. Chen, H. Zhang and W. Chen (2014). "The Roles of Moonlighting Proteins in Bacteria." Current issues in molecular biology **16**: 15-22.
- Wang, Y., C. G. Kelly, J. T. Karttunen, T. Whittall, P. J. Lehner, L. Duncan, P. MacAry, J. S. Younson, M. Singh, W. Oehlmann, G. Cheng, L. Bergmeier and T. Lehner (2001). "CD40 is a cellular receptor mediating mycobacterial heat shock protein 70 stimulation of CC-chemokines." Immunity **15**(6): 971-983.

- Wassarman, K. M. (2007). "6S RNA: a small RNA regulator of transcription." Current opinion in microbiology **10**(2): 164-168.
- Wassarman, K. M., F. Repoila, C. Rosenow, G. Storz and S. Gottesman (2001). "Identification of novel small RNAs using comparative genomics and microarrays." Genes & development **15**(13): 1637-1651.
- Waters, L. S. and G. Storz (2009). "Regulatory RNAs in bacteria." Cell **136**(4): 615-628.
- Weidenmaier, C., J. F. Kokai-Kun, S. A. Kristian, T. Chanturiya, H. Kalbacher, M. Gross, G. Nicholson, B. Neumeister, J. J. Mond and A. Peschel (2004). "Role of teichoic acids in *Staphylococcus aureus* nasal colonization, a major risk factor in nosocomial infections." Nature medicine **10**(3): 243-245.
- Welz, R. and R. R. Breaker (2007). "Ligand binding and gene control characteristics of tandem riboswitches in *Bacillus anthracis*." RNA **13**(4): 573-582.
- Winkler, W. C. and R. R. Breaker (2005). "Regulation of bacterial gene expression by riboswitches." Annual review of microbiology **59**: 487-517.
- Withey, J. H. and D. I. Friedman (2003). "A salvage pathway for protein structures: tmRNA and trans-translation." Annual review of microbiology **57**: 101-123.
- Witkowski, M., A. Weithauser, T. Tabaraie, D. Steffens, N. Krankel, B. Stratmann, D. Tschoepe, U. Landmesser and U. Rauch-Kroehnert (2016). "Micro-RNA-126 Reduces the Blood Thrombogenicity in Diabetes Mellitus via Targeting of Tissue Factor." Arteriosclerosis, thrombosis, and vascular biology **36**(6): 1263-1271.
- Xolalpa, W., A. J. Vallecillo, M. Lara, G. Mendoza-Hernandez, M. Comini, R. Spallek, M. Singh and C. Espitia (2007). "Identification of novel bacterial plasminogen-binding proteins in the human pathogen *Mycobacterium tuberculosis*." Proteomics **7**(18): 3332-3341.
- Yates, J. L., A. E. Arfsten and M. Nomura (1980). "In vitro expression of *Escherichia coli* ribosomal protein genes: autogenous inhibition of translation." Proceedings of the National Academy of Sciences of the United States of America **77**(4): 1837-1841.
- Yeaman, M. R. and N. Y. Yount (2003). "Mechanisms of antimicrobial peptide action and resistance." Pharmacological reviews **55**(1): 27-55.
- Zahner, D., K. Kaminski, M. van der Linden, T. Mascher, M. Meral and R. Hakenbeck (2002). "The ciaR/ciaH regulatory network of *Streptococcus pneumoniae*." Journal of molecular microbiology and biotechnology **4**(3): 211-216.
- Zaiou, M. (2007). "Multifunctional antimicrobial peptides: therapeutic targets in several human diseases." Journal of molecular medicine **85**(4): 317-329.
- Zengel, J. M. and L. Lindahl (1996). "A hairpin structure upstream of the terminator hairpin required for ribosomal protein L4-mediated attenuation control of the S10 operon of *Escherichia coli*." Journal of bacteriology **178**(8): 2383-2387.
- Zorgani, M. A., R. Quentin and M. Lartigue (2016). "Regulatory RNAs in the Less Studied Streptococcal Species: from Nomenclature to Identification." Frontiers in Microbiology **7**.

ANNEXES

Table S1 : DNA oligonucleotides

Name	Primer or probe (sequence 5' => 3')	Purpose
pTCV-P_{cetR}::lacZ (pBRMF103) assembly oligos		
BG1	ATGCgaattcGATTCTATCTTCAGGCCACA	Forward primer inserting P _{cetR} in pTCV-lac
BG2	GCATggatccTCTTCTCCTTAGATCATGGATAACCTCCTTG	Reverse primer inserting P _{cetR} in pTCV-lac
pTCV-P2_{cetR}::lacZ (pBRMF104) assembly oligos		
BG3	ATGCgaattcGATAAAGTTTAAATCAAGGAGGTT	Forward primer inserting P2 _{cetR} in pTCV-lac
BG4	GCATggatccTCTTCTCCTTAAGCGACCCATATTAATAAATA	Reverse primer inserting P2 _{cetR} in pTCV-lac
pTCV-P3_{cetR}::lacZ (pBRMF105) assembly oligos		
BG5	ATGCgaattcAAGGAAGACAGACAATGTGG	Forward primer inserting P3 _{cetR} in pTCV-lac
BG6	GCATggatccTCTTCTCCTTATTGGTTAAGTTTTCGTTGG	Reverse primer inserting P3 _{cetR} in pTCV-lac
RACE PCR primers		
PS1	TAACACATCTAGGGGTATAC	P _{cetR} reverse primer
PS2	GTCCCTCTTTATTTTTATTACCT	P2 _{cetR} reverse primer
PS3	TACACCTCACCTTCAACATA	P3 _{cetR} reverse primer
Northern probes		
P16S-F	TTGGTAGGAGTGGAATCT	Forward primer for 16S rRNA fragment
P16S-R	CGGTCAGAAGGATGTCAAGA	Reverse primer for 16S rRNA fragment
P _{cetR} -F	AAAAGAAAGATTGGTAGGCA	Forward primer for <i>cetR</i> fragment
P _{cetR} -R	TTCAGTGTCTCCGATTTAT	Reverse primer for <i>cetR</i> fragment

A. Annexe 1 : Méthodes supplémentaires

1. Cloning of Transcriptional Gene Reporter to *cetR* Putative Promoters

For the construction of $P_{cetR}::lacZ$ transcriptional fusion, a PCR fragment carrying -182 to +39 relative to the *cetR* transcription start site (Fig. 34. B) (oligos BG1/2) was digested by *EcoRI* and *BamHI* and ligated into likewise digested pTCV-lac plasmid. The ligated plasmid was cloned into *E. coli* and sequenced prior transformation into the WT (NEM316) strain. The same method was used to construct $P2_{cetR}::lacZ$ and $P3_{cetR}::lacZ$ transcriptional fusions (see Fig. 34. B for cloned fragments and Table S1 for oligos).

2. β -Galactosidase Assays

To monitor the *cetR* transcriptional fusions, in different growth conditions, we used *lacZ* gene reporter cloned into the plasmids pTCV- $P_{cetR}::lacZ$, pTCV- $P3_{cetR}::lacZ$ and pTCV- $P3_{cetR}::lacZ$. The β -galactosidase assays were performed as previously described (Miller, 1972). Strains containing transcriptional $P_{cetR}::lacZ$ fusions were grown at 37 °C overnight in TH medium and diluted 1:100 in fresh medium. Cultures were grown to mid-exponential or stationary growth phase, harvested and resuspended in 500 μ l of Z buffer (60 mM Na_2HPO_4 , 40 mM NaH_2PO_4 , 10 mM KCl, 1 mM $MgSO_4$) (Miller, 1972). Cells were then mechanically lysed, with glass beads in FastPrep®-24, centrifuged and supernatants were incubated with 0.5 mg/ml Dnase I (NEB) for 30 min at 37 °C. Specific β -galactosidase activity was calculated as described (Miller, 1972) and expressed with Arbitrary units/mg protein. Protein concentrations were measured using the Bio-Rad protein assay. A minimum of three independent experimental trials were performed for each strain and condition.

3. Determination of 5' ends of *CetR* by RACE PCR

RACE experiments were performed using linkers and primers listed in Table S1. *Streptococcus agalactiae* NEM316 strain grown in TH medium at 37 °C was diluted 100-fold in the same media and grown to stationary growth phase. Total RNA was extracted as described earlier and *CetR* 5' starts were determined using the 5'/3' rapid amplification of cDNA (RACE) PCR kit (2nd generation, Roch Applied Science), following manufacturer's

instructions. The generated PCR products of correct size were sequenced using the Big Dye Terminator v3.1 cycle sequencing kit (Applied Biosystems) and results were checked with the ABI Prism 310 Genetic Analyzer. The RNA antisense primers (PS1, PS2 and PS3) used for sequencing are listed in Table S1.

4. Northern Blot Analysis

Total RNA was extracted using the phenol/trizol method as described (Lamy *et al.*, 2004). For Northern blot analysis, 10 µg of total RNA was separated on agarose gels and blotted to Hybond-XL membrane (GE-Healthcare). Total RNA was then UV-cross-linked to the membrane and probed with $\gamma^{32}\text{-P}$ 5' end-labeled DNA oligonucleotides (*PcetR*-F/R).

5. Coating of microtiter plates with various ECM proteins

The 96-well microtiter plates are coated with 250 µL of ECM proteins: collagen type I, III and IV, the $\alpha 2$ laminin, insoluble fibronectin, elastin and Maxgel (ECM synthetically reconstructed from proteins purified). Wells coated with BSA ("bovine serum albumin") serve as control of specificity of adhesion to ECM proteins of muscle tissue. The proteins of the ECM are previously diluted in 0.1 M carbonate buffer (pH 9.6) to a concentration of 50 µg/mL of saturation of the wells. The plates were then incubated overnight at 4 °C then washed 3 times with TS (Tryptone Salt pH 7) before the filing of the bacterial suspension.

6. Adhesion test to ECM proteins

Bacterial cultures were prepared as described in paragraph 1.5, then their growth is stopped by adding Chloramphenicol (90 µg/mL). To assess their adherence ability, 200 µL of bacterial suspension are placed in each well of microtiter plates previously prepared. Wells containing only sterile culture medium are used to monitor for contamination. After 3h of incubation at 25 °C, the bacterial suspension is removed and the wells are washed twice with 250 µL of buffer TS (Tryptone Salt pH 7) to remove non-adherent cells. The adhered biomass is measured by staining with Crystal Violet. Briefly, cells were fixed to the support with 250 µL of absolute ethanol for 20 min. The wells are then emptied and dried under a hood for 30 min. 250 µL of an aqueous solution of crystal violet (0.1%) are then distributed in each well and incubated for 15 min at room temperature. After 2 successive washes with distilled

water to remove excess staining, adherent cells are discolored with 250 μL of acetic acid (33%). 200 μL are transferred into a new microtiter plate and the absorbance is measured at 595 nm.

B. Annexe 2 : Rôle de *CetR* dans la croissance en MCD

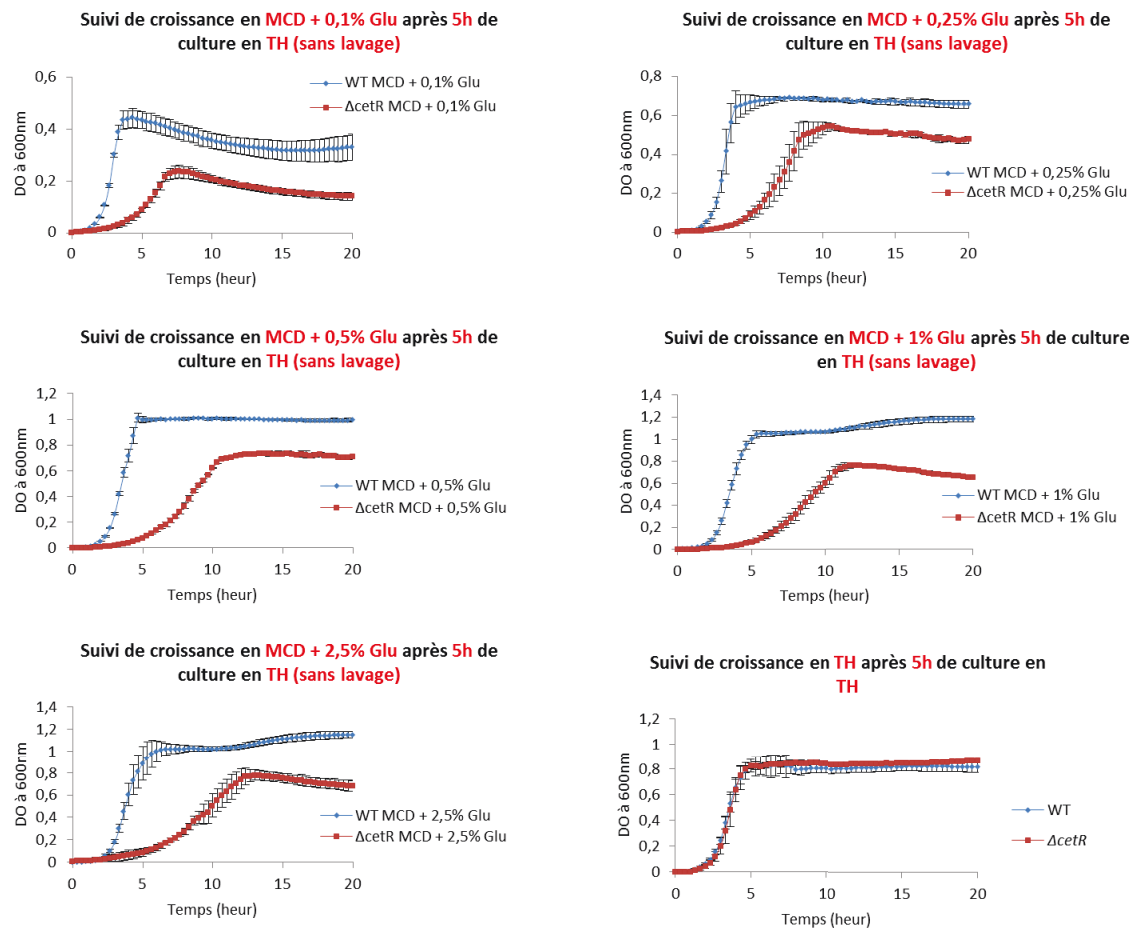


Figure S1 : Cinétique de croissance des souches WT et ΔcetR en MCD supplémenté avec différentes concentrations de glucose (Glu). La culture en MCD estensemencée avec une préculture réalisée en TH sans lavage.

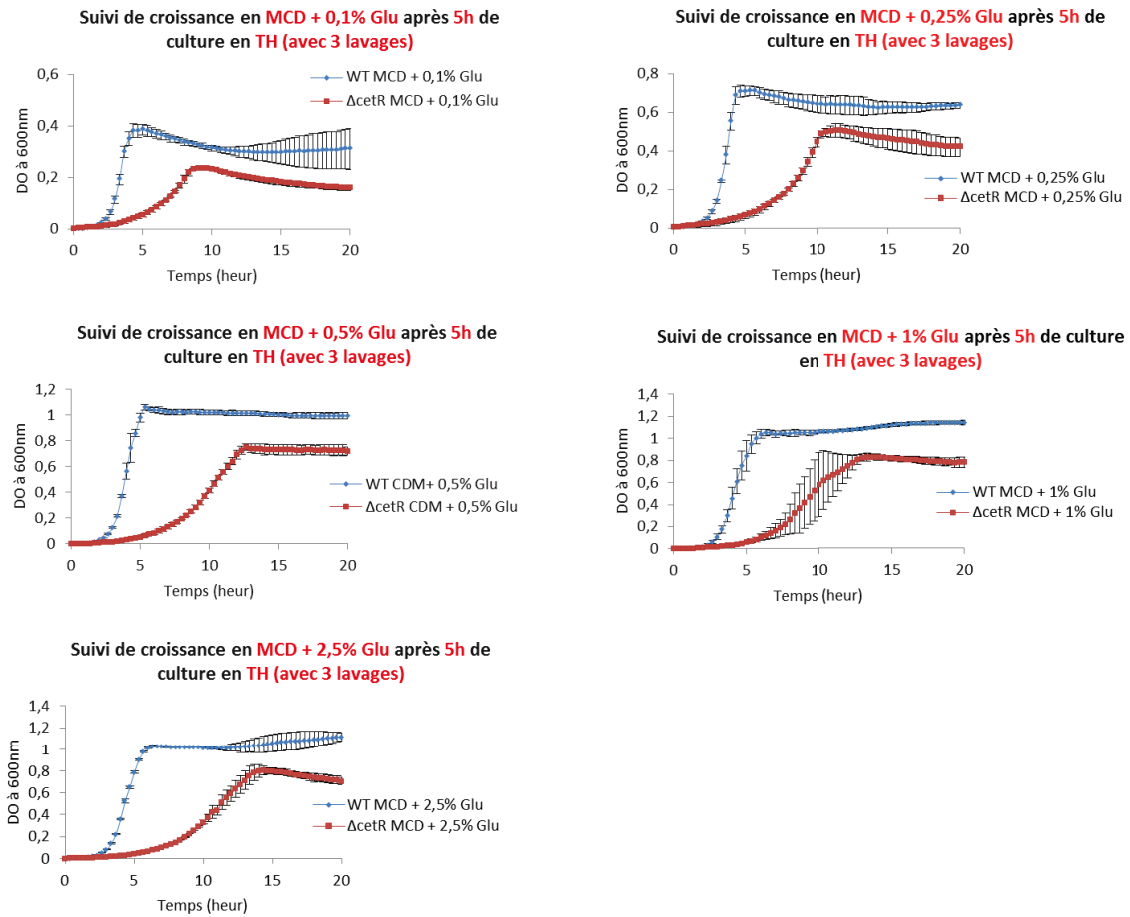


Figure S2 : Cinétique de croissance des souches WT et *ΔcetR* en MCD supplémenté avec différentes concentrations de glucose (Glu). La culture en MCD estensemencée avec la préculture réalisée en TH après 3 lavages (avec du PBS, 10 min à 6000 rpm et à 4°C).

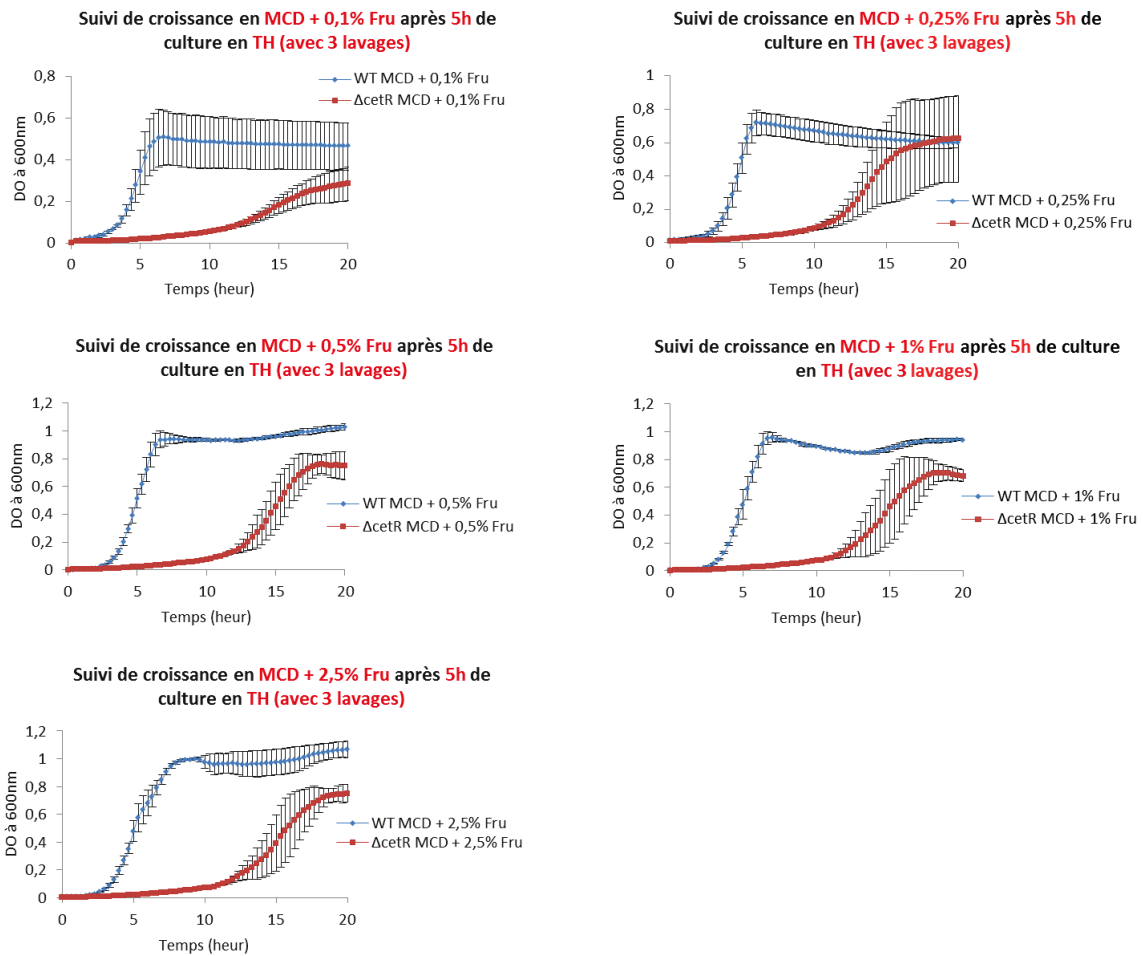


Figure S3 : Cinétique de croissance des souches WT et Δ cetR en MCD supplémenté avec différentes concentrations de fructose (Fru). La culture en MCD estensemencée avec la préculture réalisée en TH après 3 lavages (avec du PBS, 10 min à 6000 rpm et à 4°C).

C. Annexe 3 : Méthodes de l'approche protéomique

a- Bacterial strains and growth conditions

Streptococcus agalactiae strain NEM316, capsular serotype III, isolated from a neonatal case of bacteremia (Glaser *et al.*, 2002), and its derivative mutant $\Delta cetR$, were routinely grown statically in Todd Hewitt broth (TH, Difco) or in chemically defined medium at 37 °C.

For growth kinetics, bacteria were grown in 96-well microtiter plates (Greiner Bio-One, Cellstar). Overnight cultures of *S. agalactiae* strains grown in TH medium were diluted 100-fold in the same medium and grown at 37 °C to early stationary phase or overnight. Cultures were pelleted and washed three times in PBS (Phosphate Buffered Saline); and then diluted 100-fold in CDM supplemented with different concentrations (0.1%, 0.25%, 0.5%, 1% and 2.5%) of glucose (Glu) or fructose (Fru) or increasing volumes of TH.

b- Protein preparation

Bacteria were grown overnight at 37 °C in TH medium, diluted 1:1000 into 90 mL of freshly CDM supplemented with 0.5% glucose and grown at 37 °C to early stationary phase. Cells were separated from the medium by centrifugation at 6000 r.p.m for 10 min at 4 °C. The supernatants were then filtered using 0.22 µm pore-size filters (Millipore, France), mixed with 0.05 V/V of protease inhibitor cocktail (Sigma, France) to avoid degradation of proteins during the experiments, and stored at -20 °C until use. For protein precipitation, the samples were mixed to a final concentration of 0.01% of 2% DOC (Na-deoxycholate) and incubated on ice for 30 min. Subsequently, TCA (trichloroacetic acid) was added to a final concentration of 15%, vortexed immediately for 30 seconds and, after 1h, the pellets were collected by centrifugation at 15,000 r.p.m at 4 °C for 10 min. The pellets were washed twice with ice-cold acetone, centrifuged at 15,000 r.p.m at 4 °C for 10 min, dried under a slow stream of nitrogen, resuspended in 200 µL of Laemli buffer and stored at -80 °C until use.

On the other hand, the pellets were resuspended in solubilization buffer (20 mmol/L NaPO₄ pH 6.4; 0.05 V/V of protease inhibitor cocktail) to an OD (optical density) of 40 units. Cytoplasmic and membrane proteins were separated by ultracentrifugation at 50,000 r.p.m

for 30 min at 4 °C and stored at -20 °C until use. The pellets of membrane proteins were resuspended in 250 µL of solubilization buffer. Proteins yield was determined using 2D Quant kit (GE Healthcare, USA).

c- Protein identification by nano-LC-MS/MS

Short SDS-PAGE runs (precasted 4-12% Bis-Tris Mini Gels, Invitrogen, France) were performed, allowing proteins to migrate on a 0.5cm length, and gels were stained with Coomassie blue (BioRad, Marnes-la-Coquette, France). Each 1D electrophoresis lane was cut into 2-3 pieces of gel (2 mm in width) and protein identification was performed using PAPPSO platform facilities (<http://pappso.inra.fr>).

In-gel digestion was carried out according to a standard trypsinolysis protocol. Gel pieces were washed twice with 50% (v/v) acetonitrile (ACN), 25mM NH_4CO_3 , then incubated in presence of 10 mM DTT for 1 h at 56 °C. After cooling, the supernatant was removed and the samples were incubated with 55mM iodoacetamide at room temperature in the dark. Gel plugs were washed with ACN and then dried in a vacuum speed concentrator. Digestion was performed for 8 h at 37°C with 200 ng of modified trypsin (Promega) dissolved in 25mM NH_4CO_3 . Tryptic peptides were first extracted with 50% (v/v) ACN, 0.5% (v/v) trifluoroacetic acid (TFA), and then with pure ACN. Peptide extracts were dried in a vacuum speed concentrator and suspended in 25 µl of 2% (v/v) ACN, 0.05% (v/v) TFA, 0.05% (v/v) formic acid.

HPLC was performed on an Eksigent NanoLC-Ultra system. Trypsic digestion products were first concentrated and desalted on a trap column cartridge (SPE-traps C18 trap column (0.1 × 20 mm, 100 Å, 5 µm; NanoSeparations, Nieuwkoop, Netherlands) with 0.1 % HCOOH at 7.5µl.min⁻¹ for 3 min. The trap column was connected to the separating column (Stationary phase: C18 Biosphere, 3 µm. Column: 75 µm i.d., 150 mm, Nanoseparations) and the peptides were eluted with a linear gradient from 5 to 35% ACN in 0.1% HCOOH for 50 min at 300 nL/min⁻¹.

On line analysis of peptides was performed with a Q-exactive mass spectrometer (Thermo Fisher Scientific, USA), using a nanoelectrospray ion source. Ionization (1.5kV

ionization potential) was performed with stainless steel emitter (30 μ m i.d., Thermo Electron). Peptide ions were analysed using Thermo Xcalibur 2.2 with the following data-dependent acquisition steps: (1) full MS scan (mass-to-charge ratio (m/z) 400 to 1400, resolution 70000) and (2) MS/MS (normalized collision energy = 27%, resolution 17500). Step 2 was repeated for the 8 major ions detected in step 1. Dynamic exclusion was set to 40 s.

The raw mass data were first converted to mzXML format with the MSconvert software. Protein identification was performed querying MS/MS data against an *Streptococcus agalactiae* serotype III (strain NEM316) protein database (Uniprot, 01/08/2016), together with an in-house contaminant database, using the X!Tandem software (X!Tandem Cyclone (2011.12.01), <http://www.thegpm.org>) with the following parameters : one trypsin missed cleavage allowed, alkylation of cysteine and conditional oxidation of methionine, precursor and fragment ion set at 5 ppm and 0.005 Da, respectively. A refined search was added with similar parameters except that semitryptic peptides and possible N-ter acetylation of proteins were searched. All peptides matched with an E-value lower than 0.01 were parsed with X!Tandem pipeline software (<http://pappso.inra.fr/bioinfo/xtandempipeline>). Proteins identified with at least two unique peptides and a log (E-value) lower than -4 were validated.

Tableau S2: Liste des amorces utilisées pour les réactions d'amplification

Gènes	Amorces	Séquence	Taille théorique amplicon (pb)
<i>srn008</i>	srn008 F	5'-GGTGTTGCCTCCTGAGTATAC-3'	170
	srn008 R	5'-GGAGCAAAATTAGTTCGGGCT-3'	
<i>srn024</i>	srn024 F	5'-ATCCTAAATAAACTTTTCTTTTCATGT-3'	95
	srn024 R	5'-AAAAATGCAACTCAAATCATTATG-3'	
<i>srn046</i>	srn046 F	5'-CACTATTGCGCAGGGAGATTAT-3'	74
	srn046 R	5'-TCAAGTGTGACAAGAAGGTTTAC-3'	
<i>srn015</i>	srn015 F	5'-GGGTTGGAAAAAGGGTTGAG-3'	104
	srn015 R	5'-GCGGTTTATCCTAGCCGTTA-3'	
<i>srn017</i>	srn017 F	5'-TCCAACAGAAAGGAGGAGCT-3'	208
	srn017 R	5'-ACTACCCAATGCAGAACCGT-3'	
<i>srn073</i>	srn073 F	5'-GCATGCGCTAATGGTAATGA-3'	171
	srn073 R	5'-TGTTGCAAACAGTCACAAGGA-3'	
<i>gyrB</i>	gyrB F	5'-GCTCACATCAGAACTTTACTTTTAACTC-3'	123
	gyrB R	5'-TTTAATCTCACTTCCTACTTTGACACC-3'	
<i>5S</i>	5S F	5'-TAATTGCCTAGGAGATACACCTGTT-3'	112
	5S R	5'-GCTAAGCGACTACCTTATCTCACAG-3'	
<i>tmRNA</i>	tm F	5'-AAAACAGCCTGCGTCATCTT-3'	135
	tm R	5'-CAAGCCCTCACATAGCGAGT-3'	

Les amorces ont été dessinées grâce au logiciel Primer3web (<http://primer3.ut.ee/>).

F : amorce sens

R : amorce anti-sens

D. Annexe 4: Etude de la prévalence des ARNrég chez *S. agalactiae*

1. Amplification génique par PCR

L'ADN extrait pour chaque souche a été amplifié par une technique de PCR conventionnelle. Les réactions d'amplification ont été réalisées dans un thermocycleur 2720 Thermal Cycler (Applied Biosystems®). Pour un volume final de 20 µl, le mélange réactionnel contenait : 0,2 µM de chaque désoxynucléotide triphosphate (dNTPs), 0,1 µL de One Taq DNA polymerase (New England BioLabs), 1X de tampon One Taq Standard Reaction Buffer, 0,5 µM d'amorce sens et anti-sens ainsi que 200 ng d'ADN. Les 6 couples d'amorces ciblant les 6 gènes d'ARN régulateurs utilisés pour la réaction d'amplification par PCR en « point final » ont été dessinés grâce au logiciel Primer3web (<http://primer3.ut.ee/>) et sont détaillés dans le Tableau S2.

Un témoin négatif (H₂O milli-Q) ainsi qu'un témoin positif constitué de 200 ng d'ADN génomique de la souche NEM316 de *S. agalactiae* ont été utilisés dans les mêmes conditions à chaque PCR. Les amplicons obtenus ont été contrôlés par migration électrophorétique en gel d'agarose 1% préparé avec du tampon TBE (Tris Borate EDTA) 1X et contenant du bromure d'éthidium, un agent intercalant de l'ADN. Après 45 minutes de migration à 130 volts, les bandes correspondant aux produits d'amplification ont été visualisées sous ultraviolets et leurs tailles ont été appréciées comparativement au marqueur de poids moléculaire de taille connue (100 pb DNA Ladder, BioLabs).

2. Etude de la prévalence par analyse *in silico*

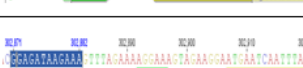
La présence des ARNrég a été recherchée, parmi la collection des 249 souches séquencées, en effectuant des recherches *blastN* sur le site internet NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) grâce aux séquences des sept ARNrég, comprenant *srn070*, obtenues à partir de notre souche de référence, NEM316. Ces sept séquences ont été également recherchées parmi les génomes séquencés de la famille *Streptococcaceae* disponibles dans la base de données.

E. Annexe 5 : ARNm cibles de Srn024 prédites *in silico*

gene	annotation	RNApredator		IntaRNA		TargetRNA		Position de l'interaction
		Position	Energie*	Position	Energie*	Position	Energie*	
<i>gbs1288</i>	Similar to pullulanase peptidoglycan bound protein (LPXTG motif)	1	-17.41	1	-16.83790	1	-14.37	
<i>gbs1396</i>	similar to mevalonate kinase	2	-17.20	2	-16.04860	5	-13.23	
<i>gbs1826</i>	-	17	-13.33	3	-14.44400	38	-9.36	
<i>gbs0281</i>	Similar to amino acid ABC transporter (permease)	4	-15.31	4	-14.40220	7	-12.28	
<i>gbs0609</i>	-	3	-15.41	10	-12.80210	46	-8.88	
<i>gbs2010</i>	sensor histidine kinase	8	-14.52	18	-11.80030	8	-11.74	
<i>gbs1434</i>	bifunctional pyrimidine regulatory protein PyrR uracil phosphoribosyltransferase	26	-13.02	20	-11.52090	9	-11.61	
<i>gbs0603</i>	DNA gyrase subunit B (gyrB)	84	-11.42	21	-11.51440	13	-11.45	
<i>gbs1624</i>	DNA replication initiation control protein YabA	14	-14.03	11	-12.48360	-	-	
<i>gbs1661</i>	transcription elongation factor GreA	21	-13.12	26	-11.26050	-	-	
<i>gbs1383</i>	ATP-dependent protease ATP-binding subunit ClpX	12	-14.19	23	-11.37100	-	-	

En Jaune : Cible prédite par les trois programmes ; En orange : Cible prédite par deux programmes

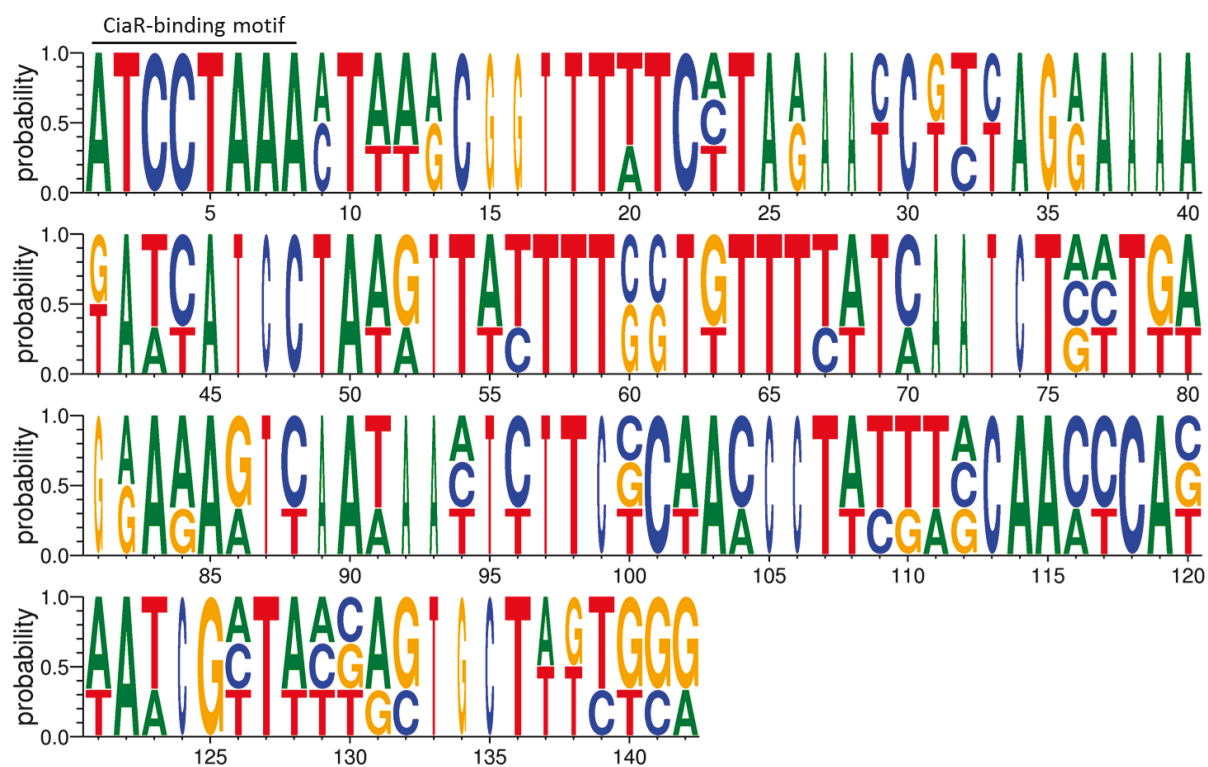
F. Annexe 6 : ARNm cibles de Srn070 prédites *in silico*

gene	annotation	RNApredator		IntaRNA		TargetRNA		Position de l'interaction
		Position	Energie*	Position	Energie*	Position	Energie*	
<i>gbs0101</i>	hypothetical protein gbs0101	1	-18.85	1	-16.09790	50	-8.52	
<i>gbs0615</i>	acid phosphatase/phosphotransferase (aphA)	2	-18.61	16	-11.55800	-	-	
<i>gbs1366</i>	hypothetical protein gbs1366	3	-17.12	7	-12.85710	-	-	
<i>gbs1989</i>	hypothetical protein gbs1989	5	-16.47	10	-12.22390	-	-	
<i>gbs0594</i>	hypothetical protein gbs0594	6	-16.24	9	-12.26720	-	-	
<i>gbs1254</i>	arsenate reductase	7	-16.06	2	-14.51280	9	-11.13	
<i>gbs1481</i>	polysaccharide biosynthesis protein	8	-15.82	24	-10.62520	-	-	
<i>gbs1552</i>	ABC transporter ATP binding protein	10	-15.77	23	-10.70180	-	-	
<i>gbs1262</i>	hypothetical protein gbs1262	13	-15.64	25	-10.51370	-	-	
<i>gbs1307</i>	laminin binding surface protein	16	-15.30	5	-13.30610	-	-	
<i>gbs0100</i>	tRNA pseudouridine synthase A	17	-15.13	42	-9.70630	-	-	
<i>gbs0270</i>	ABC transporter ATP binding protein	18	-15.09	8	-12.72110	-	-	
<i>gbs0281</i>	hypothetical protein gbs0281	19	-14.95	37	-9.87619	-	-	
<i>gbs0830</i>	hypothetical protein gbs0830	20	-14.95	4	-13.44250	47	-8.75	

En Jaune : Cible prédite par les trois programmes ; En orange : Cible prédite par deux programmes

G. Annexe 7 : Motifs conservés entre les séquences codant les ARNrég

Srn015, Srn024 et Srn070



Regulatory RNAs in the Less Studied Streptococcal Species: From Nomenclature to Identification

Mohamed A. Zorgani, Roland Quentin and Marie-Frédérique Lartigue



Regulatory RNAs in the Less Studied Streptococcal Species: From Nomenclature to Identification

Mohamed A. Zorgani^{1*}, Roland Quentin^{1,2} and Marie-Frédérique Lartigue^{1,2*}

¹ ISP, INRA, Equipe 5 "Bactéries et Risque Materno-foetal", Faculté de Médecine, UMR 1282, Université François Rabelais de Tours, Tours, France, ² Service de Bactériologie Virologie et Hygiène Hospitalière, Centre Hospitalier Régional Universitaire de Tours, Tours, France

OPEN ACCESS

Edited by:

Jörg Linde,
Leibniz-Institute for Natural Product
Research and Infection Biology
-Hans-Knoell-Institute, Germany

Reviewed by:

Manja Marz,
University of Jena, Germany
Reinhold Brückner,
Kaiserslautern University
of Technology, Germany
Gloria Del Solar,
Spanish National Research Council,
Spain

*Correspondence:

Marie-Frédérique Lartigue
marie-frederique.lartigue@univ-
tours.fr
Mohamed A. Zorgani
mohamedamine.zorgani@etu.univ-
tours.fr

Specialty section:

This article was submitted to
Infectious Diseases,
a section of the journal
Frontiers in Microbiology

Received: 02 June 2016

Accepted: 12 July 2016

Published: 26 July 2016

Citation:

Zorgani MA, Quentin R and
Lartigue M-F (2016) Regulatory RNAs
in the Less Studied Streptococcal
Species: From Nomenclature
to Identification.
Front. Microbiol. 7:1161.
doi: 10.3389/fmicb.2016.01161

Streptococcal species are Gram-positive bacteria involved in severe and invasive diseases in humans and animals. Although, this group includes different pathogenic species involved in life-threatening infections for humans, it also includes beneficial species, such as *Streptococcus thermophilus*, which is used in yogurt production. In bacteria virulence factors are controlled by various regulatory networks including regulatory RNAs. For clearness and to develop logical thinking, we start this review with a revision of regulatory RNAs nomenclature. Previous reviews are mostly dealing with *Streptococcus pyogenes* and *Streptococcus pneumoniae* regulatory RNAs. We especially focused our analysis on regulatory RNAs in *Streptococcus agalactiae*, *Streptococcus mutans*, *Streptococcus thermophilus* and other less studied *Streptococcus* species. Although, *S. agalactiae* sRNome remains largely unknown, sRNAs (small RNAs) are supposed to mediate regulation during environmental adaptation and host infection. In the case of *S. mutans*, sRNAs are suggested to be involved in competence regulation, carbohydrate metabolism, and Toxin-Antitoxin systems. A new category of miRNA-size small RNAs (msRNAs) was also identified for the first time in this species. The analysis of *S. thermophilus* sRNome shows that many sRNAs are associated to the bacterial immune system known as CRISPR-Cas system. Only few of the other different *Streptococcus* species have been the subject of studies pointed toward the characterization of regulatory RNAs. Finally, understanding bacterial sRNome can constitute one step forward to the elaboration of new strategies in therapy such as substitution of antibiotics in the management of *S. agalactiae* neonatal infections, prevention of *S. mutans* dental caries or use of *S. thermophilus* CRISPR-Cas system in genome editing applications.

Keywords: small RNAs, non-coding RNA, *Streptococcus agalactiae*, *Streptococcus mutans*, *Streptococcus thermophilus*, Regulatory RNA nomenclature

INTRODUCTION

Gene regulation is fundamental for bacteria to adapt to environmental changes. Therefore, bacteria have developed and evolved strategies to tightly control genetic networks in response to diverse extracellular stimuli (Casadesus and Low, 2006). Bacterial regulatory pathways involving small RNAs (sRNAs) are now being elucidated and their functions are coming to light,

particularly in regards to bacterial physiology, adaptation and pathogenesis as well as to plasmid replication control (Cavanagh and Wassarman, 2014; Brantl, 2015; Lopez-Aguilar et al., 2015). Beside transcriptional and translational gene regulation (activation or inhibition) through direct interaction with mRNAs, regulatory RNAs perform a variety of other mechanisms on their targets including: (i) sequestering molecules (e.g., binding to proteins; Babitzke and Romeo, 2007), (ii) affecting the expression of *cis* elements (e.g., riboswitches; Winkler and Breaker, 2005) and (iii) targeting invading cognate DNA [e.g., CRISPR RNAs (clustered regularly interspaced short palindromic repeats RNAs); Garneau et al., 2010]. The importance of sRNAs-mediated regulatory mechanisms depends on a fine-tuning of the expression of an armament of virulence factor genes (Papenfort and Vogel, 2010; Mann et al., 2012). Considering the functional roles of these RNAs, they could be targets for the development of future therapeutic drugs (Lalaouna et al., 2014).

Streptococcal species can colonize and invade humans and animals. They mostly exhibit an asymptomatic interaction within their hosts. However, several of them are well-known pathogenic species responsible for severe and life-threatening infections in humans. Infections can be exclusive to human, as it is the case for *Streptococcus pyogenes* leading to a substantial mortality and morbidity (Carapetis et al., 2005). They can also switch from a benign and commensal stage to an infectious one, like for *Streptococcus pneumoniae* (Gamez and Hammerschmidt, 2011) and the neonatal opportunistic pathogen *Streptococcus agalactiae* (Maisey et al., 2008; Krzysciak et al., 2013). In the last decades, the study of these pathogens has increased intensively due to the spreading of antibiotic resistance. Therefore, researchers are continuously invited to look for novel targets and to develop innovative therapies to combat streptococcal infections. With regard to the current knowledge on regulatory RNAs functions and modes of action, sRNAs are expected to play central roles in the bacterial regulatory networks, physiology, adaptation and virulence. Thus, they are thought-out as alternative types of regulators for adaptive gene regulation and drugs design.

Previous reviews on streptococcal sRNAs were mainly restricted to two streptococcal species: *S. pyogenes* and *S. pneumoniae* (Le Rhun and Charpentier, 2012; Patenge et al., 2013; Brantl and Bruckner, 2014; Miller et al., 2014; Wilton et al., 2015; Boulou and Repoila, 2016). However, the *Streptococcus* genus includes a large group of commensal and pathogenic Gram-positive bacteria. This review aims to provide more complete and comprehensive state of the art on the hidden part of streptococcal regulatory RNAs. It focuses on *S. agalactiae*, *Streptococcus mutans*, *Streptococcus thermophilus* sRNAs, as well as other less studied species, and brings in light their sRNome with reference to recent advances in the fast moving area of bacterial sRNAs. The clearness of scientific reports is intimately linked to the simplicity of the nomenclature been used. Thus, we dedicated the first chapter of this review to present and reemphasize the use of bacterial regulatory RNAs nomenclature. Please note that in order to put forward *Streptococci* sRNome, we will not discuss small RNAs from other Gram-positive bacteria (Romby and Charpentier, 2009; Brantl and Bruckner, 2014).

BACTERIAL REGULATORY RNAs: TERMINOLOGY ISSUES

Regulatory RNAs are commonly divided into three major groups: the first group includes *cis*-acting elements located on untranslated regions (UTRs) of a translated mRNA. They control the expression of their enclosed gene(s) through the modulation of their secondary structures or stability (Johansson et al., 2002). A second class of sRNAs are antisense RNAs (asRNAs), whose regulatory activity occurs through perfect base-pairing to the mRNA encoded by the opposite DNA strand molecule (Thomason and Storz, 2010). The last group of sRNAs, also known as *trans*-acting sRNAs, includes RNAs that target other RNAs, foreign DNA, or proteins (Gottesman and Storz, 2010). In the case of RNA targets, this group of sRNAs often shares a limited complementarity with their targets; thus, they can bind to multiple mRNAs. This type of interaction requires, often for stabilization issues, the RNA chaperone protein Hfq, which is absent in the *Streptococcus* genus and other Gram-positive bacteria (Brennan and Link, 2007; Nielsen et al., 2009).

As it is the case in proteomics and genomics (Desvaux et al., 2009), regulatory RNAs nomenclature issues raise some confusion in the field of bacterial regulatory networks. For instance, the use of the descriptors “*cis*-acting” or “*cis*-encoded” sRNA to describe an “antisense sRNA,” or the confusion between the use of “*trans*-encoded” and/or “*trans*-acting” sRNA, or even the use of different nomenclatures “Small *trans*-encoded RNA” (Nielsen et al., 2009), “Small regulatory RNA” (Frohlich and Vogel, 2009), “*trans*-encoded sRNA” (Ross et al., 2013), and “Small non-coding RNAs” (Xia et al., 2012) to cite a regulatory RNA. Thus, it is useful to reemphasize, and to remind the reader on regulatory RNA nomenclature.

In this review section, we propose a simplified and coherent nomenclature, which we believe necessary to develop comprehensible and logical thinking when describing bacterial regulatory RNAs. We propose first to revise the division of regulatory RNAs into two major categories (Table 1). The first category are the *cis*-acting RNAs. The second one, are the *trans*-acting sRNAs. This category comprises the *cis*-encoded sRNAs, known as antisense RNAs, and the *trans*-encoded sRNAs, which we refer to them in this review as “*tre*RNAs.” We propose also the use of other abbreviations for some sRNAs categories. For example, we advise the use of “*ca*RNA” and “*tra*RNA” rather than “*cis*-acting sRNA” and “*trans*-acting sRNA,” respectively. Although, few examples are listed for each category of *tra*RNAs (Table 1), the proposed list must be updated continuously since our knowledge on bacterial RNA keep changing. An exception is for the descriptor “ncRNAs” generally used to describe regulatory RNAs that do not encode for proteins, known as “non-coding RNAs,” and which can be applied to all regulatory RNA categories. As shown in Figures 1A–C regulatory RNAs nomenclature depends on their function and mode of action. The *ca*RNAs can act on transcriptional or translational levels (Figure 1A), asRNAs promote transcription attenuation or activation/inhibition of translation (Figure 1B). The *tre*RNAs can repress (Figures 1Ca–b), or activate translation (Figures 1Cc–d),

TABLE 1 | Bacterial regulatory RNAs nomenclature and mode of action.

Bacterial regulatory RNAs			
Mode of action Nomenclature	Abbreviations	Description	
cis-acting RNAs			
	caRNA	The descriptor “cis-acting RNA” is used to describe regulatory RNAs that act (Ruiz de los Mozos et al., 2013), or are located at the untranslated regions (UTRs) of mRNA molecules (Gossinger and Hartmann, 2012).	
UnTranslatedRegion	UTR	UTRs are cis-acting RNAs located on untranslated regions of a coding transcript. They can activate/inhibit translation by releasing/hindering the RBS, respectively. They can also participate in the mRNA transcript stabilization (Gossinger and Hartmann, 2012). The base pairing between 5'- and 3'-UTRs can also control mRNA translation (Ruiz de los Mozos et al., 2013).	
5' UnTranslatedRegion	5' UTR	5' UTRs are RNAs that act on the 5'-end of the mRNA transcript. They modulate the expression of the downstream genes by controlling the access of the ribosome to the translation initiation region (TIR). This class of UTRs includes Riboswitches, non-transcribed RNA, and independent transcripts. They play an important role in bacterial metabolism control. The modulation of the expression of the 5' UTR-downstream genes occurs through a metabolite-induced alteration of the RNA secondary structure, such as tRNAs (Loh et al., 2009).	
3' UnTranslatedRegion	3' UTR	3' UTRs are RNAs that act on the 3'-end of the mRNA transcript. They constitute the region of the mRNA that immediately follows the translation termination codon (TTC; Chao et al., 2012).	
trans-acting RNAs (traRNAs)			
cis-encoded (antisense)RNA	asRNA	asRNAs are cis-encoded RNAs that act on the mRNA molecule encoded by the opposite DNA strand. They present a perfect base-pairing with their mRNA target (Thomason and Storz, 2010). We also distinguish toxin-antitoxin systems (Lee and Hong, 2011) and asRNAs encoded by the streptococcal pIP501 plasmid (Brantl and Wagner, 1996), or long asRNAs described in eukaryotes (Vadaie and Morris, 2013).	
5' antisense RNA	5' asRNA	5' asRNAs are asRNAs that act on the 5'-end of the sense mRNA target. They activate or inhibit the translation of the mRNA transcript (Thomason and Storz, 2010). 5' asRNAs can also act by transcription attenuation, which is frequently found on plasmids (Thomason and Storz, 2010).	
3' antisense RNA	3' asRNA	3' asRNAs are asRNAs that act on the 3'-end of the sense mRNA target. The role of the 3' asRNAs is not yet well-deciphered. However, it was suggested that they participate in the transcript stabilization (Opdyke et al., 2004; Padalon-Brauch et al., 2008).	
trans-encoded RNA	treRNA	treRNAs are RNAs that target distant and/or different RNAs, eliminate invading cognate DNA, or form ribonucleo-protein complexes (Storz et al., 2011). We can also distinguish in this class: housekeeping RNAs such as 6S RNA, RNase P RNA (Brantl, 2009), overlapping ncRNA (Kumar et al., 2010), and operon RNAs (Kreikemeyer et al., 2001).	
cis-dependent small RNA	csRNA	csRNAs are treRNAs that are transcribed from the promoters associated to CiaRH two-component system. They present a complementarity to the Shine-Dalgarno sequence and the translation initiation codon (AUG; Marx et al., 2010).	
microRNA-size sRNA	msRNA	msRNAs are a very small RNAs (15 to 26 nt) that are first described in <i>Streptococcus mutans</i> . The role of these microRNA-like sRNAs in bacterial gene regulation is not yet understood (Lee and Hong, 2011; Kang et al., 2013).	
CRISPR RNA	crRNA	As a part of the CRISPR (clustered regularly interspaced short palindromic repeat) system, the crRNAs act as treRNAs targeting and processing foreign DNA (Karvelis et al., 2013).	
trans-activating CRISPR RNA	tracrRNA	In the CRISPR/Cas system, tracrRNAs are in charge of the activation of the pre-crRNAs (Karvelis et al., 2013).	

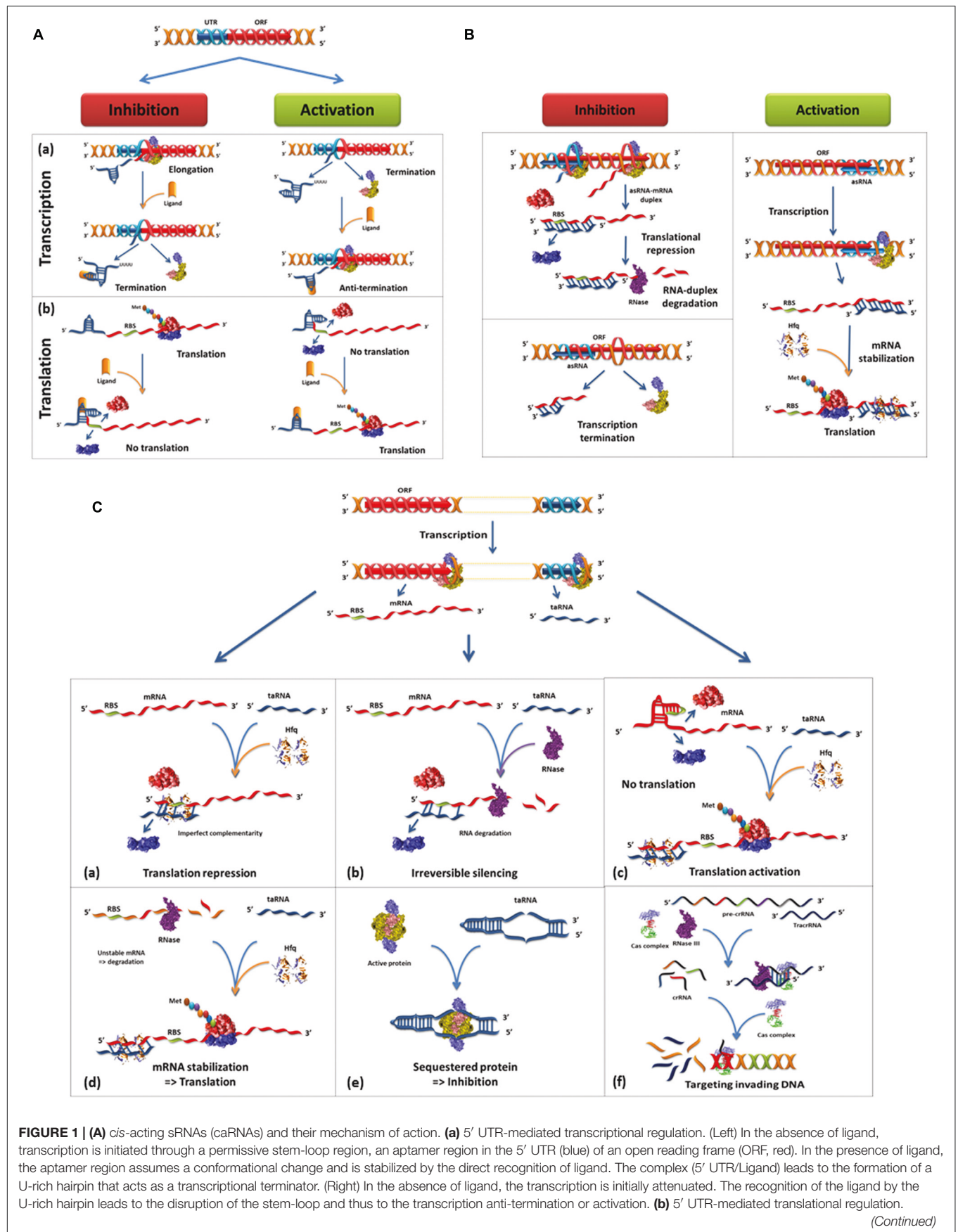


FIGURE 1 | Continued

(Left) In the absence of ligand, the aptamer region forms an anti-terminator hairpin. The ribosome binding site (RBS, green) is available, and mRNA translation is initiated. Upon ligand binding, the access to the RBS is hindered, preventing translation. (Right) In the absence of ligand, translation is inhibited due to a sequestered RBS. The recognition of the ligand by the stem-loop releases the RBS and translation can be initiated. In other cases, the ligand can be replaced by a temperature-dependent mechanism, where the mRNA adopts a secondary structure that hinders the access to the RBS (*Listeria monocytogenes*). **(B)** Mechanisms of mRNA regulation by antisense RNAs (asRNA). (Left upper) The binding of asRNA (blue) to the TIR (translation initiation region) leads to translation repression. The RNA duplex (asRNA-sense RNA) can be degraded by RNase III, a double strand specific endoribonuclease conserved in all the three kingdoms (Cho and Kim, 2015); or by RNase E which belong to the Gram-negative bacteria degradosome (Stazic et al., 2011). (Left lower) The asRNA can bind to the 5' end of the target mRNA and causes changes in the target RNA structure leading to transcription termination. (Right) The complex asRNA-Hfq can also stabilize target mRNA for translation activation. **(C)** *trans*-encoded sRNAs (treRNAs)-based regulatory mechanisms of mRNA expression. treRNAs act through an imperfect basepairing to mRNA targets. The treRNA/mRNA duplex is stabilized by Hfq and can either repress **(a–b)**, **(c)** activate or **(d)** stabilize mRNA translation. **(e)** It was also demonstrated that treRNAs can bind to proteins to inhibit and/or modify their activity. **(f)** *tracr*RNAs are a treRNA involved in the maturation of *cr*RNAs through the interaction with the complex CRISPR-Cas9, pre-*cr*RNA and RNase III. The processed *cr*RNAs target the invading DNA.

sequester proteins (**Figure 1Ce**) or participate in degrading invading DNA (**Figure 1Cf**).

***Streptococcus agalactiae*: WHAT IS BEHIND THE HIDDEN sRNome?**

Apart from its commensal colonization of the gastrointestinal and urogenital tracts, *S. agalactiae*, known also as Group B *Streptococcus* (GBS), has become a predominant cause of invasive infections in neonates and causes pneumonia, septicemia, and meningitis (Maisey et al., 2008; Krzysciak et al., 2013). However, the current knowledge on the sRNome encoded by the GBS chromosome remains limited, even though the sRNAs can play an important role in virulence of various Gram-positive bacteria including *Streptococcus* (Patenge et al., 2013). Two *in silico* studies were conducted on the sequenced genome of *S. agalactiae* NEM316. The first study aimed to predict sRNAs associated to the CiaRH two component system; we refer to them as csRNAs (Marx et al., 2010; **Table 1**). The CiaRH response regulator in *S. agalactiae* plays an important role in promoting the innate immune system, and intracellular survival (Quach et al., 2009). The search for CiaR-activated promoter (NTTAAG-N5-TTTAAG) and the transcriptional terminator sequences within the intergenic regions led to the identification of four putative csRNAs (csRNA10, csRNA11, csRNA12, and csRNA13; Marx et al., 2010). However, no experimental evidence on the expression of *S. agalactiae* csRNAs was reported in this study. The presence of complementarity to translation initiation regions in all the csRNAs suggests that they can bind to mRNAs and block translation initiation by hindering the access of ribosomes. Of note, the expression of csRNA10, csRNA11, and csRNA12 was recently confirmed and their up-regulation by acidic conditions was demonstrated using northern blot analysis in *S. agalactiae* NEM316 strain (Rosinski-Chupin et al., 2015). The control by the CiaRH two-component system of these csRNAs underlines their potential role in *S. agalactiae* acidic stress adaptation in the urogenital tract.

The second *in silico* study was based on prediction of Rho-independent terminators (RTTs) as a specific signature of sRNA gene loci. One hundred ninety-seven sRNAs were detected on the genome of *S. agalactiae* NEM316 (Pichon et al., 2011). The expression of 10 out of the 197 sRNAs was confirmed by RT-PCR and northern blot hybridization. One of the sRNAs candidates

(SQ517) detected in this study was previously reported as putative csRNA (csRNA12; Marx et al., 2010). The over-expression of SQ18 asRNA significantly decreased, more than 100-fold, the level of its target *gbs0031* mRNA, which encodes a surface immunogenic protein (Sip). A translational fusion system showed also that the over-expression of SQ18 was associated with fourfold decrease in the amount of *gbs0031::gfp* mRNA, suggesting that SQ18 acts as a negative post-transcriptional regulator.

A more recent study led to the identification of 125 sRNAs: 39 caRNAs, 39 asRNAs, and 47 treRNAs (Rosinski-Chupin et al., 2015). Ten of these sRNAs were differentially expressed in response to acidic stress. The expression of Srn015 (csRNA10), Srn024 (csRNA11), Srn070 (csRNA12), Srn085 (csRNA13), Srn071 and Srn082 was induced by more than two-fold at pH of 5.2 (natural pH of the genitourinary tract). In contrast, Srn046, Srn056, Srn057, and Srn073 were down regulated in the same experimental condition. All csRNAs identified in this study were previously predicted (Marx et al., 2010), indicating that sRNA *in silico* prediction is a powerful tool for csRNA identification in *Streptococci*.

Other sRNAs were described previously in mobile genetic elements of *S. agalactiae*, like the asRNAs involved in the replication control of the pIP501 and pMV158 plasmids (Brantl and Behnke, 1992; del Solar and Espinosa, 1992; Brantl and Wagner, 1994, 1996; Brantl, 2015; Lopez-Aguilar et al., 2015). Both pIP501 and pMV158 were originally isolated from clinical strains of *S. agalactiae* and exhibit a broad range of hosts. The asRNAs RNAIII and RNAII encoded by pIP501 and pMV158, respectively, have been found important for the inhibition of expression of the plasmids essential genes. While pIP501 RNAIII asRNA exhibits unusual stability, the RNAII encoded by pMV158 did not require the formation of a kissing complex for efficient binding to its mRNA target or for inhibition of gene expression. Due to a limited number of available genome sequences, the *S. agalactiae* sRNome remains largely unknown.

***Streptococcus mutans* sRNAs: FROM VERY SMALL TO A TOXIC MOOD**

Proposed to be the new Gram-positive paradigm, *S. mutans* is a host-associated dental pathogen (Lemos et al., 2013). This bacterium is considered to be a cariogenic organism responsible

for the formation of dental biofilm matrix and some occasional infective endocarditis. Beside its ability to survive at low pH and to produce large quantities of organic acid, *S. mutans* presents an ease of genetic manipulation. This fluency is due to the presence of a two competence regulatory system ComR/ComS, which makes the bacterium transformable in different culture conditions (Son et al., 2012). In *S. mutans*, sRNAs were first described after high-throughput experiments and a kingdom-wide prediction of bacterial sRNA-encoding genes using SIPHT (sRNA identification protocol using high-throughput technologies) method (Livny et al., 2008). This strategy identifies new treRNAs candidates by co-localizing their intergenic regions and the Rho-independent terminators. Eighteen treRNA genes of *S. mutans* were predicted, however, their functions as those of most *S. mutans* treRNAs remain elusive (Livny et al., 2008). Biocomputational studies predicted that three treRNAs were putatively regulated by CiaRH (csRNA23-1, csRNA23-2, and csRNA24; Marx et al., 2010). As mentioned earlier, csRNAs should get deeper attention owing to their potential role in *S. mutans* virulence through the interaction with the CiaRH regulon.

A new member of the treRNAs category, baptized “miRNA-size small RNAs” or “msRNAs” was recently described in *S. mutans* (Lee and Hong, 2011; **Table 1**). In eukaryotes, microRNAs (miRNAs) are 19–23 nucleotides (nts) non-coding RNAs that act as post-transcriptional regulators leading mainly to gene silencing (Bartel, 2004). However, so far no publication reports sRNAs with comparable size to miRNAs in bacteria. A deep sequencing approach suggested that detected reads corresponding to 15 and 26 nts were putative msRNAs (Lee and Hong, 2011). These RNAs may result from processed RNAs. In addition, northern profiles show discrete band of msRNA-428 suggesting that this msRNA is not produced through a random RNA degradation. Likewise, the expression of seven msRNAs was validated by qRT-PCR. Currently, the role of msRNAs in gene regulation remains unclear.

The ability of *S. mutans* to metabolize a multitude of carbohydrate sources and to colonize the tooth surface is intimately linked to its ability to survive at low pH. Xia et al. (2012) described for the first time in *S. mutans* a pH-dependent sRNA. This treRNA, L10-leader sRNA, was predicted by combining different bioinformatics approaches. A total number of 334 sRNAs were predicted, however, only the L10-leader was verified by qRT-PCR and northern blot analyses due to its high abundance. The expression level of the L10-leader was growth phase-dependent, and highly affected by pH. In clinical strains, its expression was higher in the most adhering and acidic strains and may be linked to virulence (Xia et al., 2012).

Experimental identification or bioinformatics screens alone are not enough to detect all encoded sRNAs, since their expression often depends on environmental conditions (growth phase, stress condition...). Combining *in vivo* and *in silico* approaches is the best way to identify a large number of bacterial sRNAs. Rho-independent terminator-based, transcription initiation signal-based and RNAs prediction methods were coupled to RNA-seq approach to identify the sRNome of *S. mutans* UA159, under carbohydrate repression as well as in

the presence/absence of the catabolite control protein CcpA (Zeng et al., 2013). Among the 114 sRNAs expressed, five were differentially expressed in *S. mutans* UA159 grown in glucose versus galactose, and two treRNAs were deregulated in the *ccpA* mutant. The fact that these treRNAs are expressed in carbohydrate- and/or *ccpA*-dependent manner highlights that they may be involved in the regulation of the *S. mutans* carbohydrate metabolism.

Bacteria have also developed many adaptive systems including Toxin–Antitoxin (TA) modules. These small genetic elements express a stable toxic protein (toxin), and a proximate antitoxin. The nature of the antitoxin, RNA or protein, defines the type of the TA module (Wen and Fozo, 2014). In the type I TA system, the antitoxin is an RNA molecule that controls the mRNA stability and translation of the corresponding toxin molecule. It is proposed that such systems help bacteria to survive to environmental stress, and when exposed to high concentrations of antibiotics through the formation of persister cells (Gerdes et al., 2005). Recently, a functional type I TA system was described for the first time in *S. mutans* UA159 (Koyanagi and Levesque, 2013). This TA module is located in a small intergenic region of 318 nts. The toxin (Fst-Sm) is a small hydrophobic peptide and the antitoxin (srSm) is an asRNA of approximately 70 nts in length. The 5′ end mapping of Fst-Sm and srSm transcripts indicates that they are both transcribed through a direct tandem repeat. It was also demonstrated that the expression of srSm represses the expression of the toxin when fused to a GFP gene reporter. Overexpression of the toxin Fst-Sm induced *E. coli* cell death, and was lethal in *S. mutans*. Ectopic expression of the TA module in *S. mutans* was followed with a decrease in persistent cell formation. Altogether, these results suggest that the *S. mutans* TA systems are implicated in infections caused by oral bacteria. *S. mutans* has the ability to cope with complex environments such as oral cavities, being responsible for dental caries and chronic infections. Deciphering the role of sRNAs in *S. mutans* regulatory networks may help providing some clues to design novel therapeutic strategies for dental caries.

***Streptococcus thermophilus* sRNome: TOO RICH OR TOO POOR?**

Streptococcus thermophilus is a lactic acid bacterium used in yogurt manufacturing and also commonly found in fermented milk products. *S. thermophilus* is considered as non-pathogenic for food industries (Blomqvist et al., 2006). In genetics, the name of *S. thermophilus* is intimately linked to the bacterial immune system known as CRISPR-Cas system (Deveau et al., 2008; Sapranauskas et al., 2011; Karvelis et al., 2013; Carte et al., 2014). This sRNA-based immune system protects bacteria from invading DNA, such as viruses and plasmids (**Table 1**). An important player in this process, as it is the case for the Cas9 protein and the crRNA (CRISPR RNA), is the tracrRNA (trans-activating CRISPR RNA). Northern blot profiling revealed the presence of a tracrRNA located upstream of the *cas9* gene of the CRISPR3-Cas in *S. thermophilus* LMD-9 genome (Karvelis et al., 2013). This tracrRNA is involved in crRNA biogenesis

through the pre-crRNA maturation, and is required for the Cas9-mediated interference. In another study, it was demonstrated that this trRNA, with the help of the RNase III, can stimulate the cleavage within the CRISPR repeats generating the crRNA units (Carte et al., 2014).

Other sRNAs were reported for *S. thermophilus*. Five csRNAs were predicted on *S. thermophilus* CNRZ1066 genome and one csRNA on the plasmid pSt0 (Marx et al., 2010). The expression of the latter csRNA was verified by northern blot. The *S. thermophilus* CiaRH system is inactivated by the presence of a stop codon interrupting the *ciaR* gene, and a frame shift mutation in *ciaH*. This could be a result of a particular adaptation to the dairy niche leading to loss of the function of virulence genes. Another four trRNAs were predicted by homology comparison to *S. pyogenes* MGAS315 sRNome (Tesorero et al., 2013). The expression of these trRNAs needs to be verified.

Streptococci sRNAs: A LARGE SPECIES PANEL

Currently, 67 different *Streptococcus* species are reported on the NCBI database and their genome sequenced and annotated. Nevertheless, only a few of them have been the subject of studies pointed toward the characterization of sRNAs. Bioinformatics approaches have been widely used to detect sRNAs in bacterial genomes. Tesorero et al. (2013) reported that among the 67 available genomes, 23 different *Streptococcus* genomes harbor at least one sRNA homolog to the sRNAs detected for *S. pyogenes* MGAS315 strain. Moreover, species as *Streptococcus suis*, *Streptococcus uberis*, *Streptococcus mitis*, *Streptococcus oralis*, *Streptococcus sanguinis*, *Streptococcus dysgalactiae* were also reported to harbor more than two csRNAs (Marx et al., 2010).

For *S. suis*, a pig pathogen, a differential RNA-sequencing approach was conducted on the strain P1/7 grown in rich medium and biological fluids (pig blood and cerebrospinal fluid) and allowed the identification of 29 sRNAs (Wu et al., 2014). Ten of them were predicted to be conserved in 34 pathogenic *Streptococcus* species. The deletion of five trRNAs led to the attenuation of *S. suis* virulence in a zebrafish infection model. *S. suis* mutant strains were more sensitive to killing by pig blood following the deletion of three other trRNAs as compared with

the wild-type. This study shows also that genes involved in the synthesis of capsular polysaccharide were differentially regulated in blood and cerebrospinal fluid. The role of *S. suis* sRNAs in virulence, the mechanism of interaction and regulation with their mRNA targets need to be investigated.

FUTURE PROSPECTS AND CONCLUDING REMARKS

The aim of this focused review was to bring in light the sRNomes of *Streptococcus* species. As cited earlier, understanding *S. agalactiae* sRNome can constitute one step forward in providing new strategies for the substitution of antibiotics in the treatment of neonatal infections. A major cause of dental caries, *S. mutans*, harbors a significant number of sRNAs that can play an important role in its virulence. *S. thermophilus* is the only *Streptococcus* species used in food industry. The use of CRISPR-Cas system in genome editing through genetic engineering is a very promising tool in bacterial genome new function design. The future challenges in the study of sRNAs in bacteria are the identification of their targets and the understanding of their regulatory mechanism. In *Streptococci* cited in this review, there is a lack between the identification of sRNAs and their functional characterization. Future works should be done on the mechanistic level; target identification and mechanism of regulation, in order to better appreciate these regulatory networks.

AUTHOR CONTRIBUTIONS

MZ wrote the manuscript. M-FL and RQ participated in manuscript elaboration and correction.

ACKNOWLEDGMENTS

We are grateful to Philippe Boulouc and Tatiana Rochat for critical reading of the manuscript and helpful discussion. We thank the reviewers for critical reading of the manuscript. MZ received a Ph.D. fellowship from the french Ministère de l'Éducation et de l'Enseignement Supérieur.

REFERENCES

- Babitzke, P., and Romeo, T. (2007). CsrB sRNA family: sequestration of RNA-binding regulatory proteins. *Curr. Opin. Microbiol.* 10, 156–163. doi: 10.1016/j.mib.2007.03.007
- Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297.
- Blomqvist, T., Steinmoen, H., and Havarstein, L. S. (2006). Natural genetic transformation: a novel tool for efficient genetic engineering of the dairy bacterium *Streptococcus thermophilus*. *Appl. Environ. Microbiol.* 72, 6751–6756. doi: 10.1128/AEM.01156-06
- Boulouc, P., and Repoila, F. (2016). Fresh layers of RNA-mediated regulation in Gram-positive bacteria. *Curr. Opin. Microbiol.* 30, 30–35. doi: 10.1016/j.mib.2015.12.008
- Brantl, S. (2009). Bacterial chromosome-encoded small regulatory RNAs. *Future Microbiol.* 4, 85–103. doi: 10.2217/17460913.4.1.85
- Brantl, S. (2015). Antisense-RNA mediated control of plasmid replication - pIP501 revisited. *Plasmid* 78, 4–16. doi: 10.1016/j.plasmid.2014.07.004
- Brantl, S., and Behnke, D. (1992). Copy number control of the streptococcal plasmid pIP501 occurs at three levels. *Nucleic Acids Res.* 20, 395–400. doi: 10.1093/nar/20.3.395
- Brantl, S., and Bruckner, R. (2014). Small regulatory RNAs from low-GC Gram-positive bacteria. *RNA Biol.* 11, 443–456. doi: 10.4161/rna.28036
- Brantl, S., and Wagner, E. G. (1994). Antisense RNA-mediated transcriptional attenuation occurs faster than stable antisense/target RNA pairing: an in vitro study of plasmid pIP501. *EMBO J.* 13, 3599–3607.
- Brantl, S., and Wagner, E. G. (1996). An unusually long-lived antisense RNA in plasmid copy number control: in vivo RNAs encoded by the streptococcal

- plasmid pIP501. *J. Mol. Biol.* 255, 275–288. doi: 10.1006/jmbi.1996.0023
- Brennan, R. G., and Link, T. M. (2007). Hfq structure, function and ligand binding. *Curr. Opin. Microbiol.* 10, 125–133. doi: 10.1016/j.mib.2007.03.015
- Carapetis, J. R., Steer, A. C., Mulholland, E. K., and Weber, M. (2005). The global burden of group A streptococcal diseases. *Lancet Infect. Dis.* 5, 685–694. doi: 10.1016/S1473-3099(05)70267-X
- Carte, J., Christopher, R. T., Smith, J. T., Olson, S., Barrangou, R., Moineau, S., et al. (2014). The three major types of CRISPR-Cas systems function independently in CRISPR RNA biogenesis in *Streptococcus thermophilus*. *Mol. Microbiol.* 93, 98–112. doi: 10.1111/mmi.12644
- Casadesu, J., and Low, D. (2006). Epigenetic gene regulation in the bacterial world. *Microbiol. Mol. Biol. Rev.* 70, 830–856. doi: 10.1128/MMBR.00016-06
- Cavanagh, A. T., and Wassarman, K. M. (2014). 6S RNA, a global regulator of transcription in *Escherichia coli*, *Bacillus subtilis*, and beyond. *Annu. Rev. Microbiol.* 68, 45–60. doi: 10.1146/annurev-micro-092611-150135
- Chao, Y., Papenfort, K., Reinhardt, R., Sharma, C. M., and Vogel, J. (2012). An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. *EMBO J.* 31, 4005–4019. doi: 10.1038/emboj.2012.229
- Cho, K. H., and Kim, J. H. (2015). Cis-encoded non-coding antisense RNAs in streptococci and other low GC Gram (+) bacterial pathogens. *Front. Genet.* 6:110. doi: 10.3389/fgene.2015.00110
- del Solar, G., and Espinosa, M. (1992). The copy number of plasmid pLS1 is regulated by two trans-acting plasmid products: the antisense RNA II and the repressor protein, RepA. *Mol. Microbiol.* 6, 83–94. doi: 10.1111/j.1365-2958.1992.tb00840.x
- Desvaux, M., Hebraud, M., Talon, R., and Henderson, I. R. (2009). Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue. *Trends Microbiol.* 17, 139–45. doi: 10.1016/j.tim.2009.01.004
- Deveau, H., Barrangou, R., Garneau, J. E., Labonte, J., Fremaux, C., and Boyaval, P., et al. (2008). Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* 190, 1390–1400. doi: 10.1128/JB.01412-07
- Frohlich, K. S., and Vogel, J. (2009). Activation of gene expression by small RNA. *Curr. Opin. Microbiol.* 12, 674–682. doi: 10.1016/j.mib.2009.09.009
- Gamez, G., and Hammerschmidt, S. (2011). Combat pneumococcal infections: adhesins as candidates for protein-based vaccine development. *Curr. Drug Targets* 13, 323–337. doi: 10.2174/138945012799424697
- Garneau, J. E., Dupuis, M. E., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., et al. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71. doi: 10.1038/nature09523
- Gerdes, K., Christensen, S. K., and Lobner-Olesen, A. (2005). Prokaryotic toxin-antitoxin stress response loci. *Nat. Rev. Microbiol.* 3, 371–382. doi: 10.1038/nrmicro1147
- Gossringer, M., and Hartmann, R. K. (2012). 3'-UTRs as a source of regulatory RNAs in bacteria. *EMBO J.* 31, 3958–3960. doi: 10.1038/emboj.2012.269
- Gottesman, S., and Storz, G. (2010). Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harb. Perspect. Biol.* 3:12. doi: 10.1101/cshperspect.a003798
- Johansson, J., Mandin, P., Renzoni, A., Chiaruttini, C., Springer, M., and Cossart, P. (2002). An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*. *Cell* 110, 551–561.
- Kang, S. M., Choi, J. W., Lee, Y., Hong, S. H., and Lee, H. J. (2013). Identification of microRNA-size, small RNAs in *Escherichia coli*. *Curr. Microbiol.* 67, 609–613. doi: 10.1007/s00284-013-0411-9
- Karvelis, T., Gasiunas, G., Miksys, A., Barrangou, R., Horvath, P., and Siksnys, V. (2013). crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*. *RNA Biol.* 10, 841–851. doi: 10.4161/rna.24203
- Koyanagi, S., and Levesque, C. M. (2013). Characterization of a *Streptococcus mutans* intergenic region containing a small toxic peptide and its cis-encoded antisense small RNA antitoxin. *PLoS ONE* 8:e54291. doi: 10.1371/journal.pone.0054291
- Kreikemeyer, B., Boyle, M. D., Buttaro, B. A., Heinemann, M., and Podbielski, A. (2001). Group A streptococcal growth phase-associated virulence factor regulation by a novel operon (Fas) with homologies to two-component-type regulators requires a small RNA molecule. *Mol. Microbiol.* 39, 392–406. doi: 10.1046/j.1365-2958.2001.02226.x
- Krzysciak, W., Pluskwa, K. K., Jurczak, A., and Koscielniak, D. (2013). The pathogenicity of the *Streptococcus* genus. *Eur. J. Clin. Microbiol. Infect. Dis.* 32, 1361–1376. doi: 10.1007/s10096-013-1914-9
- Kumar, R., Shah, P., Swiatlo, E., Burgess, S. C., Lawrence, M. L., and Nanduri, B. (2010). Identification of novel non-coding small RNAs from *Streptococcus pneumoniae* TIGR4 using high-resolution genome tiling arrays. *BMC Genomics* 11:350. doi: 10.1186/1471-2164-11-350
- Lalaouna, D., Eyraud, A., Chabelskaya, S., Felden, B., and Masse, E. (2014). Regulatory RNAs involved in bacterial antibiotic resistance. *PLoS Pathog.* 10:e1004299. doi: 10.1371/journal.ppat.1004299
- Le Rhun, A., and Charpentier, E. (2012). Small RNAs in streptococci. *RNA Biol.* 9, 414–426. doi: 10.4161/rna.20104
- Lee, H. J., and Hong, S. H. (2011). Analysis of microRNA-size, small RNAs in *Streptococcus mutans* by deep sequencing. *FEMS Microbiol. Lett.* 326, 131–136. doi: 10.1111/j.1574-6968.2011.02441.x
- Lemos, J. A., Quivey, R. G. Jr., Koo, H., and Abranches, J. (2013). *Streptococcus mutans*: a new Gram-positive paradigm? *Microbiology* 159:436–445. doi: 10.1099/mic.0.066134-0
- 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. doi: 10.1371/journal.pone.0003197
- Loh, E., Dussurget, O., Gripenland, J., Vaitkevicius, K., Tiensuu, T., Mandin, P., et al. (2009). A trans-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*. *Cell* 139, 770–779. doi: 10.1016/j.cell.2009.08.046
- Lopez-Aguilar, C., Romero-Lopez, C., Espinosa, M., Berzal-Herranz, A., and Del Solar, G. (2015). The 5'-tail of antisense RNAII of pMV158 plays a critical role in binding to the target mRNA and in translation inhibition of repB. *Front. Genet.* 6:225. doi: 10.3389/fgene.2015.00225
- Maisey, H. C., Doran, K. S., and Nizet, V. (2008). Recent advances in understanding the molecular basis of group B *Streptococcus* virulence. *Expert Rev. Mol. Med.* 10:27. doi: 10.1017/S1462399408000811
- Mann, B., van Opijnen, T., Wang, J., Obert, C., Wang, Y. D., Carter, R., et al. (2012). Control of virulence by small RNAs in *Streptococcus pneumoniae*. *PLoS Pathog.* 8:e1002788. doi: 10.1371/journal.ppat.1002788
- Marx, P., Nuhn, M., Kovacs, M., Hakenbeck, R., and Bruckner, R. (2010). Identification of genes for small non-coding RNAs that belong to the regulon of the two-component regulatory system CiaRH in *Streptococcus*. *BMC Genomics* 11:661. doi: 10.1186/1471-2164-11-661
- Miller, E. W., Cao, T. N., Pflughoeft, K. J., and Sumby, P. (2014). RNA-mediated regulation in Gram-positive pathogens: an overview punctuated with examples from the group A *Streptococcus*. *Mol. Microbiol.* 94, 9–20. doi: 10.1111/mmi.12742
- Nielsen, J. S., Lei, L. K., Ebersbach, T., Olsen, A. S., Klitgaard, J. K., Valentin-Hansen, P., et al. (2009). Defining a role for Hfq in Gram-positive bacteria: evidence for Hfq-dependent antisense regulation in *Listeria monocytogenes*. *Nucleic Acids Res.* 38, 907–919. doi: 10.1093/nar/gkp1081
- Opdyke, J. A., Kang, J. G., and Storz, G. (2004). GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J. Bacteriol.* 186:6698–6705. doi: 10.1128/JB.186.20.6698-6705.2004
- Padalon-Brauch, G., Hershberg, R., Elgrably-Weiss, M., Baruch, K., Rosenshine, I., Margalit, H., et al. (2008). Small RNAs encoded within genetic islands of *Salmonella typhimurium* show host-induced expression and role in virulence. *Nucleic Acids Res.* 36, 1913–1927. doi: 10.1093/nar/gkn050
- Papenfort, K., and Vogel, J. (2010). Regulatory RNA in bacterial pathogens. *Cell. Host. Microbe* 8, 116–127. doi: 10.1016/j.chom.2010.06.008
- Patenge, N., Fiedler, T., and Kreikemeyer, B. (2013). Common regulators of virulence in streptococci. *Curr. Top. Microbiol. Immunol.* 368, 111–153. doi: 10.1007/82_2012_295
- Pichon, C., du Merle, L., Caliot, M. E., Trieu-Cuot, P., and Le Bouguenec, C. (2011). An in silico model for identification of small RNAs in whole bacterial genomes: characterization of antisense RNAs in pathogenic *Escherichia coli* and *Streptococcus agalactiae* strains. *Nucleic Acids Res.* 40, 2846–2861. doi: 10.1093/nar/gkr1141
- Quach, D., van Sorge, N. M., Kristian, S. A., Bryan, J. D., Shelper, D. W., and Doran, K. S. (2009). The CiaR response regulator in group B *Streptococcus* promotes intracellular survival and resistance to innate immune defenses. *J. Bacteriol.* 191, 2023–2032. doi: 10.1128/JB.01216-08

- Romby, P., and Charpentier, E. (2009). An overview of RNAs with regulatory functions in Gram-positive bacteria. *Cell. Mol. Life Sci.* 67, 217–237. doi: 10.1007/s00018-009-0162-8
- Rosinski-Chupin, I., Sauvage, E., Sismeiro, O., Villain, A., Da Cunha, V., Caliot, M. E., et al. (2015). Single nucleotide resolution RNA-seq uncovers new regulatory mechanisms in the opportunistic pathogen *Streptococcus agalactiae*. *BMC Genomics* 16:419. doi: 10.1186/s12864-015-1583-4
- Ross, J. A., Ellis, M. J., Hossain, S., and Haniford, D. B. (2013). Hfq restructures RNA-IN and RNA-OUT and facilitates antisense pairing in the Tn10/IS10 system. *RNA* 19, 670–684. doi: 10.1261/rna.037747.112
- 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. doi: 10.1371/journal.pgen.1004001
- Sapranas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* 39, 9275–9282. doi: 10.1093/nar/gkr606
- Son, M., Ahn, S. J., Guo, Q., Burne, R. A., and Hagen, S. J. (2012). Microfluidic study of competence regulation in *Streptococcus mutans*: environmental inputs modulate bimodal and unimodal expression of comX. *Mol. Microbiol.* 86, 258–272. doi: 10.1111/j.1365-2958.2012.08187.x
- Stazic, D., Lindell, D., and Steglich, C. (2011). Antisense RNA protects mRNA from RNaseE degradation by RNA-RNA duplex formation during phage infection. *Nucleic Acids Res.* 39, 4890–4899. doi: 10.1093/nar/gkr037
- Storz, G., Vogel, J., and Wassarman, K. M. (2011). Regulation by small RNAs in bacteria: expanding frontiers. *Mol. Cell* 43, 880–891. doi: 10.1016/j.molcel.2011.08.022
- Tesorero, R. A., Yu, N., Wright, J. O., Svencionis, J. P., Cheng, Q., Kim, J. H., et al. (2013). Novel regulatory small RNAs in *Streptococcus pyogenes*. *PLoS ONE* 8:e64021. doi: 10.1371/journal.pone.0064021
- Thomason, M. K., and Storz, G. (2010). Bacterial antisense RNAs: how many are there, and what are they doing? *Annu. Rev. Genet.* 44, 167–188. doi: 10.1146/annurev-genet-102209-163523
- Vadaie, N., and Morris, K. V. (2013). Long antisense non-coding RNAs and the epigenetic regulation of gene expression. *Biomol. Concepts* 4, 411–415. doi: 10.1515/bmc-2013-0014
- Wen, J., and Fozo, E. M. (2014). sRNA antitoxins: more than one way to repress a toxin. *Toxins (Basel)* 6, 2310–2335. doi: 10.3390/toxins6082310
- Wilton, J., Acebo, P., Herranz, C., Gomez, A., and Amblar, M. (2015). Small regulatory RNAs in *Streptococcus pneumoniae*: discovery and biological functions. *Front. Genet.* 6:126. doi: 10.3389/fgene.2015.00126
- Winkler, W. C., and Breaker, R. R. (2005). Regulation of bacterial gene expression by riboswitches. *Annu. Rev. Microbiol.* 59:487–517. doi: 10.1146/annurev.micro.59.030804.121336
- Wu, Z., Wu, C., Shao, J., Zhu, Z., Wang, W., Zhang, W., et al. (2014). The *Streptococcus suis* transcriptional landscape reveals adaptation mechanisms in pig blood and cerebrospinal fluid. *RNA* 20, 882–898. doi: 10.1261/rna.041822.113
- Xia, L., Xia, W., Li, S., Li, W., Liu, J., Ding, H., et al. (2012). Identification and expression of small non-coding RNA, L10-Leader, in different growth phases of *Streptococcus mutans*. *Nucleic Acid Ther.* 22, 177–186. doi: 10.1089/nat.2011.0339
- Zeng, L., Choi, S. C., Danko, C. G., Siepel, A., Stanhope, M. J., and Burne, R. A. (2013). Gene regulation by CcpA and catabolite repression explored by RNA-Seq in *Streptococcus mutans*. *PLoS ONE* 8:e60465. doi: 10.1371/journal.pone.0060465

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

Copyright © 2016 Zorgani, Quentin and Lartigue. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

I. Annexe 9 : Publication 4

Small Non-coding RNAs in *Streptococci*

M.A Zorgani^{1*}, E Camiade¹, R. Quentin^{1, 2}, MF Lartigue^{1, 2*}

¹ ISP, INRA, Université Tours, Equipe "Bactéries et risque materno-foetal", UMR1282, F-37032 Tours, France.

² Centre Hospitalier Régional Universitaire de Tours, Service de bactériologie virologie et hygiène hospitalière, F-37044 Tours, France.

*Corresponding authors: Marie-Frédérique Lartigue, E-mail: marie-frederique.lartigue@univ-tours.fr. Mohamed Amine Zorgani, E-mail: mohamedamine.zorgani@etu.univ-tours.fr.

Université François Rabelais de Tours, Faculté de médecine, UMR1282 Equipe 5 Bactéries et Risque Materno-Fœtal, 10 Boulevard Tonnellé, F-37032 Tours, France.

Keywords: small RNAs, *Streptococcus*, *cis*-acting RNAs *trans*-acting RNAs, CRISPR RNAs

Soumis à *Frontiers in Genetics*: Research Topic Editorial
<http://journal.frontiersin.org/researchtopic/2627/small-non-coding-rnas-in-streptococci>

1 Bacterial small RNAs (sRNAs) have been identified as acting on gene posttranscriptional
2 regulation, and the mechanisms by which this can occur have begun to be understood. In
3 pathogenic bacteria, the importance of sRNAs-mediated regulatory mechanisms depends on a
4 fine-tuning of the expression of virulence factor genes. Considering the functional roles of
5 these RNAs, they could be categorized in two major classes: (i) *cis*-acting elements
6 located/act on untranslated regions (UTRs) of a translated mRNA which control the
7 expression of their enclosed gene(s) through the modulation of their secondary structures or
8 stability; and (ii) the *trans*-acting sRNAs that comprises the *cis*-encoded sRNAs, known as
9 antisense RNAs (asRNAs), and the *trans*-encoded sRNAs which generally present imperfect
10 base-pairing with the mRNA targets. The *trans*-encoded sRNAs are the most studied and
11 different types have been described, mainly, sRNAs that target invading cognate nucleic
12 acids, known as CRISPR RNAs (Clustered Regularly Interspaced Short Palindromic Repeats
13 RNAs), sRNAs that are transcribed from the promoters associated to CiaRH two-component
14 system, crRNAs (*cia*-dependent small RNAs), or even microRNA-like sRNAs described for
15 the first time in *Streptococcus mutans*. The sRNAs widespread in the bacterial kingdom
16 highlights their importance in regard to regulatory networks and pathogenesis and open new
17 perspectives for future therapeutic and industrial applications.

18 Streptococcal species are Gram-positive bacteria involved in severe and invasive diseases in
19 humans and animals. Presently, the main challenges in the study of sRNAs in *Streptococci* are
20 their identification, as well as the identification of their potential targets. However, several
21 works have been published and a large number of putative regulatory sRNAs were recently
22 identified in streptococcal species. Nevertheless, the biological functions of only few sRNAs
23 have been investigated. The goal of this special research topic is to bring together research
24 reviews and original articles on sRNAs, their identification, characterization and mode of
25 action in *Streptococci*. It is intended that this will cover all sRNAs categories cited elsewhere
26 regarding the well-studied *Streptococcus pyogenes* and *Streptococcus pneumoniae*, as well as
27 other streptococcal species, for instance, *Streptococcus agalactiae* which is an opportunistic
28 pathogen involved in neonatal infections.

First, the current knowledge on streptococcal sRNAs identification was reported by Patenge and collaborators (Patenge *et al.*, 2015). After reviewing the different streptococcal species, the authors described the screening methods used for the identification of sRNAs in the human pathogen *S. pneumoniae* and *S. pyogenes*. Bioinformatics and high throughput screening methods, like next generation sequencing, allowed the identification of an increasing number of sRNAs in both species. For instance, in *S. pyogenes*, few sRNAs had been identified and described to control bacterial virulence, such as FasX, RivX and CRISPR RNAs. Furthermore, when 63 sRNA candidates were predicted using sRNAPredict2, only eight of them were identified by whole genome tiling microarrays in *S. pneumoniae* TIGR4. This clearly indicates that focused efforts on experimental screening methods are highly needed to reduce the overlap with bioinformatics data. This effort, through using *in vivo* methods, was reviewed by Wilton and collaborators on *trans*-acting RNAs in *S. pneumoniae* (Wilton *et al.*, 2015). The *trans*-acting RNAs are sRNAs that target distant and/or different mRNAs. The authors discussed the biological functions of chromosomal- and extra-chromosomal-encoded sRNAs, their role in virulence and competence. This report revealed that many sRNAs have been identified in *S. pneumoniae*, but their function and mode of action remains to be deciphered regarding the crucial role that sRNAs may play for *S. pneumoniae* pathogenesis.

Among *trans*-acting sRNAs, Laux and co-authors demonstrated that the csRNAs are involved in competence control in *S. pneumoniae* R6 (Laux *et al.*, 2015). The csRNAs are non-coding RNAs transcribed from the promoters associated to CiaRH two-component system. They generally present a complementarity to the Shine-Dalgarno sequence and the translation initiation codon (AUG). The authors showed, by using reporter gene fusions, that deletion of a single or two csRNAs was not sufficient to impair competence gene expression. In contrast, combination of three csRNAs, or mutations in the competence gene *comC* and its complemented csRNAs, completely blocked *S. pneumoniae* competence. This study demonstrated that csRNAs are important players in competence regulation for *S. pneumoniae*. Cho & Kim (2015) reviewed the current knowledge about asRNAs in low GC Gram-positive bacterial pathogens. This category of RNAs can be involved in transcription attenuation or activation/inhibition of translation. One example of gene translation inhibition by asRNA, RNAII, was described by Lopez-Aguilar and co-authors (Lopez-Aguilar *et al.*, 2015). The RNAII encoded by pMV158 did not require the formation of a kissing complex for efficient binding to its mRNA target or for inhibition of *repB* expression. The *repB* gene encodes the replication initiator protein of the pneumococcal pMV158 plasmid. However, site-directed

mutagenesis on the 5' and 3' regions on the 5'-tail of RNAII, revealed that the entire 5'-tail on the asRNA is mandatory for efficient *repB* translation *in vitro* and pMV158 replication *in vivo*.

The CRISPR sRNAs class was reviewed by an original article on the distribution of the type II-A CRISPR-Cas system among the different *S. agalactiae* genetic lineages (Lier *et al*, 2015). The phylogenetic affiliation of the CRISPR-*cas* loci was studied in 126 clinical isolates of *S. agalactiae*. Beside the ubiquitous presence and the high sequence polymorphism of CRISPR1 spacers, one of the main finding of this study is that strains belonging to the hypervirulent sequence type 17 showed a lower number of spacers, evidence of potential role in virulence for this specific bacterial lineage.

This special research topic was the opportunity to review the different sRNAs categories and to emphasize their role in gene regulation, physiology and virulence. “*Understanding the ways that bacteria respond to and influence communities and how they survive such diverse environments will benefit from further studies of sRNAs*” (Wilton *et al*, 2015).

References

- Cho, KH., and Kim, JH. (2015). *Cis*-encoded non-coding antisense RNAs in *streptococci* and other low GC Gram (+) bacterial pathogens. *Front. Genet.* 6: 110. doi: 10.3389/fgene.2015.00110
- Laux, A., Sexauer, A., Sivaselvarajah, D., Kaysen, A., and Bruckner, R. (2015). Control of competence by related non-coding csRNAs in *Streptococcus pneumoniae* R6. *Front. Genet.* 6: 246. doi:10.3389/fgene.2015.00246
- Lier, C., Baticle, E., Horvath, P., Haguenoer, E., Valentin, AS., Glaser, P., Mereghetti, L., and Lanotte, P. (2015). Analysis of the type II-A CRISPR-Cas system of *Streptococcus agalactiae* reveals distinctive features according to genetic lineages. *Front. Genet.* 6: 214. doi:10.3389/fgene.2015.00214
- Lopez-Aguilar, C., Romero-Lopez, C., Espinosa, M., Berzal-Herranz, A., and Del Solar, G. (2015). The 5'-tail of antisense RNAII of pMV158 plays a critical role in binding to the target mRNA and in translation inhibition of *repB*. *Front. Genet.* 6: 225. doi: 10.3389/fgene.2015.00225
- Patenge, N., Pappesch, R., Khani, A., and Kreikemeyer, B. (2015). Genome-wide analyses of small non-coding RNAs in *streptococci*. *Front. Genet.* 6: 189. doi:10.3389/fgene.2015.00189
- Wilton, J., Acebo, P., Herranz, C., Gomez, A., and Amblar, M. (2015). Small regulatory RNAs in *Streptococcus pneumoniae*: discovery and biological functions. *Front. Genet.* 6: 126. doi: 10.3389/fgene.2015.00126

The Adc/Lmb system mediates zinc acquisition and contributes to *Streptococcus agalactiae* growth and survival

*Pauline Moulin, Kévin Patron, Camille Cano, **Mohamed A. Zorgani**, Emilie Camiade, Elise Borezée-Durant, Agnès Rosenau, Laurent Mereghetti and Aurélia Hiron*

The Adc/Lmb system mediates zinc acquisition in
Streptococcus agalactiae and contributes to bacterial growth
and survival

Pauline Moulin ¹, Kévin Patron ¹, Camille Cano ¹, Mohamed Amine Zorgani ¹,
Emilie Camiade ¹, Elise Borezée-Durant ², Agnès Rosenau ^{1, 3}, Laurent
Mereghetti ^{1, 3}, and Aurélia Hiron ^{1, #}

¹ ISP, Université François Rabelais Tours, INRA, UMR1282, F-37032 Tours, France

² Micalis institute, AgroParis Tech, Université Paris-Saclay, 78350 Jouy en Josas,
France

³ CHRU de Tours, Service de Bactériologie-Virologie et Hygiène, F-37044 Tours,
France

For correspondence

E-mail: aurelia.hiron@univ-tours.fr

Tel: (+33) 2 47 36 61 97

Running title: Characterization of a zinc transporter of *Streptococcus agalactiae*

ABSTRACT

The Lmb protein of *Streptococcus agalactiae* is described as an adhesin that binds laminin, a component of the human extracellular matrix. In this study, we reveal a new role for this protein in zinc uptake. We also identified two Lmb homologs, AdcA and AdcAll, redundant binding proteins that combine with the AdcCB translocon to form a zinc-ABC transporter. Expression of this transporter is controlled by the zinc concentration of the medium through the zinc-dependent regulator AdcR. Triple deletion of *lmb*, *adcA*, and *adcAll*, or that of the *adcCB* encoding genes, impaired growth and cell separation in a zinc-restricted environment. Moreover, we showed that this Adc zinc-ABC transporter promotes *S. agalactiae* growth or survival in some human biological fluids, suggesting that it contributes to the infection process. These results indicate that zinc has biologically vital functions in *S. agalactiae* and that, under the conditions tested, the Adc/Lmb transporter constitutes the main zinc-acquisition system of the bacterium.

IMPORTANCE

A zinc transporter, composed of three redundant binding-proteins (Lmb, AdcA and AdcAll) was characterized in *Streptococcus agalactiae*. This system was shown to be essential for bacterial growth and morphology in zinc-restricted environments, including human biological fluids.

INTRODUCTION

Streptococcus agalactiae (Group B *Streptococcus*, GBS) is a Gram-positive commensal bacterium of the human gastro-intestinal and uro-genital tracts. GBS carriage is mostly asymptomatic in healthy adults and this bacterium is detected in the vagina of approximately 30% of pregnant women. Maternal carriage is the main source of transmission to neonates in which *S. agalactiae* can cause invasive infections (pneumonia, septicaemia, and meningitis) with an overall mortality rate of approximately 10% (1, 2). GBS is also an emergent pathogen among the elderly or adults with underlying diseases (1, 3).

The ability of GBS to colonize different niches and cause infection is multifactorial and many virulence-associated proteins have been identified (4, 5). Binding of GBS adhesins to components of the extracellular matrix constitutes a crucial first step in the process of infection (6–9). Among them, the Lmb protein has been identified as a GBS receptor for laminin, a glycosylated multidomain protein found in all human tissues (10). The gene encoding Lmb is located on a transposon with the *scpB* and *sht* genes, encoding a C5a peptidase and a histidine triad protein, respectively (11, 12). The *lmb* promoter region is a hot spot for the integration of two mobile genetic elements. One of them is associated with increased expression of *lmb* resulting in increased binding of strains harboring the transposon to laminin (13). It has also been shown that Lmb may promote bacterial invasion using human brain microvascular endothelial cell lines (14).

Lmb is clustered by sequence homology as a metal-binding receptor. Indeed, Lmb has strong homology with the zinc-binding proteins AdcA and Lbp of other streptococcal species (15, 16). The crystal structure of Lmb has also been resolved and revealed the presence of a bound zinc in a metal-binding crevice (17, 18). Until

now, studies have only been focused on its function as an adhesin, despite evidence suggesting that Lmb may be involved in zinc uptake.

Zinc (Zn^{2+}) is a trace element that is essential for most living cells. It is a cofactor for a number of essential prokaryotic enzymes and transcriptional regulators (19). In the human body, the zinc concentration varies from 1.5 μM in cerebrospinal fluid to over 100 μM in lung tissue (20, 21). Pathogenic bacteria must adapt zinc transport to accommodate these differences to both avoid toxicity and meet their requirements for this metal. An important determinant of resistance against elevated levels of Zn determinant in *Streptococcus pneumoniae*, is the CzcD-SczA system, which is conserved among *Streptococcus* species (22). On the other side, during starvation, zinc acquisition in Streptococci is mostly performed by an ATP-binding cassette (ABC) transporter. It is composed of one or several metal-binding proteins (AdcA, Lbp, or Lmb), an integral membrane component (AdcB), and an ATPase (AdcC) (21). The streptococcal AdcR repressor, a MarR family regulatory protein is involved in the regulation of zinc uptake genes (23). In the presence of adequate intracellular zinc concentrations, AdcR binds to its target genes and inhibits their expression. AdcR repression is relieved during zinc starvation, allowing bacteria to efficiently adapt the expression of zinc acquisition systems to their needs for this metal (24–27).

In several pathogenic streptococci, deletion of one or several components of the Adc zinc-ABC transporter results in lower growth under zinc-restricted conditions as well as decreased virulence, adhesion, and biofilm formation, underlying the importance of zinc metabolism during colonization and infection (16, 21, 28–30). Concerning *S. agalactiae*, the role of zinc is still poorly documented and no zinc transporter has yet been characterized.

90 In this study, we showed that Lmb, together with the AdcA, AdcAll binding-
91 proteins, and the AdcCB translocon, compose a zinc transporter. The expression and
92 the role of this transporter in the bacterium physiology were examined.

MATERIALS AND METHODS

Bacterial strains and growth conditions

The reference wild type (WT) GBS strain used in this study is A909, a ST 7, CC 7 and serotype III clinical isolate from a human case of septicaemia. *E. coli* and *S. agalactiae* strains used in this study are listed in Table S1. *E. coli* strains were routinely grown in Luria Bertani (LB) medium (MP, Solon, USA; Cat. n° 1005317) at 37°C with agitation (220 rpm), or on LB-agar plate (1.5% agar). *S. agalactiae* strains were routinely grown in Todd Hewitt (TH) broth (Sigma-Aldrich, St-Louis, Mo; Cat. n° T1438) at 37°C without agitation, and on TH-agar or Brain Heart Infusion-agar (BHI-agar) (AES, Bruz, France, Cat. n° EAB140102). When necessary, *E. coli* and *S. agalactiae* strains were grown with erythromycin (150 µg/ml for *E. coli* or 10 µg/ml for *S. agalactiae*).

Growth of bacteria in chemically defined medium (CDM)

S. agalactiae strains were cultured in a liquid chemically defined medium (CDM) [8.3 g/l Dulbecco's Modified Eagle medium base (Cat n° D5030, Sigma-Aldrich), 1 X BME vitamins, 74 µM adenine, 89.2 µM uracil, 65.7 µM xanthine, 66.2 µM guanine, 1123.5 µM DL-alanine, 757 µM L-asparagine, 1127 µM L-aspartic acid, 684.5 µM L-glutamine, 1019.5 µM L-glutamic acid, 868.6 µM L-proline, 734.9 µM L-tryptophan, 4125.4 µM L-cysteine, 12 µM lipoic acid, 1 µM pyruvate; pH 7.4]. All CDM components were from Sigma-Aldrich.

To obtain zinc-restricted CDM, 500 µM of the chelating agent EDTA (Sigma-Aldrich, Mo; Cat. n° E5134) was added to remove residual traces of metal ions. Addition of EDTA had no discernible effect on *S. agalactiae* growth compared to growth in CDM without EDTA or growth in CDM with a mix of metals [17.4 µM ZnSO₄, 10.5 µM CoCl₂, 0.4 µM CuSO₄, 147.9 µM MnSO₄].

Bacterial growths were performed in 96-well microtiter plates (Greiner Bio-One, Cellstar®) (300 µl culture volume). A bacterial over-night culture grown in TH was used to inoculate a zinc-restricted CDM ($OD_{600nm} = 0.005$) grown during 8 hours that was used to inoculate wells ($OD_{600nm} = 0.005$) containing zinc-restricted CDM with various $ZnSO_4$ concentrations (from 0 to 100 µM). Before each inoculation, cells were pelleted and washed in MilliQ water. Plates were incubated for 18 hours at 37°C in a Eon™ thermoregulated spectrophotometer plate reader (BioTek® Instruments). OD_{600nm} was measured every hour after a double orbital shaking of the plate during 5 seconds.

Chromosomal and plasmidic DNA purification

Chromosomal DNA of *S. agalactiae* cultured overnight without agitation in TH (37°C) was purified as previously described (31).

E. coli plasmids were purified with a NucleoSpin® Plasmid kit (Macherey-Nagel®), according to the manufacturer's instructions.

DNA concentration was measured with a NanoDrop™ Lite Spectrophotometer (Thermo scientific®). The ratio of absorbance at 260 nm and 280 nm was used to assess its purity.

DNA sequencing

PCR products purified with the NucleoSEQ® kit (Macherey-Nagel®) were sequenced on both strands using the Big Dye Terminator v3.1 cycle sequencing kit from Applied Biosystems and the ABI Prism 310 Genetic Analyzer.

Polymerase Chain Reaction (PCR)

Oligonucleotides (Eurogentec® and Sigma-Aldrich®) used in this study are listed in Table 1. Analytical PCR used standard OneTaq polymerase (NEB®) and

PCR for cloning or sequencing was carried out with Q5 High-Fidelity DNA polymerase (NEB®). The resulting PCR fragments were purified with a NucleoSpin® Gel and PCR clean-up kit (Macherey-Nagel®), according to the manufacturer's instructions.

RNA extraction

Total RNAs were extracted from mid-exponential ($OD_{600nm} = 0.5$) growing cells in CDM. The bacteria were lysed mechanically with glass beads in a FastPrep®-24 instrument, and total RNAs were extracted with a phenol/trizol-based purification method as previously described (32). Concentration and purity of RNA were assessed with a NanoDrop™ Lite Spectrophotometer (Thermo Scientific), and analysis of the A260/A280 ratio. A DNase (TURBO DNA-free DNase™, Ambion®) treatment of the purified RNAs was then realized. Control PCRs, using 50 µg of the purified RNAs, were performed to check for DNA contamination (Primers pairs are detailed in Table 1).

Reverse Transcription and quantitative Reverse-Transcriptase PCR (qRT-PCR)

The RNAs were reverse-transcribed by using the iScript™ cDNA Synthesis kit (Bio-Rad®, Hercules, CA), according to the manufacturer's instructions.

Primers were selected with Primer3web software (<http://bioinfo.ut.ee/primer3/>) in order to design 100 to 200-bp amplicons (Table 1 shows the primer sequences). qRT-PCRs were performed in a 20 µl reaction volume containing 40 ng of cDNA, 0.5 µl of gene-specific primers (10 µM), and 1 X of LightCycler® 480 SYBR Green I mix (Roche®). PCR amplification, detection, and analysis were performed with the LightCycler® 480 PCR Detection System and the LightCycler® 480 Software

(Roche®). PCR conditions included an initial denaturation step at 95°C for 5 min, followed by 40-cycle amplification (95°C for 10 s, 60°C for 20s, 72°C for 20 s). The specificity of the amplified product and the absence of primer dimer formation were verified by generating a melting curve (65-98°C in continuous). The crossing point (C_p) was defined for each sample. The expression levels of the tested genes were normalized using the *gyrB* (primers OLM021 and OLM022) or *recA* (primers OLM321 and OLM322) genes of *S. agalactiae* as an internal standard whose transcript level did not vary under our experimental conditions (Table 1 shows the primer sequences). The change (n -fold) in the transcript level was calculated using the following equations: $\Delta C_p = C_{p(\text{target gene})} - C_{p(\text{reference gene, recA and gyrB})}$, $\Delta\Delta C_p = \Delta C_{p(\text{reference condition})} - \Delta C_{p(\text{tested condition})}$, and $\text{ratio} = 2^{-\Delta\Delta C_p}$. Each assay was performed in duplicate and repeated with at least three independent RNA samples.

5'-RACE PCR

S. agalactiae A909 strain was grown in zinc-restricted CDM at 37°C without agitation until the mid-exponential phase ($OD_{600nm} = 0.5$) of growth. Total RNAs of growing cells were extracted as described above and the 5'-end of the *lmb-sht* operon mRNA was determined using the 5'/3' rapid amplification of cDNA ends (RACE) PCR kit, second generation (Roche Applied Science®). Two *lmb* antisense specific primers (OAH068 and OAH069) were designed (Table 1) to produce cDNAs. Control RT-PCRs, omitting reverse transcriptase, were performed to check for DNA contamination of the RNA preparation with appropriate primers (Table 1).

Construction of *lmb*, *adcA*, *adcAll*, *adcCB* and *adcR* deletion mutants

S. agalactiae A909 Δ *lmb* is a non-polar mutant of strain A909 deleted of the entire coding sequence of *lmb* (*sak_1319*) by allelic exchange. Upstream and

downstream flanking regions of *lmb* were amplified by PCR with primers OAH024/OAH025 and OAH026/OAH027, respectively. A recombination cassette, consisting of a fusion between these two regions was obtained using splicing-by-overlap-extension PCR with primers OAH024 and OAH027 (Table 1). To carry out chromosomal gene inactivation, appropriate PCR fragments (overlap-extension) were cloned into the *EcoRI/BamHI* restriction sites of the thermosensitive shuttle plasmid pG+host1 (Table S1).

The same cloning strategy was applied to obtain the $\Delta adcA$ (*sak_0685*), the $\Delta adcAll$ (*sak_1898*), the $\Delta adcCB$ (*sak_0218* and *_0219*) and the $\Delta adcR$ (*sak_0217*) mutant strains except that fusion between upstream and downstream regions of *adcAll* were obtained using the *PstI* restriction site. Upstream and downstream regions of *adcA* were amplified by PCR with primers OAH086/OAH076 and OAH077/OAH087, respectively. Upstream and downstream regions of *adcAll* were amplified by PCR with primers OAH040/OAH062 and OAH063/OAH043, respectively. Upstream and downstream regions of *adcR* were amplified by PCR with primers OAH003/OAH004 and OAH005/OAH006, respectively. Upstream and downstream regions of *adcCB* were amplified by PCR with primers OAH074/OAH015 and OAH16/OAH075, respectively.

The recombinant plasmids were electroporated in *E. coli* for amplification, purified and then electroporated in strain A909 (Table S1). Allelic exchange was performed as previously described (33). Deletions of genes were confirmed by PCR and sequencing.

Construction of *lmb*, *adcA* and *adcAll* complementation

To complemente the Δlmb mutant strain, the entire coding sequence of *lmb* was amplified by PCR with primers OAH115 and OAH102. The PCR fragment was

cloned into the *Pst*//*Bgl*III restriction sites of the pTCV-P_{Tet} plasmid, a derivative of the shuttle vector pTCV, carrying a constitutively expressed Gram-positive promoter sequence (34; Table S1).

In the same way, the Δ *adcA* mutant strain was complemented using primers OAH120 and OAH121 and cloned with the *Xba*I//*Bgl*III restriction sites.

The Δ *adcAII* mutant strain was complemented using primers OAH119 and OAH104 and cloned with the *Bgl*III//*Pst*I restriction sites.

The recombined pTCV-P_{Tet} plasmids were electropored in *E. coli* for amplification, purified, and gene sequences were confirmed by sequencing before electroporation in A909 strain.

Construction of *lacZ* transcriptional fusions

Plasmid pTCV-*lac* that carries a promoterless *lacZ* gene (35) was used to construct transcriptional *lacZ* reporter fusions (Table S1).

Promoter region of *lmb* was amplified by PCR using primers OAH122 and OAH009 and the PCR fragment was cloned into the *Eco*RI//*Bam*HI restriction sites of the plasmid pTCV-*lac* (Table 1).

Sequence of the conserved 10-pb palindromic motif potentially constituting the AdcR binding site (AdcR box) was obtained from <http://regprecise.lbl.gov>. Site-tagged mutagenesis was performed to obtain the *adcR box*₁* and *adcR box*₂* coding sequences, in which 3 nucleotides for *adcR box*₁* and 4 nucleotides *adcR box*₂* were replaced so as to destroy the palindromic operator sequence without disturbing RNA polymerase fixation. For each construction, two DNA fragments were generated by PCR using primers OAH122/OAH034 and OAH035/OAH009 for the *adcR box*₁* mutation or OAH122/OAH022 and OAH023/OAH009 for the *adcR box*₂* mutation. The point mutations were obtained using oligonucleotides containing 3 or 4

mismatches (OAH035 and OAH023) and also carrying *BsaI* restriction sites, a type IIS restriction endonuclease, which cleaves after its restriction site, generating DNA fragments with tetranucleotide cohesive ends (Table 1). Following digestion with *BsaI* (2h at 50°C), the two fragments were purified with a NucleoSpin® PCR clean-up kit (Macherey-Nagel®) and ligated using the sticky-end instant ligase (NEB®), seamlessly fusing the fragments together without adding any nucleotides. The resulting fragments were re-amplified by PCR using the external oligonucleotides OAH122 and OAH009 (Table 1) and cloned in the pTCV-*lac* plasmid.

Electroporation

Electrocompetent *E. coli* and *S. agalactiae* cells were produced as previously described (36, 37). Both bacterial species were then transformed by electroporation using the Micropulser™ (Biorad®) and the Ec2 conditions (2.5 kV) with 1 to 2 µg of appropriate plasmids, respectively. Transformants were selected on LB-agar (*E. coli*) or on 5% horse blood tryptic soy (TSH) agar plates (BioMérieux) (*S. agalactiae*), with the appropriate antibiotics.

β-galactosidase assays

For β-galactosidase assays, bacteria were grown in CDM and harvested (10 ml samples) during the mid-exponential phase of growth ($OD_{600nm} = 0.5$). Cells were resuspended in 500 µl of Z buffer (38), lysed mechanically with glass beads in a FastPrep®-24 instrument, and cell debris were eliminated by centrifugation (5 min; 8000 g). Supernatant was incubated with 0.5 mg/ml DNase I (NEB®) at 37°C for 30 min and used for assays. Assays were performed as previously described and β-galactosidase specific activities were expressed as arbitrary units/mg protein (38).

Protein concentrations were determined using the Bio-Rad protein assay (Bio-Rad®, Hercules, CA). All experiments were carried out in triplicate.

Cell size and quantification of chain length

Bacteria were grown in CDM containing 1 or 10 μM of zinc. At mid-exponential phase of growth, bacteria were deposited onto glass slides and examined in bright field condition at X1000 magnification with the Nikon® eclipse 80i optical microscope. Image acquisition and processing were performed using the NIS-Elements D software. Cell size (μm^2) of 100 bacteria from three different pictures of three independent experiments was measured. To determine the bacteria chain length 100 chains of three different pictures of three independent experiments were counted. Chain length values were calculated as percentage of all counted chains.

Competition assays in human biological fluids

To easily discriminate the wild type and mutant strains in mixed cultures, we used the pTCV-*lac* vector which carries a promoterless *lacZ* gene or the pTCV-*lac*-P_{Cyl} vector with a *lacZ* gene under the control of the strong and constitutively active promoter P_{Cyl} (34; Table S1). When plated on TH agar containing erythromycin (10 $\mu\text{g}/\text{ml}$) and X-Gal (60 $\mu\text{g}/\text{ml}$), cells carrying pTCV-*lac*-P_{Cyl} vector appeared as blue colonies while cells carrying pTCV-*lac* remained white.

Each strain was transformed with pTCV-*lac* and pTCV-*lac*-P_{Cyl} plasmids. Mixed cultures were inoculated alternately with wild type strain carrying pTCV-*lac* / mutant strain carrying pTCV-*lac*-P_{Cyl} combination or wild type carrying pTCV-*lac*-P_{Cyl} / mutant strain carrying pTCV-*lac* to ensure that plasmids had no deleterious effect on bacterial growth. All strains were grown during 8 hours in TH with erythromycin 10 $\mu\text{g}/\text{ml}$, then in zinc-restricted CDM with erythromycin 10 $\mu\text{g}/\text{ml}$ overnight, and mixed cultures were inoculated with approximately 5.10^5 CFU/ml of each strain in human

biological fluids supplemented with erythromycin 10 µg/ml. Bacterial growth was monitored during 48 hours and wild type and mutant strains were discriminated by plating culture dilutions on TH agar containing erythromycin (10 µg/ml) and X-Gal (60 µg/ml). For each competition experiment between the WT and a mutant strain, both combination (WT pTCV-*lac* / Δ pTCV-*lac*-P_{Cyl} or WT pTCV-*lac*-P_{Cyl} / Δ pTCV-*lac*) were tested. Each experiment was repeated at least three times.

Human plasma was obtained from healthy donors of the Etablissement Français du Sang (EFS Centre Atlantique, France). It was decomplexed by heating at 56°C during 30 min. Amniotic fluid and cerebrospinal Fluid (CSF) were collected in the Bretonneau University Hospital (Tours, France) from respectively 2 and 3 donors who were not taking medications that would influence the analysis, such as antimicrobial agents.

Macrophage survival assays

Bacteria were grown in zinc-restricted CDM overnight. Cultures were washed in PBS and adjusted to the desired inoculum in RPMI 1640 medium (Gibco®), and CFU counts were verified by plating serial dilutions on TH plates. Macrophages RAW 264.7 cells, grown to confluence in RPMI-10% Fetal Calf Serum (Gibco®), were counted and incubated with bacteria (multiplicity of infection, MOI ~10) in RPMI 1640 at 37°C with 5% CO₂ for 1 h to allow bacterial phagocytosis. They were then incubated in RPMI-10% Fetal Calf Serum-gentamycin (500 µg/ml) / streptomycin (100 µg/ml) / penicillin (100 U/ml) in order to kill extracellular bacteria during 2 hours, which represent the time zero (T0) of the assay. At the indicated times (6 and 24 hours after T0), infected macrophages were washed once with RPMI and then lysed by incubation in 1 ml of ice-cold MilliQ water for 30 min. CFU counts were determined

311 by plating serial dilutions on TH plates. Assays were performed in duplicate and were
312 repeated three times.

313 **Statistical analyses**

314 Data are presented as the mean $\pm$ standard error of the mean. Statistical analyses
315 were performed using the unpaired Student's *t* test. A probability value of less than
316 0.05 was considered statistically significant.

317

318

RESULTS

Two homologs of the Lmb protein are present in *Streptococcus agalactiae*

Lmb is located on a 16-kb composite transposon containing the *scpB*, *lmb*, and *sht* genes (11, 12). We performed BlastN searches on available genomic DNA sequences of *S. agalactiae* (n = 271), and identified the presence of the *lmb*-containing transposon in 96.8 % of the 187 human isolates and only 26.7 % of the 84 animal isolates (mainly fish and cattle), in agreement with previous observations of 30 human and 38 bovine strains (11).

We performed BlastP searches on the *S. agalactiae* A909 genome, and identified two putative homologs of Lmb (SAK_1319). We named them AdcA (SAK_0685) and AdcAll (SAK_1898) based on sequence and functional homology with *S. pneumoniae* proteins. Lmb and AdcAll proteins are very similar, sharing 58% identity. The *lmb* and *adcAll* genes exhibit the same transcriptional organization as they are each co-transcribed with a gene encoding the histidine-triad proteins Sht (SAK_1318), and ShtII (SAK_1897), respectively (Fig.1). The AdcA protein shares only 34 % identity with Lmb, and its corresponding gene is monocistronic (Fig.1). In contrast to *lmb*, which is more specific to human isolates, the *adcA* and *adcAll* genes are highly conserved (at least 98 % identity) in all of the 271 available *S. agalactiae* genomes.

The AdcA/Lmb family proteins are conserved among most Streptococci and share a high degree of homology with zinc-binding proteins, in particular the AdcA and AdcAll proteins of *S. pneumoniae* (Table 2), whose role in zinc transport has been extensively studied (16, 39, 40). Following sequence alignment, we observed that the residues forming the Zn²⁺ ion containing binding site of Lmb (His66, His142, His206 and Glu281) are conserved in the AdcA and AdcAll proteins (18) (Fig. 2).

Analysis of the structure-based alignment between the proteins also revealed that AdcA possesses a histidine-rich loop and an extended C-terminal region exhibiting homology with the *Escherichia coli* ZinT protein and containing three conserved histidine residues that could potentially form a supplementary Zn-binding site (Fig. 2). The histidine-rich loop and ZinT domain of the *S. pneumoniae* AdcA homolog have been suggested to aid in recruiting Zn²⁺ (40). The structural differences between *S. agalactiae* AdcA and Lmb/AdcAll proteins are similar to those observed between *S. pneumoniae* AdcA and AdcAll proteins, and correspond to distinct zinc acquisition mechanisms (40, 41).

Regulation of the *lmb-sht*, *adcA*, and *adcAll-shtII* genes is Zn²⁺-dependent

Zinc-binding proteins in other bacteria are upregulated under zinc limitation (26). We therefore examined *lmb-sht* expression in the presence of various concentrations of extracellular zinc, using a transcriptional *lacZ* fusion approach. We grew the *S. agalactiae* A909 strain in a metal-restricted, chemically defined medium (CDM) containing no zinc in its formulation and named hereafter zinc-restricted CDM. This medium was supplemented with up to 100 μM Zn²⁺. *S. agalactiae* A909 grew equally under all conditions (Fig. S1). We measured *lmb-sht* promoter activity during mid-exponential growth as there were no significant growth-phase dependent differences under all tested conditions (data not shown). In contrast, β-galactosidase activity was significantly and gradually repressed in the presence of 0.5 to 100 μM Zn²⁺ ($P<0.01$; Fig. 3A); *lmb-sht* promoter activity was totally repressed in the presence of 10 μM added Zn²⁺ (approximately 20 fold less activity than in zinc-restricted CDM).It was also fully repressed in rich medium (Todd Hewitt or Brain Heart Infusion), in which the Zn²⁺ concentration is approximately 20 μM (data not shown). The repression of *lmb* expression by zinc appears to be specific to this metal

as supplementation with various concentrations of Mn^{2+} , Cu^{2+} , Ni^{2+} , Fe^{2+} , or Co^{2+} did not changed *Imb-sht* promoter activity (Fig. 3A).

Using qRT-PCR, we confirmed the strong repression of *Imb* in the presence of 10 μM extracellular Zn^{2+} , as well as that of the *adcA* and *adcAll* genes, suggesting that these genes are co-regulated (Fig. 3B). The *Imb* gene also appears to be the most highly expressed gene relative to *adcA* and *adcAll* when bacteria were grown in zinc-restricted CDM, but the difference was significant only with *adcA* (about 4 more highly expressed, $P < 0.05$, Fig. 3B).

Zn^{2+} -repression of the *Imb-sht*, *adcA*, and *adcAll-shtII* genes is mediated by AdcR

We have shown that the expression of the *Imb*, *adcA*, and *adcAll* genes is zinc-dependent. Zn^{2+} -dependent regulation is mediated in *S. pneumoniae* and *S. pyogenes* by the AdcR regulator (26, 27, 42). *S. pneumoniae* AdcR has a unique effector in zinc and its crystal structure revealed the presence of two metal-binding sites (43). A putative AdcR protein (SAK_0217; 53% identity with its pneumococcal homolog) containing the conserved metal-binding residues, is also present in *S. agalactiae* (Fig. S2). The AdcR encoding gene is present in all *S. agalactiae* sequenced genomes within an *adcRCB* operon (Fig. 1). We obtained a mutant of the *adcR* gene ($\Delta adcR$) in *S. agalactiae* A909 and measured the effect of this mutation on *Imb-sht* promoter activity using the β -galactosidase assay. In the absence of AdcR, *Imb-sht* promoter activity was no longer repressed by zinc (Fig. 4A). The same results were confirmed for *Imb* and observed for the *adcA*, and *adcAll* genes by qRT-PCR (Fig. 4B).

Our results indicate that the expression of the three putative zinc-binding proteins is controlled by AdcR. We thus analyzed their promoter regions to identify

AdcR operator sites (TTAACNNGTTAA) (23, 42). One conserved 10-bp palindromic motif is present in the promoter region of the *adcAll-shtII* and *adcRCB* operons and two overlapping copies were also present upstream from the *lmb-sht* and *adcA* genes (Fig. 1). We performed RACE-PCR to identify the transcription start site of the *lmb-sht* operon. We found it to be preceded by appropriately spaced -10 and -35 promoter regions which are included in the first putative AdcR binding motif (AdcR box) (Fig. 5). We introduced point mutations, designed to destruct the palindromic operator sequence without disturbing RNA polymerase fixation, by PCR-mediated site-tagged mutagenesis, into each or both putative AdcR boxes. Mutation of either AdcR box led to weaker repression of *lmb-sht* promoter activity by zinc, indicating that both boxes are necessary for optimal AdcR fixation (Fig. 5). When both boxes were mutated, *lmb-sht* promoter activity was completely derepressed in the presence of zinc, as observed in the Δ *adcR* strain (Fig. 5). Altogether, these results strongly suggest that Zn²⁺-dependent repression of the *lmb-sht* promoter is mediated by the AdcR regulator, which binds to the TTAACNNGTTAA target sequence in the presence of zinc.

We used this sequence to identify potential AdcR-regulated genes in *S. agalactiae* using the Search Pattern function of SagaList software (<http://genolist.pasteur.fr/SagaList>). In addition to genes encoding the Adc/Lmb system, we found a perfect AdcR box in the promoter region (400 bp upstream from the start codon) of 7 supplementary genes or operons (Table 3); in particular, upstream of the *adh* gene (*sak_0087*), encoding a potential zinc-containing alcohol dehydrogenase, which has already been identified to be a direct binding target of AdcR in *S. pneumoniae* (26, 42).

Lmb, AdcA, and AdcAll are involved in zinc acquisition

We constructed mutant strains containing single and combined deletions of each gene encoding a putative zinc-binding protein in *S. agalactiae* A909 to evaluate the relative contribution of Lmb, AdcA, and AdcAll in zinc transport. Bacteria in which the *lmb*, *adcA*, and *adcAll* genes were fully expressed were grown in zinc-restricted CDM (Fig. 3B). The growth of all single and double mutant strains was similar to that of the wild type strain (Fig. S3). In contrast, we observed almost no growth for the triple mutant of the binding proteins ($\Delta lmb/adcA/adcAll$) (Fig. 6A). We cloned the *lmb*, *adcA*, and *adcAll* genes downstream of the constitutive promoter P_{Tet} (34) to generate complementing vectors (Table S1). We introduced each construction in the $\Delta lmb/adcA/adcAll$ triple mutant strain and complemented bacteria were grown in zinc-restricted CDM. Recombinant plasmids expressing any of the three proteins fully restored the growth of the $\Delta lmb/adcA/adcAll$ triple mutant, suggesting that the three proteins are redundant as zinc suppliers (Fig. S4).

We then added increasing concentrations of Zn²⁺, from 0.5 to 100 μ M, to the medium. We observed no difference in growth for the wild type strain (Fig. S1) but the growth of the $\Delta lmb/adcA/adcAll$ mutant strain was gradually restored by Zn²⁺ addition reaching a maximum restored growth rate from a concentration of 10 μ M (Fig. 6A). The addition of 10 or 100 μ M Mn²⁺, Cu²⁺, Ni²⁺, Fe²⁺, Mg²⁺, or Co²⁺ had no impact on growth (Fig. S5), suggesting that the detrimental effect observed in the triple mutant is specifically due to zinc deprivation.

Lmb, AdcA, and AdcAll use the same AdcCB translocon

Lmb, AdcA, and AdcAll belong to the metal binding receptor class of proteins that classically associate with an AdcCB translocon, consisting of a permease, called AdcB (SAK_0219), and an ATPase, AdcC (SAK_0218), to form an ABC transporter. In all sequenced *S. agalactiae* strains, genes encoding a putative AdcCB translocon

are present within an *adcRCB* operon which lacks a substrate-binding encoding gene (Fig. 1). We constructed the mutant strain of the translocon $\Delta\textit{adcCB}$ to test whether Lmb, AdcA, and AdcAll all use the AdcCB proteins to import zinc. The $\Delta\textit{adcCB}$ mutant was unable to grow in zinc-restricted CDM, but its growth was gradually restored by the addition of Zn^{2+} (Fig. 6B). The growth of the $\Delta\textit{lmb/adcA/adcAll}$ triple mutant and that of the $\Delta\textit{adcCB}$ mutant were similar for all tested Zn^{2+} concentrations. This was also the case for the mutant strain of the entire putative transporter $\Delta\textit{adcCB/lmb/adcA/adcAll}$ (data not shown). Altogether, these results strongly suggest that the three binding proteins use the same translocon to transport zinc.

The absence of the Adc/Lmb zinc transporter affects cell morphology

In bacteria, several proteins involved in cell division, cell wall synthesis, and remodeling, including peptidoglycan deacetylases and metalloproteases, use zinc as a cofactor (21, 44). We examined whether the absence of the Lmb/AdcA/AdcAll zinc-binding proteins or the AdcCB translocon affects cell morphology and/or chain formation. We cultured *S. agalactiae* strains in zinc-restricted CDM with 1 or 10 μM of added zinc, examined them by optical microscopy, and calculated the size of individual bacteria and the number of cells per chain for each strain. We observed no significant difference between the individual cell size for any strain (data not shown). However, under low zinc concentrations (1 μM Zn^{2+}), the $\Delta\textit{lmb/adcA/adcAll}$ and $\Delta\textit{adcCB}$ mutant strains formed aggregates (clumps) or were mostly in the form of single cells, whereas the wild type strain formed chains of mostly 2 to 4 cells per chain (Fig. 7). In the presence of 10 μM of added zinc in the medium, the $\Delta\textit{lmb/adcA/adcAll}$ and $\Delta\textit{adcCB}$ strains regained the capacity to form chains with a size distribution comparable to that of the wild type strain, indicating that the observed abnormal cell morphology was due to zinc deficiency (Fig. 7).

The Adc/Lmb zinc transporter does not promote bacterial survival in murine macrophages

Phagocytic cells sequester iron to improve the clearance of pathogens, including bacteria (45). Other transition metals, specifically manganese and zinc, are also actively sequestered by the host during infection to hinder bacterial growth (46) and a Mn^{2+}/Fe^{2+} transporter has been shown to play a role in GBS intracellular survival inside macrophages (47). We compared the survival of the *S. agalactiae* wild type strain and its isogenic $\Delta lmb/adcA/adcAll$ mutant strain in the widely used murine macrophage cell line RAW 264.7. This phagocytic cell line efficiently killed *S. agalactiae* with both strains displaying similar survival kinetics (Fig. 8). This result indicates that the Adc/Lmb zinc transporter has no significant role in bacterial survival in macrophages.

The Adc/Lmb zinc transporter supports growth and survival of *Streptococcus agalactiae* in two human biological fluids

During infection, *S. agalactiae* has to cross or colonize several human anatomic sites that have variable zinc concentrations: approximately 1.5 μM in the amniotic fluid, more than 14 μM in plasma, and 2.3 μM in cerebrospinal fluid (CSF) (21, 48, 49). We thus examined the effect of the Adc/Lmb transporter on the ability of *the bacteria* to develop and survive in these human biological fluids. We performed a competition assay between the $\Delta lmb/adcA/adcAll$ or $\Delta adcCB$ mutant strains and their wild-type parent by inoculating human biological fluids (amniotic fluid, plasma, or CSF) with equal amounts and monitoring the proportion of each strain. Both mutant strains were clearly outcompeted by the wild-type strain in CSF and amniotic fluid. Indeed, the decrease of the mutant strains began during the growth phase and was accentuated during the persistence stage until they constituted less than 10% of the

global *S. agalactiae* population after 48 hours (Fig. 9). In contrast, the wild type/mutant strain ratio remained relatively unchanged after growth in plasma, suggesting that the Adc/Lmb zinc transporter is dispensable in this biological fluid (Fig. 9).

The *lmb* gene appears to be specific to human isolates. We thus performed a competition assay between the wild type and simple Δlmb mutant strain in human cerebrospinal and amniotic fluids. There was no significant difference in their growth (Fig. S6), indicating that the Lmb and AdcA/AdcAll proteins were also redundant in these human biological fluids.

DISCUSSION

In this study, we identified the Lmb/AdcA/AdcAll lipoproteins and the AdcCB translocon as components of a zinc transporter in *Streptococcus agalactiae*. We showed that expression of the *lmb*, *adcA*, and *adcAll* genes is controlled by the zinc concentration of the medium through the zinc-dependent regulator AdcR. Bacteria were unable to grow in zinc-restricted chemically defined medium in the absence of the Adc/Lmb transporter. To our knowledge, this is the first time that this metal has been shown to be essential for optimal growth of *S. agalactiae*. Zinc is a cofactor for a number of prokaryotic enzymes. In bacteria, there are several important enzymes that are zinc-containing proteins including alcohol dehydrogenase, peptidoglycan deacetylase, metalloprotease, or ribosomal proteins (21, 44). Homologs of these proteins are present in *S. agalactiae* and are also likely to require zinc for their proper activity. Thus, a severe zinc restriction could impair the structure and function of essential proteins, explaining the growth delay observed for the Adc/Lmb zinc-transporter mutant strains and the abnormal cell division. Indeed, the $\Delta lmb/adcA/adcAll$ and $\Delta adcCB$ mutant strains only formed aggregates or very short chains under zinc restriction conditions. The origin of this defect is still unknown and may be related to mis-metallation or under-metallation of key enzyme(s) involved in cell division. For example, the L,D-carboxypeptidase DacB of *Streptococcus pneumoniae*, which is essential for peptidoglycan turnover and crucial to preserve cell shape, contains a mononuclear Zn^{2+} catalytic centre (50). An homolog of this enzyme is present in *S. agalactiae*, and an alteration of their activity may explain part of the division phenotype and of the observed growth deficiency.

Bacteria without the Adc/Lmb transporter grew similarly to the wild type strain when 10 μM zinc was added to the medium, suggesting the presence of one or

several secondary zinc transporters with lower affinity. A good candidate may be the MtsA manganese-transporter, of which the gene is present in the *S. agalactiae* genome, and which has been shown to bind zinc with low affinity in *S. pneumoniae* (51, 52).

We showed that the growth and survival of $\Delta lmb/adcA/adcAll$ and $\Delta adcCB$ mutant strains in human cerebrospinal and amniotic fluids were clearly impaired. These biological fluids represent potential infection sites. In contrast, we observed no significant role of the zinc transporter in human plasma. The variable importance of the Adc/Lmb transporter, depending on the human biological fluid, could be easily correlated with their zinc content. Indeed, the zinc concentration in plasma varies between 12 and 16 μM , which is much higher than that usually observed in amniotic and cerebrospinal fluids (approximately 1.5 and 2.3 μM , respectively). Thus, the Adc/Lmb zinc transporter may be dispensable in plasma because zinc concentrations are high enough to allow import by secondary transporters. Moreover, recent studies pointed out that vertebrates could sequester zinc upon bacterial infection, in particular by the action of the human calprotectin released from neutrophils (46). This strategy has been shown to efficiently inhibit microbial growth inside abscessed tissues and contributes to host defense by rendering bacterial pathogens more sensitive to host immune effectors (53, 54). Thus, *in vivo*, available zinc concentrations in a defined host compartment could vary during infection. In these conditions, the presence of a high-affinity zinc uptake system as the Adc/Lmb transporter of *S. agalactiae* could be essential for the bacteria and open interesting perspectives for future studies.

Our results support the concept that the Adc/Lmb zinc-transporter is composed of three binding proteins Lmb, AdcA, and AdcAll that all use the same

553 AdcCB translocon. This multiplicity of substrate binding proteins is common amongst
554 ABC transporters. In the three pathogenic Streptococci, *S. pneumoniae*, *S.*
555 *pyogenes*, and *S. gordonii*, the Adc system is composed of one AdcCB translocon
556 and two zinc-binding proteins, the genes of which are found either within an
557 *adcRCBA* operon or independently of the *adc* cluster (21). In *S. pneumoniae*, either
558 AdcA or AdcAll are sufficient for zinc acquisition during growth *in vitro* and for
559 systemic virulence *in vivo*, but both are necessary for optimal nasopharynx
560 colonization (40). A recent study, again in *S. pneumoniae*, showed that the absence
561 of AdcAll, but not AdcA, negatively affected early colonization of the nasopharynx
562 (55). These results suggest that during infection, *S. pneumoniae* encounters
563 environments with various zinc concentrations and that, in specific niches and
564 conditions, the bacteria benefit from having two distinct zinc-binding proteins. The
565 particularity of the *S. agalactiae* Adc/Lmb transporter is the presence of an additional
566 binding protein, Lmb, which is almost specific to human isolates. In *S. agalactiae*,
567 Lmb, AdcA and AdcAll are redundant for zinc acquisition in chemically defined
568 medium which is consistent with what was observed in *S. pneumoniae* (40). In
569 human biological fluids, bacterial growth was not affected by the loss of the Lmb
570 protein alone, whereas the $\Delta lmb/adcA/adcAll$ mutant strain was clearly outcompeted
571 by the wild type strain in cerebrospinal and amniotic fluids. Thus, the three *S.*
572 *agalactiae* zinc-binding proteins were redundant under the conditions we tested.
573 However, we cannot exclude that, in specific environments or when in competition
574 with other bacteria, Lmb, AdcA, and AdcAll could be complementary, and possessing
575 three copies of these zinc-binding proteins could give to *S. agalactiae* an advantage.
576 *In vivo* experiments would be useful to answer these questions.

Another explanation for the multiplicity of these proteins may be that they have a secondary function beyond their involvement in zinc transport. Indeed, Lmb is a laminin-binding adhesin (10, 18). We expected to observe a difference between the binding of wild type and $\Delta lmb/adcA/adcAll$ mutant strains to laminin because of the strong homology between Lmb and AdcAll (73% similarity) and because we determined conditions (zinc-restricted CDM) in which these proteins were optimally expressed. We performed binding assays to immobilized human laminin, fibrinogen, or collagen and did not observe any significant difference between the wild type strain, Δlmb , or $\Delta lmb/adcA/adcAll$ mutant strains after bacterial growth in zinc-restricted CDM or TH broth (data not shown). There was also no effect on bacterial adhesion to A549 human alveolar basal epithelial cells (data not shown). These results do not agree with previous observations (10, 18). A possible explanation may be that we did not use the same *S. agalactiae* genetic background that was used in previous studies and that, despite strong conservation of the Lmb sequence, the ability to bind laminin may be strain dependent. It may be informative to test purified recombinant Lmb, AdcA, and AdcAll proteins for their binding ability, as previously done for the Lmb protein of *S. agalactiae* or Lbp of *S. pyogenes* (15, 18, 56). However, a recent study on Lmb homologs of *S. pneumoniae* also failed to show any binding to laminin (55).

Both the *S. agalactiae* *lmb* and *adcAll* genes are encoded within an operon containing the *sht* and *shtII* genes. These genes encode polyhistidine-triad proteins and *S. agalactiae* Sht has been shown to promote complement degradation by binding to factor H (12). In *S. pneumoniae*, the Sht homologs, called Pht proteins, aid in zinc delivery to the ABC transporter substrate-binding protein AdcAll (40, 41), and play a role in pneumococcal adhesion to the respiratory epithelium (57). The

602 involvement of *S. agalactiae* Sht proteins in zinc transport, their interaction with the
603 Adc/Lmb system components, and their putative redundancy are now under study
604 and should provide new insights to better understand mechanisms of zinc acquisition
605 in *S. agalactiae*.

606 **ACKNOWLEDGEMENTS**

607 We are grateful to the Patrick Trieu-Cuot's team and in particular to Elise
608 Caliot and Shaynoor Dramsi for providing us pG+host1 and pTCV-*lac*. A special
609 thanks to Arnaud Firon for pTCV-P_{Tet} and pTCV-*lac*-P_{Cyl} plasmids and for helpful
610 discussions. We also thank Vanessa Rong and Daniel Niquet for technical
611 assistance, Julien Gaillard for assistance in microscopy and Eric Morello for providing
612 us the RAW 264.7 cell line.

613 Pauline Moulin received a PhD fellowship from the Région Centre.

614

615

616

617

618

REFERENCES

1. **Verani JR, McGee L, Schrag SJ.** 2010. Prevention of perinatal group B streptococcal disease--revised guidelines from CDC, 2010. *MMWR Recomm Rep* **59**:1–36.
2. **Edmond KM, Kortsalioudaki C, Scott S, Schrag SJ, Zaidi AKM, Cousens S, Heath PT.** 2012. Group B streptococcal disease in infants aged younger than 3 months: systematic review and meta-analysis. *Lancet* **379**:547–556.
3. **Phares CR, Lynfield R, Farley MM, Mohle-Boetani J, Harrison LH, Petit S, Craig AS, Schaffner W, Zansky SM, Gershman K, Stefonek KR, Albanese BA, Zell ER, Schuchat A, Schrag SJ, Active Bacterial Core surveillance/Emerging Infections Program Network.** 2008. Epidemiology of invasive group B streptococcal disease in the United States, 1999–2005. *JAMA* **299**:2056–2065.
4. **Tazi A, Bellais S, Tardieux I, Dramsi S, Trieu-Cuot P, Poyart C.** 2012. Group B *Streptococcus* surface proteins as major determinants for meningeal tropism. *Curr Opin Microbiol* **15**:44–49.
5. **Rosa-Fraile M, Dramsi S, Spellerberg B.** 2014. Group B streptococcal haemolysin and pigment, a tale of twins. *FEMS Microbiol Rev* **38**:932–946.
6. **Schubert A, Zakikhany K, Pietrocola G, Meinke A, Speziale P, Eikmanns BJ, Reinscheid DJ.** 2004. The fibrinogen receptor FbsA promotes adherence of *Streptococcus agalactiae* to human epithelial cells. *Infect Immun* **72**:6197–6205.
7. **Rosenau A, Martins K, Amor S, Gannier F, Lanotte P, van der Mee-Marquet N, Mereghetti L, Quentin R.** 2007. Evaluation of the ability of *Streptococcus agalactiae* strains isolated from genital and neonatal specimens to bind to human fibrinogen and correlation with characteristics of the *fbsA* and *fbsB* genes. *Infect Immun* **75**:1310–1317.
8. **Tazi A, Disson O, Bellais S, Bouaboud A, Dmytruk N, Dramsi S, Mistou M-Y, Khun H, Mechler C, Tardieux I, Trieu-Cuot P, Lecuit M, Poyart C.** 2010. The surface protein HvgA mediates group B streptococcus hypervirulence and meningeal tropism in neonates. *J Exp Med* **207**:2313–2322.
9. **Buscetta M, Firon A, Pietrocola G, Biondo C, Mancuso G, Midiri A, Romeo L, Galbo R, Venza M, Venza I, Kaminski P-A, Gominet M, Teti G, Speziale P, Trieu-Cuot P, Beninati C.** 2016. PbsP, a cell wall-anchored protein that binds plasminogen to promote hematogenous dissemination of group B *Streptococcus*. *Mol Microbiol* **101**:27–41.
10. **Spellerberg B, Rozdzinski E, Martin S, Weber-Heynemann J, Schnitzler N, Lütticken R, Podbielski A.** 1999. Lmb, a protein with similarities to the Lral adhesin family, mediates attachment of *Streptococcus agalactiae* to human laminin. *Infect Immun* **67**:871–878.
11. **Franken C, Haase G, Brandt C, Weber-Heynemann J, Martin S, Lämmle C, Podbielski A, Lütticken R, Spellerberg B.** 2001. Horizontal gene transfer and host specificity of beta-haemolytic streptococci: the role of a putative composite transposon containing *scpB* and *lmb*. *Mol Microbiol* **41**:925–935.
12. **Maruvada R, Prasad Rao NV, Rubens CE.** 2009. Acquisition of factor H by a novel surface protein on group B *Streptococcus* promotes complement degradation. *FASEB* **23**:3967–3977.

- 658 13. **Al Safadi R, Amor S, Hery-Arnaud G, Spellerberg B, Lanotte P, Mereghetti L, Gannier F,**
659 **Quentin R, Rosenau A.** 2010. Enhanced expression of *lmb* gene encoding laminin-binding
660 protein in *Streptococcus agalactiae* strains harboring IS1548 in *scpB-lmb* intergenic region. PloS
661 One **5**:e10794.
- 662 14. **Tenenbaum T, Spellerberg B, Adam R, Vogel M, Kim KS, Schroten H.** 2007. *Streptococcus*
663 *agalactiae* invasion of human brain microvascular endothelial cells is promoted by the laminin-
664 binding protein Lmb. Microbes Infect **9**:714–720.
- 665 15. **Linke C, Caradoc-Davies TT, Young PG, Proft T, Baker EN.** 2009. The laminin-binding protein
666 Lbp from *Streptococcus pyogenes* is a zinc receptor. J Bacteriol **191**:5814–5823.
- 667 16. **Bayle L, Chimalapati S, Schoehn G, Brown J, Vernet T, Durmort C.** 2011. Zinc uptake by
668 *Streptococcus pneumoniae* depends on both AdcA and AdcAll and is essential for normal
669 bacterial morphology and virulence. Mol Microbiol **82**:904–916.
- 670 17. **Ragunathan P, Spellerberg B, Ponnuraj K.** 2009. Structure of laminin-binding adhesin (Lmb)
671 from *Streptococcus agalactiae*. Acta Crystallogr D Biol Crystallogr **65**:1262–1269.
- 672 18. **Ragunathan P, Sridaran D, Weigel A, Shabayek S, Spellerberg B, Ponnuraj K.** 2013. Metal
673 binding is critical for the folding and function of laminin binding protein, Lmb of *Streptococcus*
674 *agalactiae*. PloS One **8**:e67517.
- 675 19. **Andreini C, Bertini I.** 2012. A bioinformatics view of zinc enzymes. J Inorg Biochem **111**:150–
676 156.
- 677 20. **Jacobsen FE, Kazmierczak KM, Lisher JP, Winkler ME, Giedroc DP.** 2011. Interplay
678 between manganese and zinc homeostasis in the human pathogen *Streptococcus pneumoniae*.
679 Metallomics **3**:38–41.
- 680 21. **Shafeeq S, Kuipers OP, Kloosterman TG.** 2013. The role of zinc in the interplay between
681 pathogenic streptococci and their hosts. Mol Microbiol **88**:1047–1057.
- 682 22. **Kloosterman TG, van der Kooi-Pol MM, Bijlsma JJE, Kuipers OP.** 2007. The novel
683 transcriptional regulator SczA mediates protection against Zn²⁺ stress by activation of the Zn²⁺-
684 resistance gene *czcD* in *Streptococcus pneumoniae*. Mol Microbiol **65**:1049–1063.
- 685 23. **Panina EM, Mironov AA, Gelfand MS.** 2003. Comparative genomics of bacterial zinc regulons:
686 enhanced ion transport, pathogenesis, and rearrangement of ribosomal proteins. Proc Natl Acad
687 Sci U S A **100**:9912–9917.
- 688 24. **Aranda J, Garrido ME, Cortés P, Llagostera M, Barbé J.** 2008. Analysis of the protective
689 capacity of three *Streptococcus suis* proteins induced under divalent-cation-limited conditions.
690 Infect Immun **76**:1590–1598.
- 691 25. **Ogunniyi AD, Grabowicz M, Mahdi LK, Cook J, Gordon DL, Sadlon TA, Paton JC.** 2009.
692 Pneumococcal histidine triad proteins are regulated by the Zn²⁺-dependent repressor AdcR and
693 inhibit complement deposition through the recruitment of complement factor H. FASEB **23**:731–
694 738.
- 695 26. **Shafeeq S, Kloosterman TG, Kuipers OP.** 2011. Transcriptional response of *Streptococcus*
696 *pneumoniae* to Zn²⁺ limitation and the repressor/activator function of AdcR. Metallomics **3**:609–
697 618.

- 698 27. **Sanson M, Makthal N, Flores AR, Olsen RJ, Musser JM, Kumaraswami M.** 2015. Adhesin
699 competence repressor (AdcR) from *Streptococcus pyogenes* controls adaptive responses to zinc
700 limitation and contributes to virulence. *Nucleic Acids Res* **43**:418–432.
- 701 28. **Loo CY, Mittrakul K, Voss IB, Hughes CV, Ganeshkumar N.** 2003. Involvement of the *adc*
702 operon and manganese homeostasis in *Streptococcus gordonii* biofilm formation. *J Bacteriol*
703 **185**:2887–2900.
- 704 29. **Weston BF, Brenot A, Caparon MG.** 2009. The metal homeostasis protein, Lsp, of
705 *Streptococcus pyogenes* is necessary for acquisition of zinc and virulence. *Infect Immun*
706 **77**:2840–2848.
- 707 30. **Tedde V, Rosini R, Galeotti CL.** 2016. Zn²⁺ Uptake in *Streptococcus pyogenes*:
708 Characterization of *adcA* and *Imb* Null Mutants. *PloS One* **11**:e0152835.
- 709 31. **Patron K, Gilot P, Camiade E, Mereghetti L.** 2015. An homolog of the Frz
710 Phosphoenolpyruvate:carbohydrate phosphoTransferase System of extraintestinal pathogenic
711 *Escherichia coli* is encoded on a genomic island in specific lineages of *Streptococcus agalactiae*.
712 *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis* **32**:44–50.
- 713 32. **Lamy M-C, Zouine M, Fert J, Vergassola M, Couve E, Pellegrini E, Glaser P, Kunst F,**
714 **Msadek T, Trieu-Cuot P, Poyart C.** 2004. CovS/CovR of group B streptococcus: a two-
715 component global regulatory system involved in virulence. *Mol Microbiol* **54**:1250–1268.
- 716 33. **Biswas I, Gruss A, Ehrlich SD, Maguin E.** 1993. High-efficiency gene inactivation and
717 replacement system for Gram-positive bacteria. *J Bacteriol* **175**:3628–3635.
- 718 34. **Firon A, Tazi A, Da Cunha V, Brinster S, Sauvage E, Dramsi S, Golenbock DT, Glaser P,**
719 **Poyart C, Trieu-Cuot P.** 2013. The Abi-domain protein Abx1 interacts with the CovS histidine
720 kinase to control virulence gene expression in group B Streptococcus. *PLoS Pathog*
721 **9**:e1003179.
- 722 35. **Poyart C, Trieu-Cuot P.** 1997. A broad-host-range mobilizable shuttle vector for the
723 construction of transcriptional fusions to beta-galactosidase in Gram-positive bacteria. *FEMS*
724 *Microbiol Lett* **156**:193–198.
- 725 36. **Dower WJ, Miller JF, Ragsdale CW.** 1988. High efficiency transformation of *E. coli* by high
726 voltage electroporation. *Nucleic Acids Res* **16**:6127–6145.
- 727 37. **Ricci ML, Manganelli R, Berneri C, Orefici G, Pozzi G.** 1994. Electrotransformation of
728 *Streptococcus agalactiae* with plasmid DNA. *FEMS Microbiol Lett* **119**:47–52.
- 729 38. **Miller JH.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory.
- 730 39. **Loisel E, Jacquamet L, Serre L, Bauvois C, Ferrer JL, Vernet T, Di Guilmi AM, Durmort C.**
731 2008. AdcAll, a new pneumococcal Zn-binding protein homologous with ABC transporters:
732 biochemical and structural analysis. *J Mol Biol* **381**:594–606.
- 733 40. **Plumptre CD, Eijkelkamp BA, Morey JR, Behr F, Couñago RM, Ogunniyi AD, Kobe B,**
734 **O'Mara ML, Paton JC, McDevitt CA.** 2014. AdcA and AdcAll employ distinct zinc acquisition
735 mechanisms and contribute additively to zinc homeostasis in *Streptococcus pneumoniae*. *Mol*
736 *Microbiol* **91**:834–851.

- 737 41. **Bersch B, Bougault C, Roux L, Favier A, Vernet T, Durmort C.** 2013. New insights into
738 histidine triad proteins: solution structure of a *Streptococcus pneumoniae* PhtD domain and zinc
739 transfer to AdcAll. *PLoS One* **8**:e81168.
- 740 42. **Reyes-Caballero H, Guerra AJ, Jacobsen FE, Kazmierczak KM, Cowart D, Koppolu UMK,**
741 **Scott RA, Winkler ME, Giedroc DP.** 2010. The metalloregulatory zinc site in *Streptococcus*
742 *pneumoniae* AdcR, a zinc-activated MarR family repressor. *J Mol Biol* **403**:197–216.
- 743 43. **Guerra AJ, Dann CE, Giedroc DP.** 2011. Crystal structure of the zinc-dependent MarR family
744 transcriptional regulator AdcR in the Zn(II)-bound state. *J Am Chem Soc* **133**:19614–19617.
- 745 44. **Sun X, Xiao C, Ge R, Yin X, Li H, Li N, Yang X, Zhu Y, He X, He Q-Y.** 2011. Putative copper-
746 and zinc-binding motifs in *Streptococcus pneumoniae* identified by immobilized metal affinity
747 chromatography and mass spectrometry. *Proteomics* **11**:3288–3298.
- 748 45. **Weinberg ED.** 2009. Iron availability and infection. *Biochim Biophys Acta* **1790**:600–605.
- 749 46. **Kehl-Fie TE, Skaar EP.** 2010. Nutritional immunity beyond iron: a role for manganese and zinc.
750 *Curr Opin Chem Biol* **14**:218–224.
- 751 47. **Shabayek S, Bauer R, Mauerer S, Mizaikoff B, Spellerberg B.** 2016. A streptococcal NRAMP
752 homologue is crucial for the survival of *Streptococcus agalactiae* under low pH conditions. *Mol*
753 *Microbiol* **100**:589–606.
- 754 48. **Meret S, Henkin RI.** 1971. Simultaneous direct estimation by atomic absorption
755 spectrophotometry of copper and zinc in serum, urine, and cerebrospinal fluid. *Clin Chem*
756 **17**:369–373.
- 757 49. **Tamura GS, Kuypers JM, Smith S, Raff H, Rubens CE.** 1994. Adherence of group B
758 streptococci to cultured epithelial cells: roles of environmental factors and bacterial surface
759 components. *Infect Immun* **62**:2450–2458.
- 760 50. **Abdullah MR, Gutiérrez-Fernández J, Pribyl T, Gisch N, Saleh M, Rohde M, Petruschka L,**
761 **Burchhardt G, Schwudke D, Hermoso JA, Hammerschmidt S.** 2014. Structure of the
762 pneumococcal l,d-carboxypeptidase DacB and pathophysiological effects of disabled cell wall
763 hydrolases DacA and DacB. *Mol Microbiol* **93**:1183–1206.
- 764 51. **Bray BA, Sutcliffe IC, Harrington DJ.** 2009. Expression of the MtsA lipoprotein of
765 *Streptococcus agalactiae* A909 is regulated by manganese and iron. *Antonie Van Leeuwenhoek*
766 **95**:101–109.
- 767 52. **McDevitt CA, Ogunniyi AD, Valkov E, Lawrence MC, Kobe B, McEwan AG, Paton JC.** 2011.
768 A molecular mechanism for bacterial susceptibility to zinc. *PLoS Pathog* **7**:e1002357.
- 769 53. **Corbin BD, Seeley EH, Raab A, Feldmann J, Miller MR, Torres VJ, Anderson KL, Dattilo**
770 **BM, Dunman PM, Gerads R, Caprioli RM, Nacken W, Chazin WJ, Skaar EP.** 2008. Metal
771 chelation and inhibition of bacterial growth in tissue abscesses. *Science* **319**:962–965.
- 772 54. **Kehl-Fie TE, Chitayat S, Hood MI, Damo S, Restrepo N, Garcia C, Munro KA, Chazin WJ,**
773 **Skaar EP.** 2011. Nutrient metal sequestration by calprotectin inhibits bacterial superoxide
774 defense, enhancing neutrophil killing of *Staphylococcus aureus*. *Cell Host Microbe* **10**:158–164.

55. **Brown LR, Gunnell SM, Cassella AN, Keller LE, Scherkenbach LA, Mann B, Brown MW, Hill R, Fitzkee NC, Rosch JW, Tuomanen EI, Thornton JA.** 2016. AdcAll of *Streptococcus pneumoniae* Affects Pneumococcal Invasiveness. *PLoS One* **11**:e0146785.
56. **Terao Y, Kawabata S, Kunitomo E, Nakagawa I, Hamada S.** 2002. Novel laminin-binding protein of *Streptococcus pyogenes*, Lbp, is involved in adhesion to epithelial cells. *Infect Immun* **70**:993–997.
57. **Kallio A, Sepponen K, Hermand P, Denoël P, Godfroid F, Melin M.** 2014. Role of Pht proteins in attachment of *Streptococcus pneumoniae* to respiratory epithelial cells. *Infect Immun* **82**:1683–1691.

REFERENCES RELATED TO SUPPLEMENTAL MATERIAL

1. **Biswas I, Gruss A, Ehrlich SD, Maguin E.** 1993. High-efficiency gene inactivation and replacement system for Gram-positive bacteria. *J Bacteriol* **175**:3628–3635.
2. **Firon A, Tazi A, Da Cunha V, Brinster S, Sauvage E, Dramsi S, Golenbock DT, Glaser P, Poyart C, Trieu-Cuot P.** 2013. The Abi-domain protein Abx1 interacts with the CovS histidine kinase to control virulence gene expression in group B *Streptococcus*. *PLoS Pathog* **9**:e1003179.
3. **Poyart C, Trieu-Cuot P.** 1997. A broad-host-range mobilizable shuttle vector for the construction of transcriptional fusions to beta-galactosidase in Gram-positive bacteria. *FEMS Microbiol Lett* **156**:193–198.
4. **Tettelin H, Massignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, Angiuoli SV, Crabtree J, Jones AL, Durkin AS, Deboy RT, Davidsen TM, Mora M, Scarselli M, Margarit y Ros I, Peterson JD, Hauser CR, Sundaram JP, Nelson WC, Madupu R, Brinkac LM, Dodson RJ, Rosovitz MJ, Sullivan SA, Daugherty SC, Haft DH, Selengut J, Gwinn ML, Zhou L, Zafar N, Khouri H, Radune D, Dimitrov G, Watkins K, O'Connor KJB, Smith S, Utterback TR, White O, Rubens CE, Grandi G, Madoff LC, Kasper DL, Telford JL, Wessels MR, Rappuoli R, Fraser CM.** 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial “pan-genome.” *Proc Natl Acad Sci USA* **102**:13950–13955.

804 **Table 1. Oligonucleotides used in this study**

Name	Sequence ^a	Name	Sequence ^a
Primers used for the deletion of <i>adcR</i>		Primers used for the complementation of <i>adcAll</i>	
OAH003	TTAATC GAATTC GGAATCTTCAAATGGACTATG	OAH119	AAAATA AGATCT GGAGGTTTAATGAAAAGG
OAH004	ATGATTACCCTTCTAGTTCTAAACTGTCATATATACCTC	OAH104	TTATAT CTGCAG CTATTGATTAAACGATTG
OAH005	TGACAGTTTTAGAACTAGAAGGGTAATCATGCGATAT	Primers used for the complementation of <i>adcA</i>	
OAH006	ACATTG GGATCCC GCCATGGTAAAGATTGGTTCC	OAH120	GTAAAA AGATCT GGAGGAAAAATGAGAAAG
Primers used for the deletion of <i>adcCB</i>		OAH121	ATCTTT CTAG ATTAATGAGACATAAGGTC
OAH074	ATGACAG GAATTC GAACAAAAATTAGACCATTTAGTGAG	Primers used for transcriptional fusion	
OAH015	TGCGAAGCAAGCCTACGCATGATTACCCTTCTAATTCTCT	OAH122	TAGTTAG GAATTC CTTTCTTCTTGGGATTAGTAGC
OAH016	AAGGGTAATCATGCGTAGGCTTGCTTCGCAAACGTTAA	OAH009	TTC GGATCCC CTCCTCAATTATAATTAAACCAGTTATTAAC
OAH075	ATACTT GGATCCC CAGATACAAATAATGTAGCTCCCC	OAH034	GGTCTC GCCAGTTAATTTTACTCCTTTATCAATGC
Primers used for the deletion of <i>lmb</i>		OAH035	GGTCTC GCTGGCCCATAACTGGTTAAATTATAATTGAGG
OAH024	AGGCTG GAATTC TGGAAGGCGCTACTGTTCC	OAH022	GGTCTC GCCCCCAGTTATTAACCAGTTAA
OAH025	CTCCTTTACTTCAACCCTTTTTTTCATAGTACCTCCTCAATT	OAH023	GGTCTC GGGGGATTATAATTGAGGAGGTACTATG
OAH026	TACTATGAAAAAAGGGTTGAAGTAAAGGAGATTATTAGTGAAG	Primers used for 5' RACE PCR	
OAH027	GGCTTG GGATCC AGCTAGCTCACTTGGAGAC	OAH068	GCACCTGATTGGATCATCCTCAC

Primers used for the deletion of *adcAll*

OAH040 AAAATC**GAATT**CCAACGTGTTAATCAAGCAAGTG
OAH062 TA**AGGTACCT**CCGTATCCTTTTCATTAAACCTCC
OAH063 TT**AGGTACCT**AGGTAGTTATATAAAGAAAGGACG
OAH043 TACTA**AGGATCC**ATCTCCCATGTCATCAATGAC

Primers used for the deletion of *adcA*

OAH086 AATCGT**GGTACC**GCAGCTCTAGCAGATCCACAC
OAH076 ATA**AAGCTT**GAAATTTCTTTCTCATTTTTTCTCC
OAH077 ATG**AAGCTT**CATTAATATTTAAAAGATGATATCGG
OAH087 CATCC**AGGATCCC**GTCCAGTTGTTTTCTTAGATAC

Primers used for the complementation of *Imb*

OAH115 TAAATT**AGATCT**GGAGGTACTATGAAAAA
OAH102 TAATA**ACTGCAGT**TACTTCAACTGTTGATAGAGC

OAH069 GCTGGTTACA**ACTGACATGCCTTG**

Primers used for qRT-PCR

OAH053 *Imb* TAGTAATGATAGCAGGGTGTGATAAGTC
OAH054 *Imb* AGATACTTCTTTTGTGTCATCGCATACATT
OAH047 *adcA* TTGGGATTACAAAGCTAAATCTAAAAA
OAH048 *adcA* ATCTTGATTGATTCTACGTCAGTCTTG
OAH051 *adcAll* GCATAAAATCATAGGAAAGCATATCAA
OAH052 *adcAll* GCTAAGTATGAGAATGCTGTATGTGAAG
OLM321 *recA* CTGGTGGTCGTGCTTTGAAA
OLM322 *recA* TATGCTCACCAGTCCCCTTG
OLM021 *gyrB* GCTCACATCAGAACTTTACTTTTAACTC
OLM022 *gyrB* TTTAATCTCACTTCCTACTTTGACACC

^a Added restriction site sequences are shown in bold

Table 2. Identity (%) between the Lmb, AdcA, and AdcAll proteins of *Streptococcus agalactiae* and the AdcA and AdcAll proteins of *Streptococcus pneumoniae*

S. <i>agalactiae</i> proteins	Identity with S. <i>pneumoniae</i> proteins	
	(%)	
	AdcA	AdcAll
Lmb	35	67
AdcA	62	37
AdcAll	37	61

Table 3. Identification of potential AdcR regulon members by *in silico* *Streptococcus agalactiae* A909 genome scanning

Position ^a	Gene ID and/or name	Function	Pattern sequence ^b
-126	<i>sak_0087; adh</i>	Alcohol dehydrogenase	ttaaag attaaccagtta agaata
-28	<i>sak_0217; adcR</i>	Repressor, zinc metabolism	tataa tttactggtta caaaa
-96	<i>sak_0252</i>	Binding protein, peptide/nickel ABC transporter	tcta attaaccagtta agtaat
-28 and -39	<i>sak_0685; adcA</i>	Binding protein, zinc ABC transporter	aaata ttaaccggtaaataacggg ttaaataaa
-50	<i>sak_0693</i>	Similar to surface proteins, (LPXTG motif)	aaaact taaccggtta attatt
-42	<i>sak_1023</i>	Similar to histidine triad protein, putative internalin	tataa ttaactagtta actaaa
-32 and -43	<i>sak_1319; lmb</i>	Binding protein, zinc ABC transporter	taaa attaactggtaataactg ttaaattat
-52	<i>sak_1468</i>	Similar to flavoprotein, involved in K ⁺ transport	tatt gtaactggtta agtatt
-39	<i>sak_1469</i>	Similar to ammonium transporter	aata cttaaccagtta acaata
-340	<i>sak_1542</i>	Binding protein, peptide/nickel ABC transporter	tata cttaactggtta agtata
-38	<i>sak_1898; adcAII</i>	Binding protein, zinc ABC transporter	aaat gtaactggtta agtatt

^a Nucleotides from the start codon of the genes to the end of the putative AdcR binding site

^b Nucleotides in bold indicate the putative AdcR binding site

FIGURES

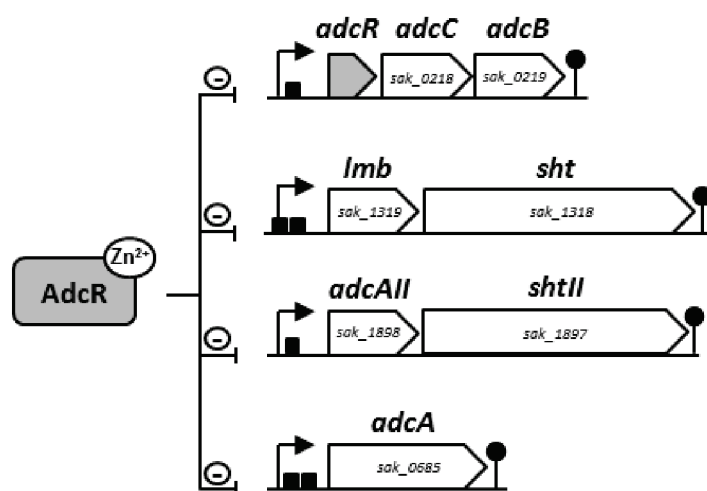


Figure 1. Transcriptional organization of the *lmb*, *adcA*, *adcAll*, and *adcRCB* genes and model of Zn-dependent regulation by AdcR in *Streptococcus agalactiae*

In the presence of Zn²⁺, AdcR can bind to the AdcR-binding motifs (TTAACNNGTTAA) located in the promoter regions of the *lmb-sht*, *adcAll-shtII*, and *adcRCB* operons and *adcA* gene, causing repression of their expression.

($\rightarrow$) Promoter; (●) Transcriptional terminator; (■) AdcR-binding box; (⊖) Repression

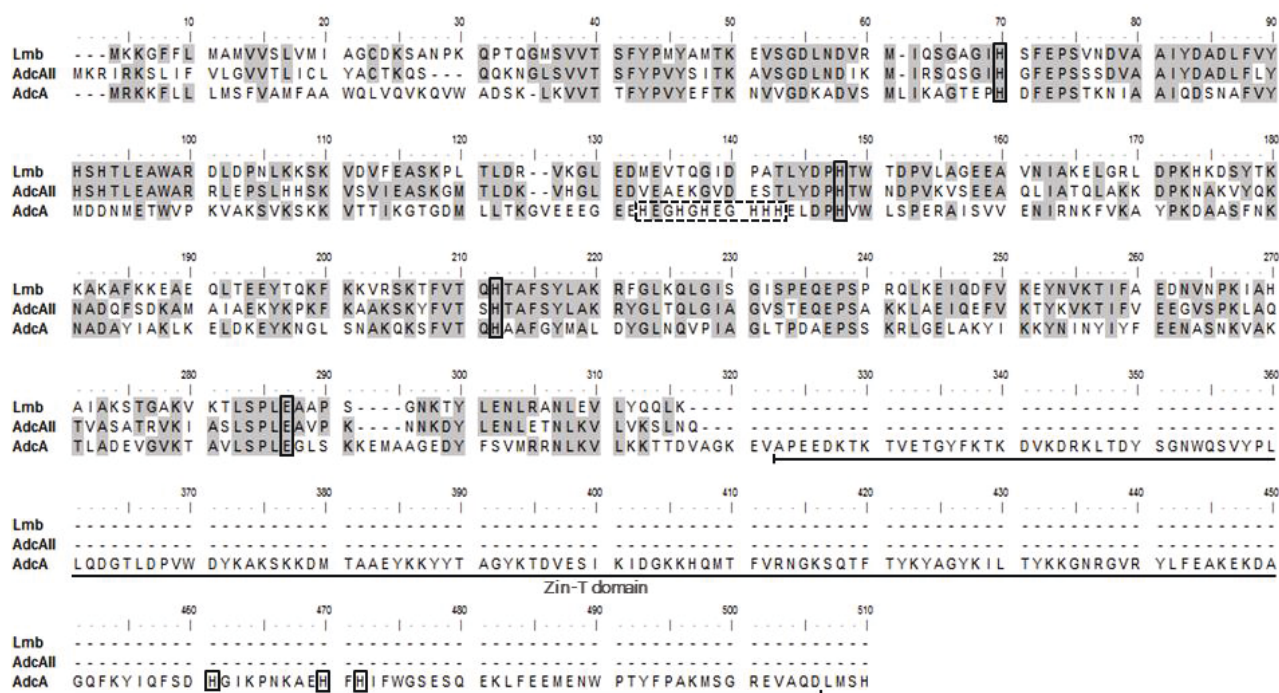


Figure 2. Amino acid sequence alignment of Lmb, AdcA, and AdcAll proteins of *Streptococcus agalactiae*

Identical residues are highlighted in grey. Potential zinc-binding residues are boxed in black. The His-rich loop of AdcA, which has been suggested to aid in recruiting Zn²⁺, is boxed by the black dotted lines. The Zin-T domain of AdcA, that may contain a supplementary zinc-binding site, is underlined. Sequence alignment was performed using the BioEdit program.

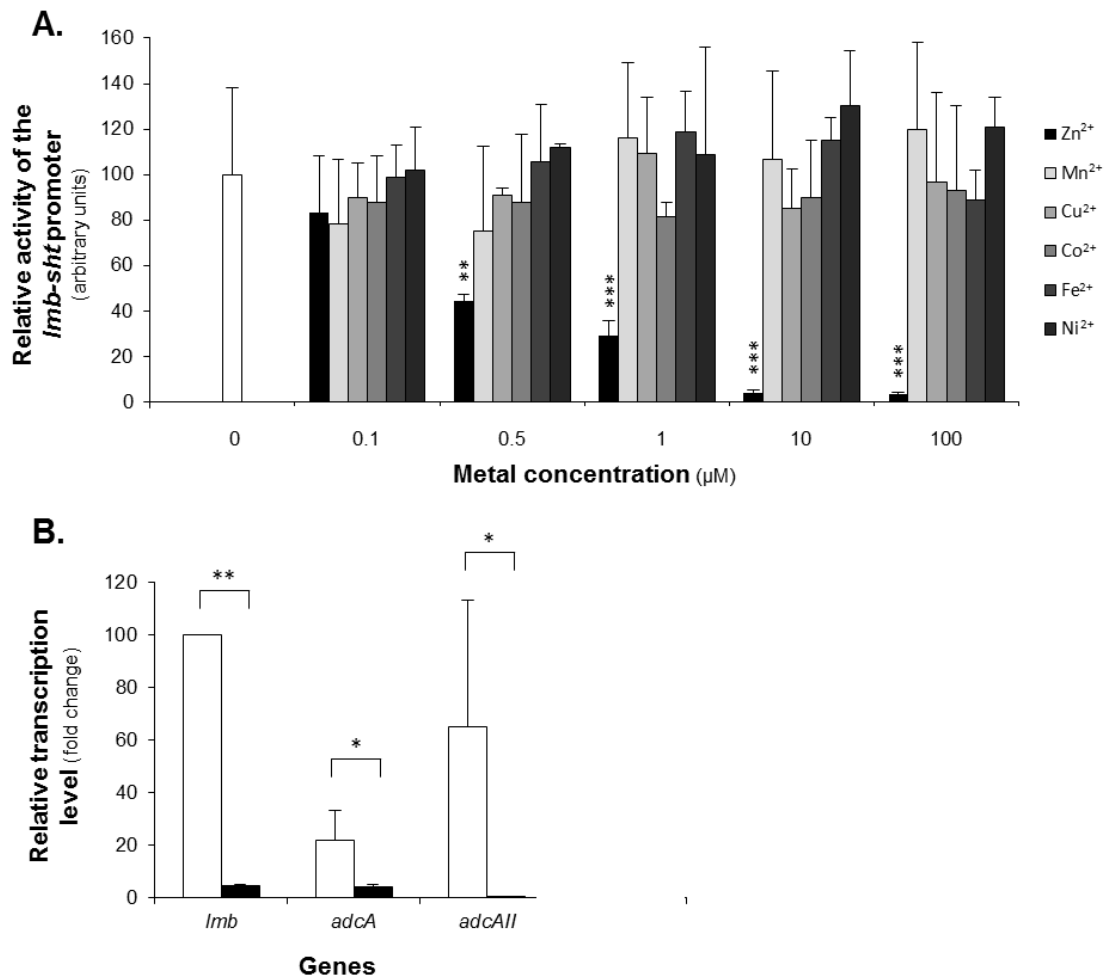


Figure 3. Regulation of the *Streptococcus agalactiae* *lmb*, *adcA*, and *adcAll* genes is Zn dependent

(A) *lmb-sht* promoter activity was measured in Zn-restricted CDM supplemented with various amounts of added Zn²⁺ (0 to 100 μM). A909 cells were grown until the mid-exponential phase of growth (OD_{600nm} = 0.5) and β-galactosidase assays were performed as described in experimental procedures. The reference value (100) is the *lmb-sht* promoter activity of cells grown in zinc-restricted CDM (0 μM Zn²⁺ added, white bar) and activities were calculated based on this reference. The values shown are presented as the mean ± standard deviation. The asterisks indicate *P*-values obtained using unpaired Student's *t*-test between promoter activity of WT cells grown in zinc-restricted CDM and in CDM with the various added Zn²⁺ concentrations (** *P*<0.01, *** *P*<0.001).

(B) *lmb*, *adcA*, and *adcAll* expression were measured in zinc-restricted CDM containing 0 (white bars) or 10 μM added Zn²⁺ (black bars). qRT-PCR was performed on RNA extracts of *S. agalactiae* A909 grown until the mid-exponential phase. The amount of transcripts of each gene was normalized against *recA* transcripts levels. The reference value (100) is the level of *lmb* transcripts after growth in zinc-restricted

CDM (0 μM added Zn^{2+}). Gene expression is presented as n-fold changes. Results are presented as the mean $\pm$ standard deviation of three independent experiments. The asterisks indicate *P*-values obtained using unpaired Student's *t*-test between gene expression of cells grown in zinc-restricted CDM and with 10 μM of added Zn^{2+} (* *P*<0.05; ** *P*<0.01).

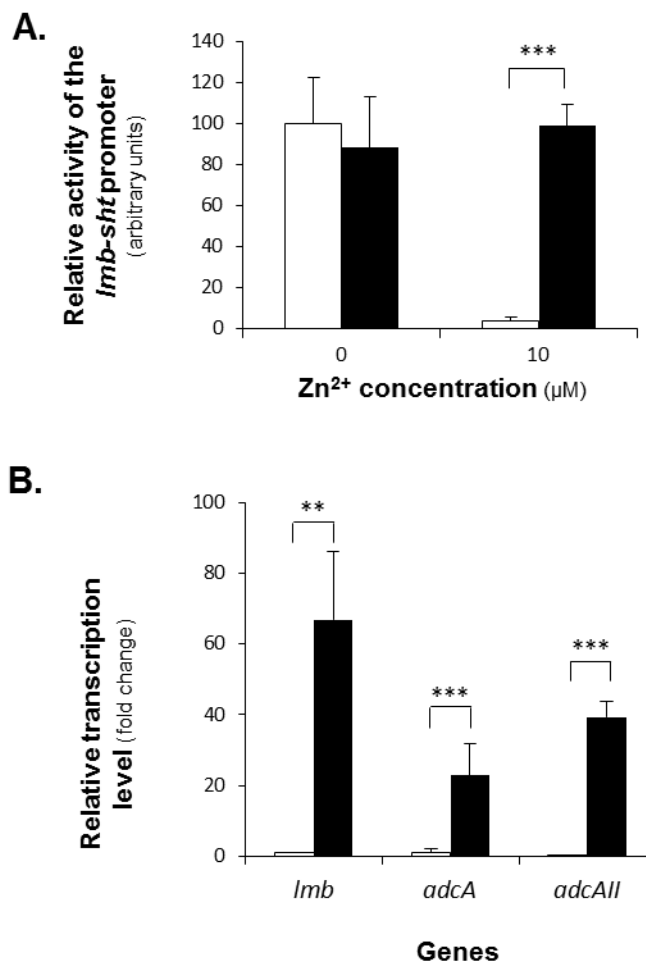


Figure 4. AdcR is required for Zn-dependent repression of *lmb*, *adcA* and *adcAll* genes.

(A) WT A909 (white bars) or its isogenic $\Delta adcR$ mutant (black bars) strains were grown either in zinc-restricted CDM or in medium supplemented with 10 μM Zn^{2+} until reaching the mid-exponential phase ($OD_{600nm} = 0.5$). β -galactosidase assays were performed as described in experimental procedures. The reference value (100) is the *lmb-sht* promoter activity of WT cells grown in zinc-restricted CDM and the results are presented as n-fold changes against this reference. The values shown are the means $\pm$ standard deviations of three independent assays.

(B) WT A909 (white bars) or its isogenic $\Delta adcR$ mutant (black bars) strains were grown in zinc-restricted CDM supplemented with 10 μM Zn^{2+} until reaching the mid-exponential phase ($OD_{600nm} = 0.5$). The amount of transcripts of each gene was normalized against *recA* transcripts levels. The reference value (1) is the level of *lmb* transcripts of the wild type strain. Gene expression is presented as n-fold changes. Results are presented as the mean $\pm$ standard deviation of three independent experiments.

The asterisks indicate *P*-values obtained using unpaired Student's *t*-test between promoter activity (A) or gene expression (B) of the WT strain and that of its isogenic $\Delta adcR$ mutant (* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$).

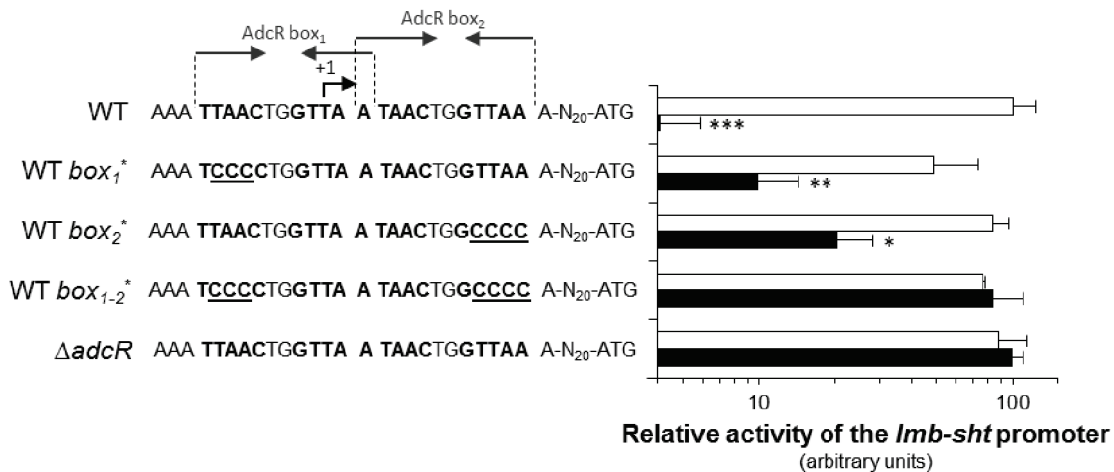


Figure 5. Both putative AdcR-boxes within the *lmb-sht* promoter are required for full Zn-dependent promoter repression

Transcriptional *lacZ* fusions with the *lmb-sht* promoter region carrying point mutations destroying an inverted repeat of the first putative AdcR-box (*box1*^{*}), or point mutations destroying an inverted repeat of the second putative AdcR-box (*box2*^{*}), or combined mutations (*box1-2*^{*}) were constructed. *lacZ* fusions containing the native *lmb-sht* promoter region and its derivatives were introduced in the WT strain. The native *lmb-sht* promoter region was also introduced in a Δ *adcR* mutant strain.

β -galactosidase assays were performed as described in experimental procedures. The relative activity of the promoters was measured in zinc-restricted CDM containing 0 (white bars) or 10 μ M added Zn²⁺ (black bars). The reference value (100) is the *lmb-sht* promoter activity of WT cells grown in zinc-restricted CDM and the results are presented as n-fold changes against this reference. The values shown are the means $\pm$ standard deviations of three independent assays.

The positions are numbered with respect to the start codon and the transcription initiation site is labeled +1 (↗). Less relevant nucleotides are replaced with N.

Nucleotides in bold represent the two putative AdcR-binding sites and underlined nucleotides indicate the introduced point mutations. The asterisks indicate *P*-values obtained using unpaired Student's *t*-test between promoter activity of the WT strain grown in zinc-restricted CDM and the other tested strains and conditions (* *P*<0.05; ** *P*<0.01; *** *P*<0.001).

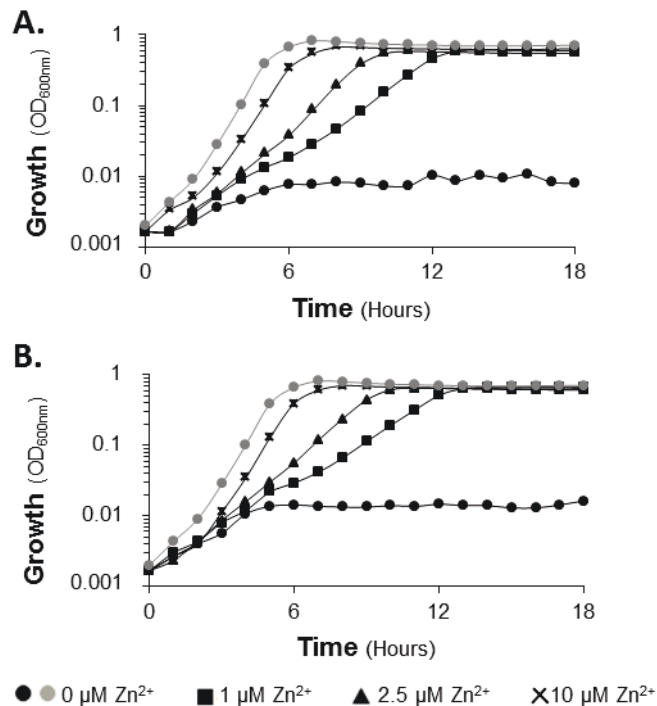


Figure 6. The Lmb, AdcA, and AdcAll proteins are involved in zinc acquisition in association with the AdcCB translocon

A909 cells **(A)** WT (in grey) and $\Delta lmb/adcA/adcAll$ mutant strains (in black) or **(B)** WT (in grey) and $\Delta adcCB$ mutant strains (in black), were grown in zinc-restricted CDM with various amounts of added Zn²⁺ (from 0 to 10 μM). The growth was monitored by assessing OD_{600nm} at 1 h intervals during 18 h. Data are representative mean OD_{600nm} measurements from three independent experiments.

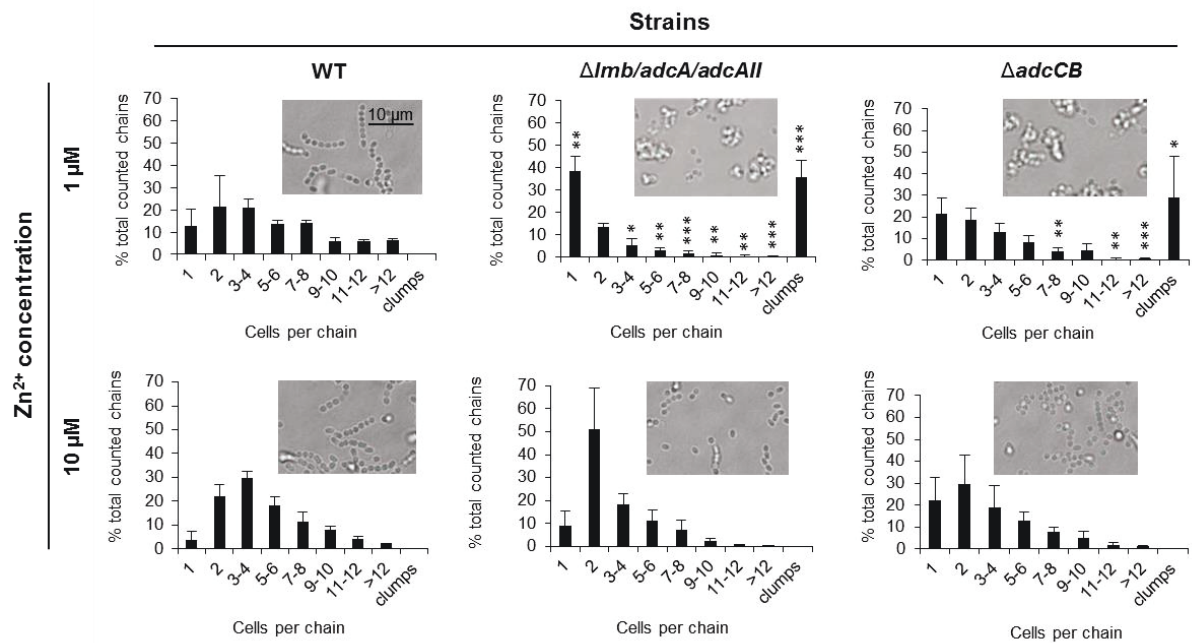


Figure 7. The Adc/Lmb zinc transporter affects cell morphology

Bacterial-chain length of the A909 wild type, $\Delta lmb/adcA/adcAll$, and $\Delta adcCB$ mutant strains was observed after growth in zinc-restricted CDM with 1 μM (top) or 10 μM (bottom) added zinc. Visualization of chain length was performed at a magnification of X1000. Cells were collected during the mid-exponential phase. Images were captured from three separate experiments, and at least 100 chains were counted from each set, for a total of 300 or more chains counted for each strain. Chain length values were distributed between arbitrarily set numerical categories and calculated as a percentage of all counted chains. Results are presented as the mean $\pm$ standard deviation for three independent counts. The asterisks indicate P -values obtained using unpaired Student's t -test between chain-length counts of the WT strain and the $\Delta lmb/adcA/adcAll$ and $\Delta adcCB$ mutant strains (* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$).

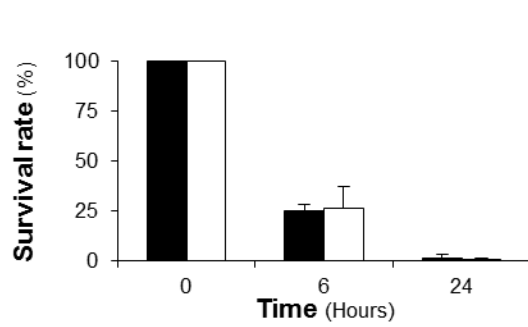


Figure 8. Intracellular survival of *Streptococcus agalactiae* in RAW 264.7 macrophages

RAW 264.7 cells were infected with A909 WT (black bars) or $\Delta lmb/adcA/adcAll$ (white bars) strains (MOI = 10) and phagocytosis was allowed to proceed for 1 h. Antibiotics were then added and the cells incubated for a period of 2 h to kill the extracellular bacteria. This initial antibiotic treatment, which represents time 0 of the experiment, was extended for different times up to 24 h and the cells were lysed to quantify the intracellular survival rates of the bacteria. Results are presented as the mean $\pm$ standard deviation of three independent experiments performed in triplicate for each strain.

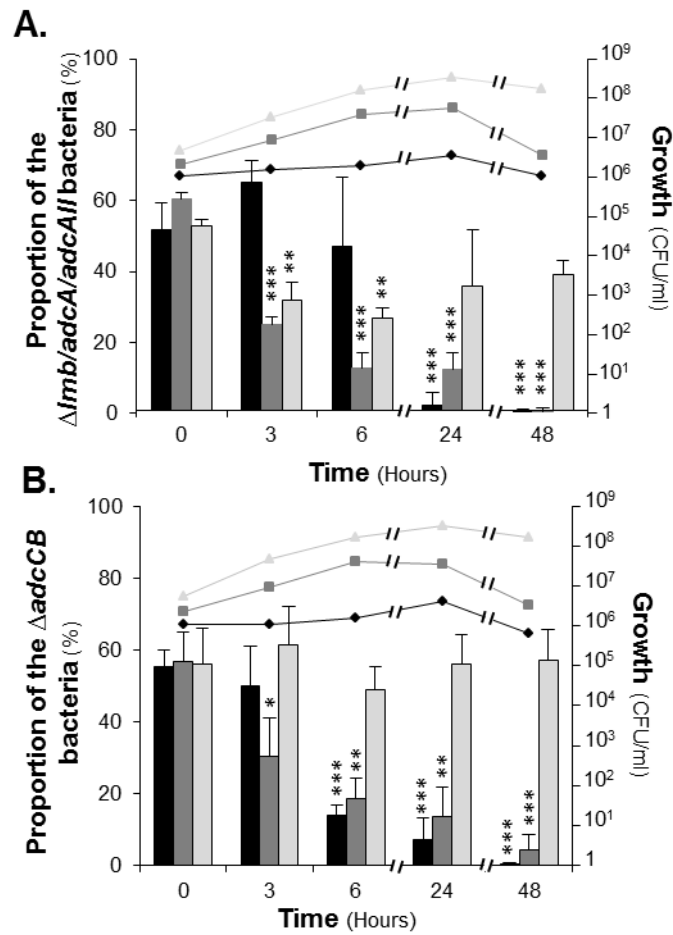


Figure 9. The Adc/Lmb zinc transporter is beneficial for *Streptococcus agalactiae* survival in two human biological fluids

A909 cells **(A)** WT and $\Delta lmb/adcA/adcAll$ mutant strains or **(B)** WT and $\Delta adcCB$ mutant strains, were inoculated in equivalent numbers into human cerebrospinal fluid (black bars and curve with diamond symbols), amniotic fluid (dark grey bars and curve with square symbols), or plasma (light grey bars and curve with triangular symbols). Co-cultures were incubated at 37°C for 48 h without agitation, the CFU counted, and the growth curves traced. The proportion of each strain was monitored by plating diluted cultures on TH agar containing erythromycin (10 μ g/ml) and X-Gal (60 μ g/ml) (see Experimental procedures section). Results are presented as the mean $\pm$ standard deviation for three independent co-cultures. The asterisks indicate *P*-values obtained using unpaired Student's *t*-test between proportion of the strain at T0 and its proportion at the indicated times (* *P*<0.05; ** *P*<0.01; *** *P*<0.001).

SUPPLEMENTAL MATERIAL

Table S1. Bacterial strains and plasmids used in this study

Strain or plasmid	Genotype or description	Source or reference
<u>Escherichia coli strain</u>		
XL1-blue	<i>endA1 gyrA96</i> (Nal ^R) <i>thi-1 recA1 relA1 lac glnV44</i> <i>hsdR17</i> (r _K ⁻ m _K ⁺) <i>F'</i> [::Tn10 (Tet ^R) <i>proAB</i> ⁺ <i>lacI</i> ^q ZΔM15]	Stratagene
<u>Streptococcus agalactiae strains</u>		
A909 (WT)	Isolated from a septic human neonate in 1934	(Tettelin <i>et al.</i> , 2005)
A909Δ <i>lmb</i>	Isogenic <i>lmb</i> (<i>sak_1319</i>) deletion mutant of A909	This study
A909Δ <i>lmb</i> - pTCV-P _{Tet} :: <i>lmb</i>	<i>lmb</i> (<i>sak_1319</i>) plasmid complementation of A909Δ <i>lmb</i> /Δ <i>adcA</i> /Δ <i>adcAll</i>	This study
A909Δ <i>adcA</i>	Isogenic <i>adcA</i> (<i>sak_0685</i>) deletion mutant of A909	This study
A909Δ <i>adcA</i> -pTCV-P _{Tet} :: <i>adcA</i>	<i>adcA</i> (<i>sak_0685</i>) plasmid complementation of A909Δ <i>lmb</i> /Δ <i>adcA</i> /Δ <i>adcAll</i>	This study
A909Δ <i>adcAll</i>	Isogenic <i>adcAll</i> (<i>sak_1898</i>) deletion mutant of A909	This study
A909Δ <i>adcAll</i> -pTCV-P _{Tet} :: <i>adcAll</i>	<i>adcAll</i> (<i>sak_1898</i>) plasmid complementation of A909Δ <i>lmb</i> /Δ <i>adcA</i> /Δ <i>adcAll</i>	This study
A909Δ <i>lmb</i> /Δ <i>adcA</i>	Isogenic <i>lmb</i> and <i>adcA</i> deletion double mutant of A909	This study
A909Δ <i>lmb</i> /Δ <i>adcAll</i>	Isogenic <i>lmb</i> and <i>adcAll</i> deletion double mutant of A909	This study
A909Δ <i>adcA</i> /Δ <i>adcAll</i>	Isogenic <i>adcA</i> and <i>adcAll</i> deletion double mutant of A909	This study
A909Δ <i>lmb</i> /Δ <i>adcA</i> /Δ <i>adcAll</i>	Isogenic <i>lmb</i> , <i>adcA</i> and <i>adcAll</i> deletion triple mutant of A909	This study
A909Δ <i>adcR</i>	Isogenic <i>adcR</i> (<i>sak_0217</i>) deletion mutant of A909	This study
A909Δ <i>adcCB</i>	Isogenic <i>adcCB</i> (<i>sak_0218</i> and <i>sak_0219</i>) deletion double mutant of A909	This study
A909Δ <i>lmb</i> /Δ <i>adcA</i> /Δ <i>adcAll</i> /Δ <i>adcCB</i>	Isogenic <i>lmb</i> , <i>adcA</i> , <i>adcAll</i> and <i>adcCB</i> deletion quintuple mutant of A909	This study
<u>Plasmids</u>		
pG+host1 ^{TS}	Replication-thermosensitive shuttle (^{TS}) plasmid, Ery ^R	(2)
pTCV-P _{Tet}	Mob+(IncP); <i>oriR</i> <i>pACYC184</i> ; <i>oriR</i> <i>pAM_1</i> ; complementation vector, promoter P _{Tet}	(3)
pTCV-P _{Tet} :: <i>lmb</i>	<i>lmb</i> complementation vector, promoter P _{Tet}	This study
pTCV-P _{Tet} :: <i>adcA</i>	<i>adcA</i> complementation vector, promoter P _{Tet}	This study
pTCV-P _{Tet} :: <i>adcAll</i>	<i>adcAll</i> complementation vector, promoter P _{Tet}	This study
pTCV- <i>lac</i>	Promoter probe plasmid carrying a <i>lacZ</i> gene devoided of a promoter and the <i>ermB</i> gene (Ery ^R)	(4)
pTCV- <i>lac</i> -P _{Cyl}	Plasmid carrying a <i>lacZ</i> gene under the control of the constitutive promoter P _{Cyl}	(3)
pTCV- <i>lac</i> :: <i>lmb</i> ^{IR}	pTCV/ <i>lac</i> containing the entire intergenic region (^{IR}) of the <i>S. agalactiae</i> A909 <i>lmb-sht</i> operon in the <i>EcoRI</i> / <i>BamHI</i> site upstream of the <i>lacZ</i> gene	This study
pTCV- <i>lac</i> :: <i>lmb</i> ^{IR-box1*}	pTCV- <i>lac</i> :: <i>lmb</i> ^{IR} with 3 nucleotide substitutions in the AdcR box ₁ .	This study
pTCV- <i>lac</i> :: <i>lmb</i> ^{IR-box2*}	pTCV- <i>lac</i> :: <i>lmb</i> ^{IR} with 4 nucleotide substitutions in the AdcR box ₂ .	This study
pTCV- <i>lac</i> :: <i>lmb</i> ^{IR-box1*box2*}	pTCV- <i>lac</i> :: <i>lmb</i> ^{IR} with combined AdcR box ₁ and box ₂ nucleotide substitutions.	This study

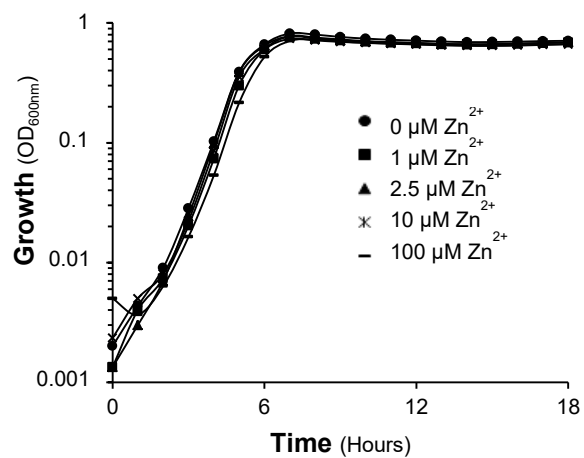


Figure S1. Zinc concentration in the medium does not affect wild type growth

The A909 WT strain was grown in zinc-restricted CDM with various amounts of added Zn^{2+} (from 0 to 100 μM). Growth was monitored by taking $\text{OD}_{600\text{nm}}$ measurements at 1 h intervals during 18 h. Data are representative mean $\text{OD}_{600\text{nm}}$ measurements from three independent experiments.

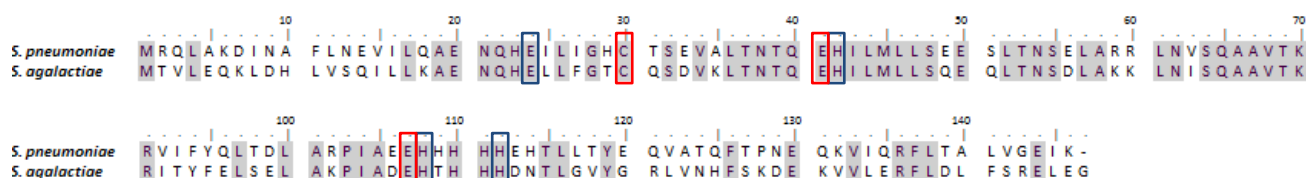


Figure S2. Amino acid sequence alignment of AdcR proteins of *Streptococcus pneumoniae* and *Streptococcus agalactiae*

Identical residues are highlighted in grey. Conserved metal-binding residues of primary sensing site 1 are boxed in blue, residues forming the second site are boxed in red. Sequence alignment was performed using the BioEdit program.

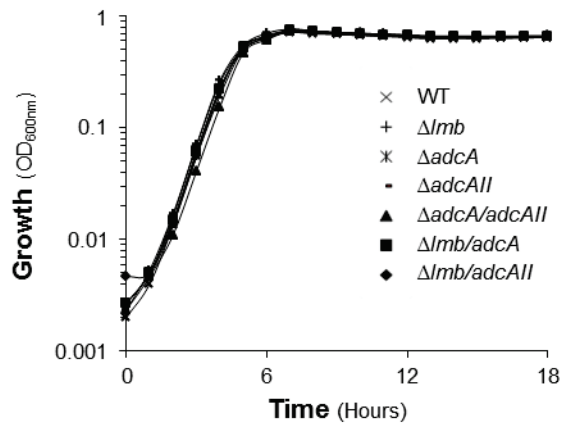


Figure S3. Absence of one or two of the zinc-binding proteins Lmb, AdcA, or AdcAll does not affect bacterial growth in zinc-restricted conditions

A909 WT and mutant strains were grown in zinc-restricted CDM. Growth was monitored by taking OD_{600nm} measurements at 1 h intervals during 18 h. Data are representative mean OD_{600nm} measurements from three independent experiments.

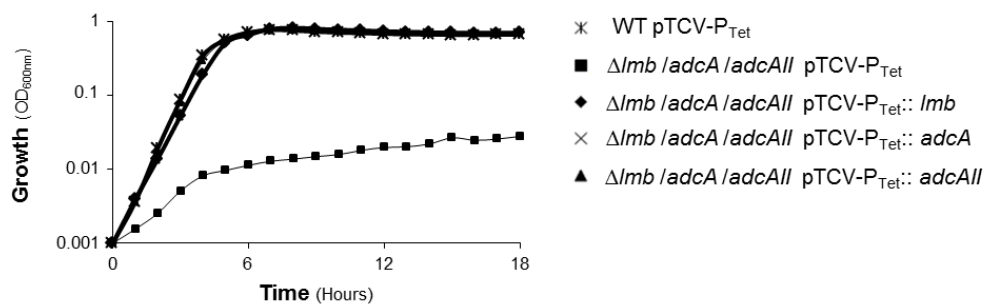


Figure S4. Expression of any of the three binding-proteins fully restored the growth of the $\Delta lmb/adcA/adcAll$ triple mutant.

A909 WT and mutant strains carrying either the empty pTCV-P_{Tet} plasmid or the pTCV-P_{Tet} expressing *lmb*, *adcA* or *adcAll* were grown in zinc-restricted CDM with 10 μ g/ml erythromycin. Growth was monitored by taking OD_{600nm} measurements at 1 h intervals during 18 h. Data are representative mean OD_{600nm} measurements from three independent experiments.

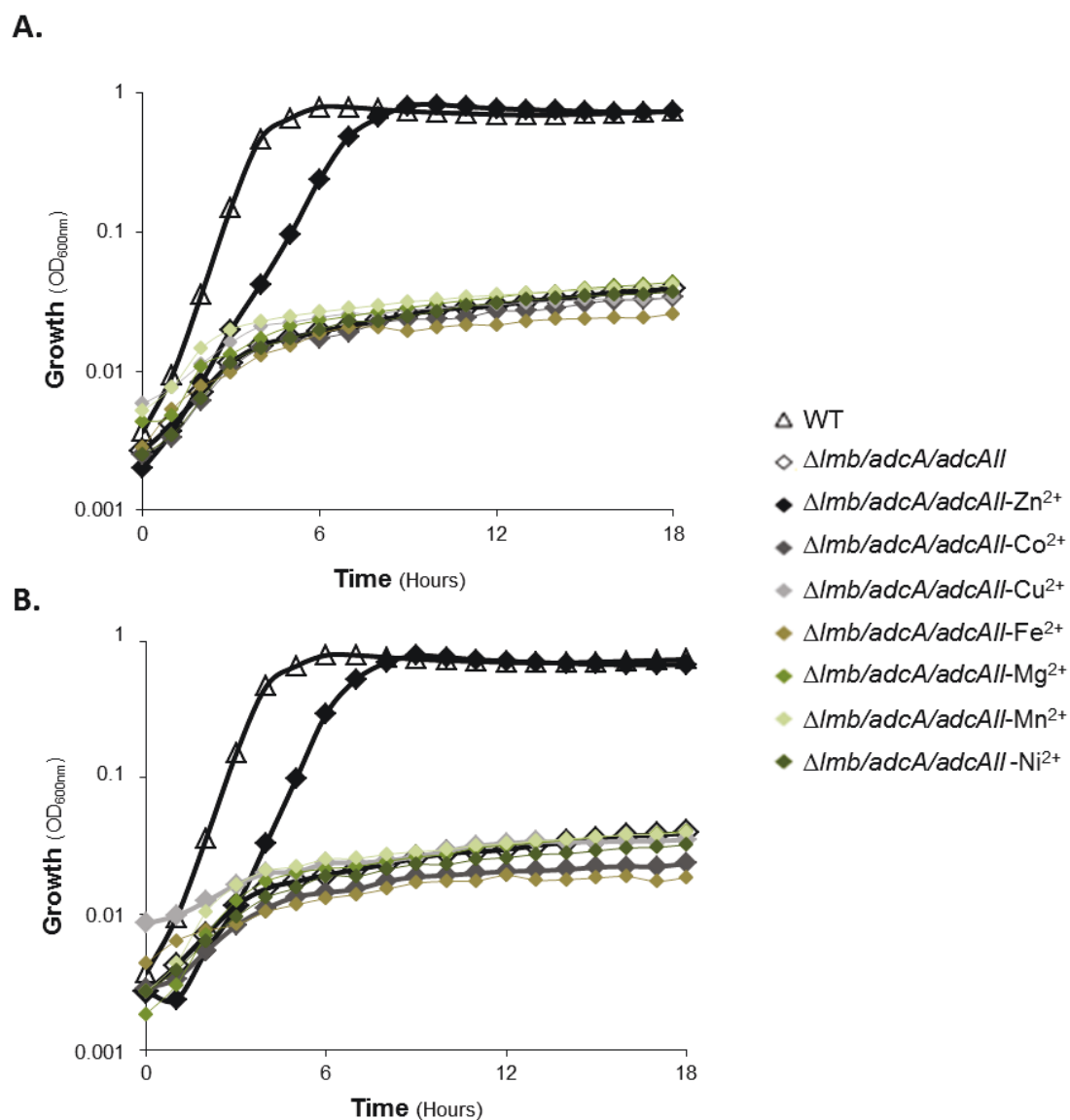


Figure S5. Growth of the *S. agalactiae* $\Delta lmb/adcA/adcAll$ is not restored by addition of metals, except zinc.

A909 cells WT (triangles) or $\Delta lmb/adcA/adcAll$ mutant strain (diamonds) were grown in zinc-restricted CDM with 10 μ M **(A)** or 100 μ M **(B)** of various added metals. The growth was monitored by assessing OD_{600nm} at 1 h intervals during 18 h. Data are representative mean OD_{600nm} measurements from three independent experiments.

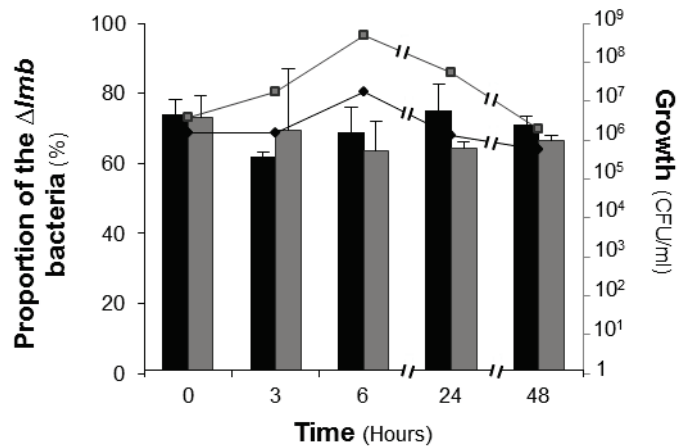


Figure S6. Absence of the Lmb protein does not impair *Streptococcus agalactiae* survival in two human biological fluids

A909 cells WT and Δlmb mutant strains were inoculated in equivalent numbers into human cerebrospinal fluid (black bars and curve with black diamond symbols) or amniotic fluid (dark grey bars and curve with dark grey square symbols). Co-cultures were incubated at 37°C for 48 h without agitation, the CFU counted, and the growth curves traced. The proportion of each strain was monitored by plating diluted cultures on TH agar containing erythromycin (10 μ g/ml) and X-Gal (60 μ g/ml) (see Experimental procedures section). Results are presented as the mean $\pm$ standard deviation for three independent co-cultures.

New insight in the structural features of haloadaptation in α -amylases from halophilic Archaea following homology modeling strategy: folded and stable conformation maintained through low hydrophobicity and highly negative charged surface

Mohamed A. Zorgani, Kevin Patron and Mickaël Desvaux

New insight in the structural features of haloadaptation in α -amylases from halophilic *Archaea* following homology modeling strategy: folded and stable conformation maintained through low hydrophobicity and highly negative charged surface

Mohamed Amine Zorgani · Kevin Patron ·
Mickaël Desvaux

Received: 13 February 2014 / Accepted: 16 May 2014
© Springer International Publishing Switzerland 2014

Abstract Proteins from halophilic archaea, which live in extreme saline conditions, have evolved to remain folded, active and stable at very high ionic strengths. Understanding the mechanism of haloadaptation is the first step toward engineering of halostable biomolecules. Amylases are one of the main enzymes used in industry. Yet, no three-dimensional structure has been experimentally resolved for α -amylases from halophilic archaea. In this study, homology structure modeling of α -amylases from the halophilic archaea *Haloarcula marismortui*, *Haloarcula hispanica*, and *Halalkalicoccus jeotgali* were performed. The resulting models were subjected to energy minimization, evaluation, and structural analysis. Calculations of the amino acid composition, salt bridges and hydrophobic interactions were also performed and compared to a set of non-halophilic counterparts. It clearly appeared that haloarchaeal α -amylases exhibited lower propensities for helix formation and higher propensities for coil-forming regions. Furthermore, they could maintain a folded and stable conformation in high salt concentration through highly negative charged surface with over representation of acidic residues, especially Asp, and low

hydrophobicity with increase of salt bridges and decrease in hydrophobic interactions on the protein surface. This study sheds some light on the stability of α -amylases from halophilic archaea and provides strong basis not only to understand haloadaptation mechanisms of proteins in microorganisms from hypersalines environments but also for biotechnological applications.

Keywords α -Amylase · Archaea · Halophile · Haloadaptation · Stability · Homology modeling

Introduction

Strategies adopted by some microorganisms to colonize extreme environments have attracted a lot of attention [1]. These extremophiles are adapted to thrive and survive in extreme environments characterized by the presence of hostile conditions, such as high or low temperatures, extreme values of pH, high pressure or high salt concentrations. Considering the latter, there is a wide range of halophilic microorganisms belonging to the domains of *Archaea* and *Bacteria* [2, 3]. To be able to survive, actively grow and propagate at high salinity these microorganisms must keep their cytoplasm at least isosmotic with their external environment, in which salt concentration, i.e. NaCl, in the cytoplasm can reach up to 4 M [4].

The stable and unique native structure of a protein is a basic requirement for its proper catalytic activity [5]. Understanding the haloadaptation mechanisms of proteins in microorganisms from hypersalines environments, i.e. salt lakes, is not only important to address fundamental problems related to protein folding, stability and solubility but also of particular interest given their great potential

Electronic supplementary material The online version of this article (doi:10.1007/s10822-014-9754-y) contains supplementary material, which is available to authorized users.

M. A. Zorgani · M. Desvaux
UR454 Microbiologie, INRA, 63122 Saint-Genès Champanelle,
France

Present Address:

M. A. Zorgani (✉) · K. Patron
UMR1282 Bactérie et risques materno-fœtal, Faculté de
médecine, INRA, Université François Rabelais de Tours, 10
Boulevard Tonnellé, 37032 Tours, France
e-mail: mohamedamine.zorgani@etu.univ-tours.fr

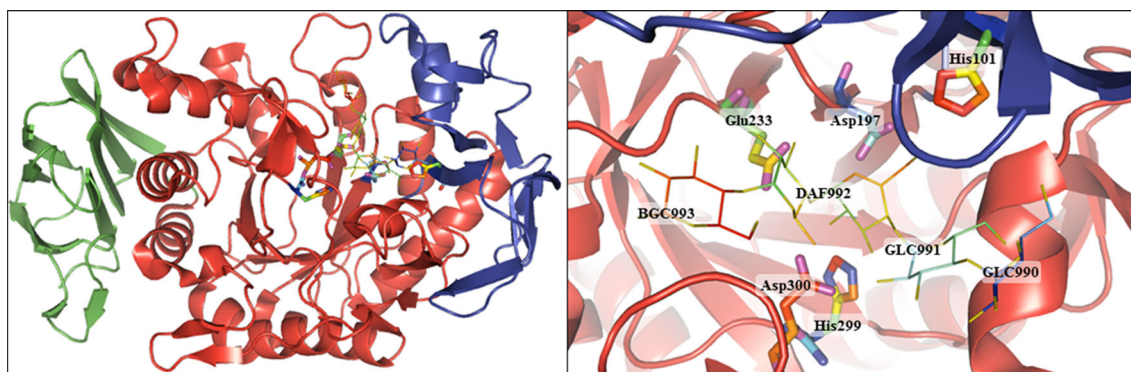


Fig. 1 Structure of a mammalian α -amylase (PDB ID: 1PPI) in a complex with a carbohydrate inhibitor [11]. Left panel shows the domain organization of the α -amylase. Domains A, B and C are shown in red, blue and green respectively. Right panel shows the active center highlighting the amino acids involved in the catalytic site (His101, Asp197, Glu233, His299 and Asp300) in complex with acarbose (a pseudotetrasaccharide α -amylase inhibitor composed of GLC990,

GLC991, DAF992 and BGC993). The carboxylic oxygens of the catalytic residues Glu233 and Asp300 form hydrogen bonds with the glycosidic component of the acarviosine group (DAF992 and BGC993). The third residue of the catalytic triad Asp197 is located on the opposite side of the inhibitor binding cleft. GLC: α -D-Glucose, BGC: β -D-glucose, DAF: 4,6-dideoxy-4-[[[(1S,5R,6S)-3-formyl-5,6-dihydroxy-4-oxocyclohex-2-en-1-yl]amino]- α -D-xylo-hex-5-enopyranose

application in the biotechnology industry. For instance, haloarchaea and their enzymes can be used as biocatalysts in application requiring low water activity (A_w), e.g. biochemical reactions in high salt or organic solvent concentrations [6]. Furthermore, these enzymes from halophilic microorganisms could be used in many harsh industrial processes where the concentrated salt solutions used would otherwise inhibit many enzymatic conversions. For instance, halophilic α -amylases have valuable applications in many industries such as pharmaceutical, textile, sugar, brewing, food, and detergents [7, 8]. Even that, halophilic proteins have not been studied as widely as thermophilic proteins due to the difficulty of crystallizing them at high ionic strengths; most halophilic α -amylases are significantly thermotolerant and remain stable at room temperature over long periods [9].

α -Amylases (α -1,4-glucan-4-glucanohydrolase, EC 3.2.1.1) are endoenzymes that act on internal α -glycosidic bonds of polysaccharides, e.g. starch, glycogen, and hydrolyze these bonds to produce α -anomeric mono- or oligosaccharides [10]. On structural level, α -amylases present a single polypeptide chain folded into three domains named A, B, and C (Fig. 1) [11]. The N-terminal of the polypeptide is constituted of the catalytic domain A, which is made up of eight parallel β -strands arranged in a barrel encircled by eight α -helices called TIM barrel (Triose Phosphate Isomerase barrel). Domain B connects the third β -strand and the third α -helix of the TIM barrel presented as extension of domain A. Domain C constitutes the C-terminal part of the polypeptide and it consists of eight anti-parallel β -strands [10]. The TIM barrel is made up of four regions involved in the active site and highly conserved among α -amylase family [11, 12]: (1) the His positioned in C-terminal

of the third β -strand interacts with the glucose residues of the substrate; (2) the Asp of the fourth β -strand plays the role of nucleophile during catalysis; (3) the fifth β -strand with Glu acting as proton donor/acceptor; and (4) the fourth region at the seventh β -strand with His and Asp residues that may form hydrogen bonds with glucose residues of the substrate.

The most intensively studied α -amylase for which a three-dimensional (3D) structure has been solved are isolated from *Alteromonas haloplantis* (AHA) [13], pig pancreatic (PPA) [14], *Aspergillus oryzae* (TAKA) [15], *Bacillus subtilis* (BSUA) [16], and *Pyrococcus woesei/furiosus* (PWA/pfa) [17]. However, no 3D structure has been resolved experimentally for any of α -amylases isolated from the halophilic archaea, namely *Haloarcula marismortui*, *Haloarcula hispanica*, and *Halalkalicoccus jeotgali*. Therefore, there is not any available information about the structural features allowing these haloarchaeal α -amylases to fold in active and stable conformation in very high salt concentrations. However, it has been noted in several studies that the high frequency of negatively charged residues on protein surface is one of the most prominent features in proteins from halophilic micro-organisms [3, 18, 19]. Relatively low hydrophobicity has been pointed as another explanation for the adaptation to high salt environments [19]. Overall, structural insights gathered from all known halophilic protein structures suggest that they exhibit a notable instability in low salt concentration, and rather than being unfolded in high salt concentration they appear to be dependent on the presence of salts [20]. The current study aimed at characterizing the haloadaptation phenomenon and molecular stability in salty environments of some haloarchaeal α -amylases through 3D structure homology modeling strategy using different bioinformatics approaches.

Methods

Sequence retrieval and similarity search

Three sequences of α -amylase isolated from the halophilic archaea *H. marismortui* (amyHm), *H. hispanica* (amyHh), and *H. jeotgali* (amyHj) were retrieved from UniProtKB (<http://www.uniprot.org/>) under accession numbers: Q5UZY3, Q4A3E0, and D8J7H2, respectively. No 3D structure has been experimentally resolved for any of the three haloarchaeal α -amylases. In order to find the most similar sequence with known 3D structure (template) to each of the target sequences, similarity search were performed using BLASTp [21]. The 3D structures of the templates were extracted from the Protein Data Bank (PDB) (<http://www.rcsb.org/>). The multiple alignments with the targets and the templates were performed using ESPript v3.0 [22].

In silico structure modeling

With a view to study the 3D structures of the selected haloarchaeal α -amylases, comparative modeling was performed using the program MODELLER 9v8 [23]. The encoded 3D structures extracted from the PDB database 1WZA; 1MXG; and 1UOK, were used as templates for the construction of a set of 20 models for the target sequences amyHm; amyHh; and amyHj, respectively and separately. The template 1WZA is an α -amylase isolated from the species *Halothermothrix orenii*, and exhibits halophilic and thermophilic characteristics at the same time with no acidic-surface [24]. The template with the PDB ID 1MXG is an α -amylase produced by the hyperthermophilic *Pyrococcus furiosus*, and has the ability to bind both metals zinc and calcium [17]. The last template with the PDB ID 1UOK isolated from *Bacillus cereus*, is an oligo-1,6-glycosidase (dextrin 6- α -glucanohydrolase, EC 3.2.1.10) very close to α -amylase family [25].

Based on the templates chosen for each haloarchaeal α -amylase, the resulting models generated by MODELLER were conducted into steepest descent energy minimization using Gromacs 4.0.5 with the GROMOS96 45a3 force field [26].

The quality of our models was assessed by two types of energy functions: a semi-empirical force field GROMOS and a knowledge-based energy function Anolea [27]. The GROMOS and Anolea scores were calculated and normalized into *Z score* as described by [28]. We have selected for each target, the best model presenting the lowest average *Z score* for further structural analysis.

Structural analysis and electrostatic potential calculations

Pairwise structural alignments for each of our selected haloarchaeal models with a set of 19 non-halophilic

Table 1 PDB IDs of the non-halophilic α -amylases used in the structural comparison with the haloarchaeal models

α -Amylase origin	PDB ID
<i>Pyrococcus woesei</i>	1MW0
<i>Sulfolobus solfataricus</i>	1EH9
<i>Bacillus subtilis</i>	1UA7
<i>Bacillus amyloliquefaciens</i>	3BH4
<i>Bacillus halmapalus</i>	2JGP
<i>Bacillus sp KR 8104</i>	3DCO
<i>Bacillus stearothermophilus</i>	1HVX
<i>Bacillus licheniformis</i>	1OB0
<i>Bacillus sp KSM-K38</i>	1UD2
<i>Deinococcus radiodurans</i>	2BHU
<i>Lactobacillus plantarum</i>	3DHU
<i>Thermoactinomyces vulgaris</i>	2D0F
<i>Thermotoga maritima</i>	2B5D
<i>Aspergillus niger</i>	2AAA
<i>Aspergillus oryzae</i>	3KWX
<i>Homo sapiens</i>	3OLD
<i>Hordeum vulgare</i>	3BSG
<i>Sus scrofa</i>	3L2L
<i>Tenebrio molitor</i>	1CLV

homologues α -amylases with known 3D structure (Table 1) were performed to identify structural similarities and differences using Dalilite [29]. The secondary structure was predicted using DSSP program implemented in Dalilite [30]. A visual inspection of the constructed models regarding the amino acid composition and localization was also performed with Pymol v1.1 [31]. The surface electrostatic potential of the α -amylases was calculated using the program ABPS (Adaptive Poisson-Boltzmann Solver) implemented in Pymol [32].

Surface amino acids, hydrophobic interactions, and salt bridges calculations

The average number of charged amino acids (Asp, Glu, His, Lys, and Arg) that lies on the surface of the haloarchaeal models, was calculated and compared to the average number of charged amino acids that lies on the surface of the above mentioned set of non-halophilic structures, using a cut-off of 7 as relative solvent accessibility value (*rsav*). An amino acid that has all atoms *rsav* less than or equal to 7 was considered as buried, whereas all other residues with >7 *rsav* were considered to lie on the surface of the structures. This procedure was carried out with PIC program using NACCESS as calculator [33].

The potential number of intraprotein hydrophobic interactions, as well as the number of salt bridges on the constructed models of haloarchaeal α -amylases was calculated

utilizing PIC program and compared to that of non-halophilic homologues. The parameters of calculations for both hydrophobic interactions and salt bridges were set to default [33].

Amino acid composition of halophilic and non-halophilic α -amylases

To find the differences in amino acid composition (Δ AAC) between haloarchaeal and non-halophilic α -amylases, the average value of AAC from the haloarchaeal α -amylases sequences was calculated and compared to the average value of AAC from 5 sets of non-halophilic α -amylases sequences deposited in UniProtKB under EC number 3.2.1.1. This procedure was carried out using MEGA v5.05 [34]. After reducing sequence redundancy using UniRef100 from UniProt, these 5 sets includes 806 α -amylase sequences: 58 sequences from *Archaea*, 525 sequences from *Bacteria* species, 34 sequences from fungi, 139 sequences from metazoan group, and 50 α -amylase sequences belong to plants. The differences in the relative hydrophobic values (Δ RHV) of polar (D, E, H, R, K, Q, N, T, S) and apolar amino acids (V, A, I, L, P, C, M, F, W, Y, G) that composed α -amylase sequences were calculated as described by [35].

Statistical analysis

Differences between the structural properties, AAC and RHV regarding haloarchaeal α -amylases and non-halophilic homologues were calculated whenever was possible. For instance, Δ AAC indicates haloarchaeal minus non-halophilic amino acid composition. The differences significance was evaluated using the *t* test where the null hypothesis is that the average Δ was 0 at 0.05 *p* value. A *p* value <0.05 is considered as significant. A *p* value <0.005 is considered as very significant. A *p* value <0.0005 is considered as very high significant.

Results and discussion

Sequence similarity analysis

Similarity search were performed in order to find the most similar sequence with known structure to generate 3D

structure models for the sequences of the haloarchaeal α -amylases isolated from *H. marismortui* (amyHm), *H. hispanica* (amyHh), and *H. jeotgali* (amyHj). The two sequences amyHm and amyHh shares similar sequence identity with the sequences of 1WZA and 1MXG, respectively, other homologues with experimentally determined structure. Whereas, the haloarchaeal α -amylase secreted by *H. jeotgali* shares high sequence identity and similarity with an oligo-1,6-glycosidase secreted by *Bacillus cereus* (Table 2). In general, the overall sequence similarity of α -amylases is low and can fall below 10 % with some members [17].

The multiple sequence alignment of the three haloarchaeal α -amylase sequences (amyHm, amyHh, and amyHj) with their respective template sequences shows that they present conserved active site residues known among α -amylase family (Fig. 2). The overall conservation of the catalytic residues among the haloarchaeal α -amylases suggests that their enzymatic mechanism is very similar to the double-displacement catalytic mechanism of other known α -amylases [11, 12]. Of note those α -amylases are most certainly secreted via the Sec export pathway as indicated by the presence of N-terminal signal peptide in amyHh for instance [36–38]. In general, aspects related to protein secretion in *Archaea* would require further in-depth investigations. Furthermore, the secondary structure alignment shows that all the 3D structures used as templates (1WZA, 1MXG and 1UOK) to build the haloarchaeal α -amylase models conserve the same global structural organization known among α -amylase family, domains A (TIM barrel), B, and C (Fig. 2). Except for the template 1UOK, where in addition to the TIM barrel, it presents five more α -helix positioned as it follows: α -helix (Nalpha6') between the sixth β -strand (Nbeta6) and the sixth α -helix (Nalpha6); Nalpha7' between Nbeta7 and Nalpha7; and three α -helix (Nalpha8', Nalpha8'', and Nalpha8''') between Nbeta8 and Nalpha8 [25]. In order to get further insight into the sequence similarity with regards to structural resemblance and as non-structural information is available in the literature; the structure of haloarchaeal α -amylases was modeled using homology modeling and the three selected templates were used to generate the haloarchaeal models.

Table 2 Sequence identity and similarity between the sequences of haloarchaeal α -amylase and their respective template sequences. The *rmsd* values (atoms superimposed/atoms aligned) are calculated after superposition with Pymol [31]

Target ID	Template PDB ID	Sequence identity (%)	Sequence similarity (%)	E value ^a	<i>rmsd</i>
amyHm	1WZA	30.0	46.0	2.0e ⁻³⁵	0.359 (241/343)
amyHh	1MXG	27.0	45.0	2.0e ⁻³³	0.273 (226/346)
amyHj	1UOK	46.0	64.0	1.0e ⁻¹⁵⁰	0.173 (442/545)

^a The E value indicates the reliability of the alignment, as match is lower as the alignment is significant

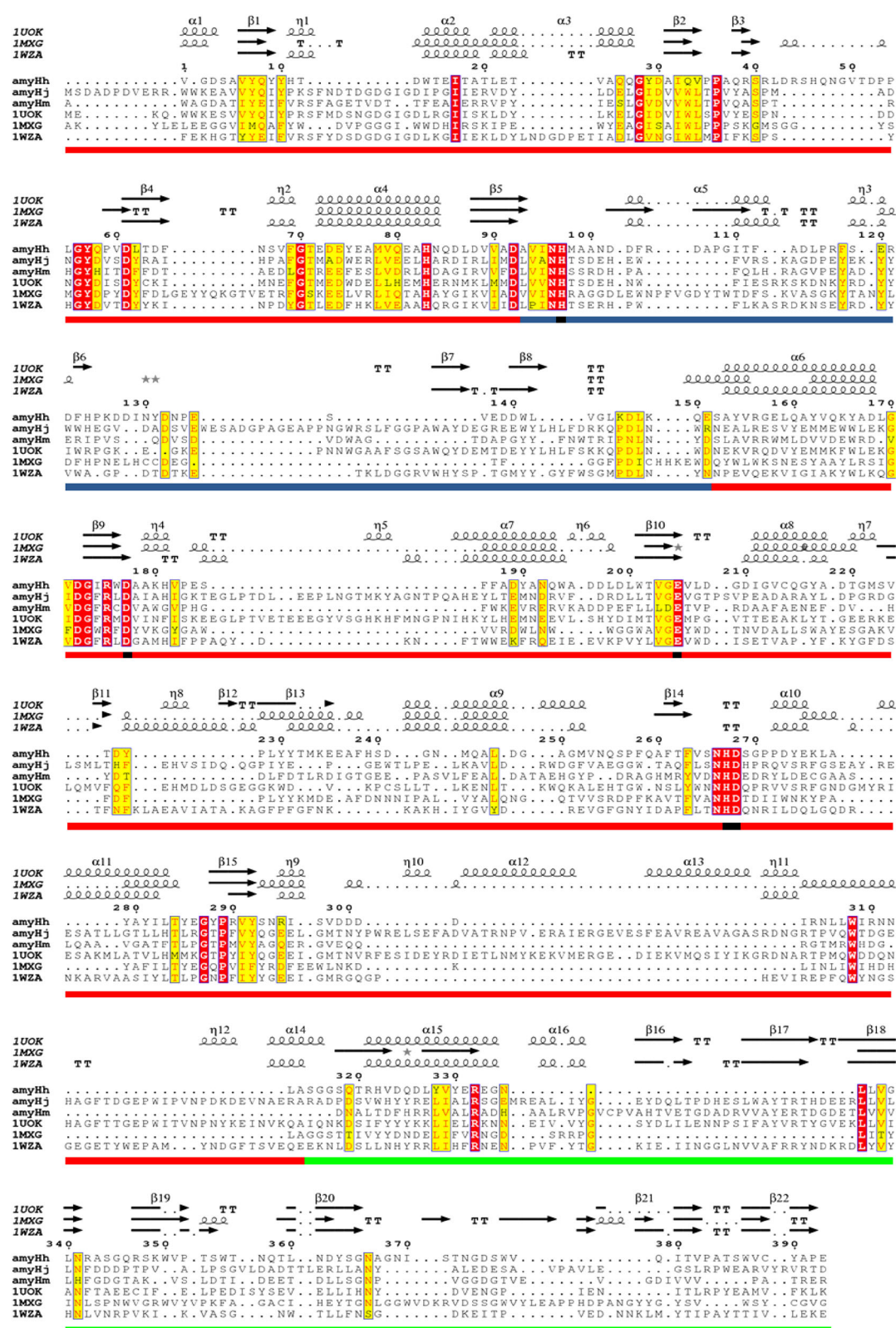
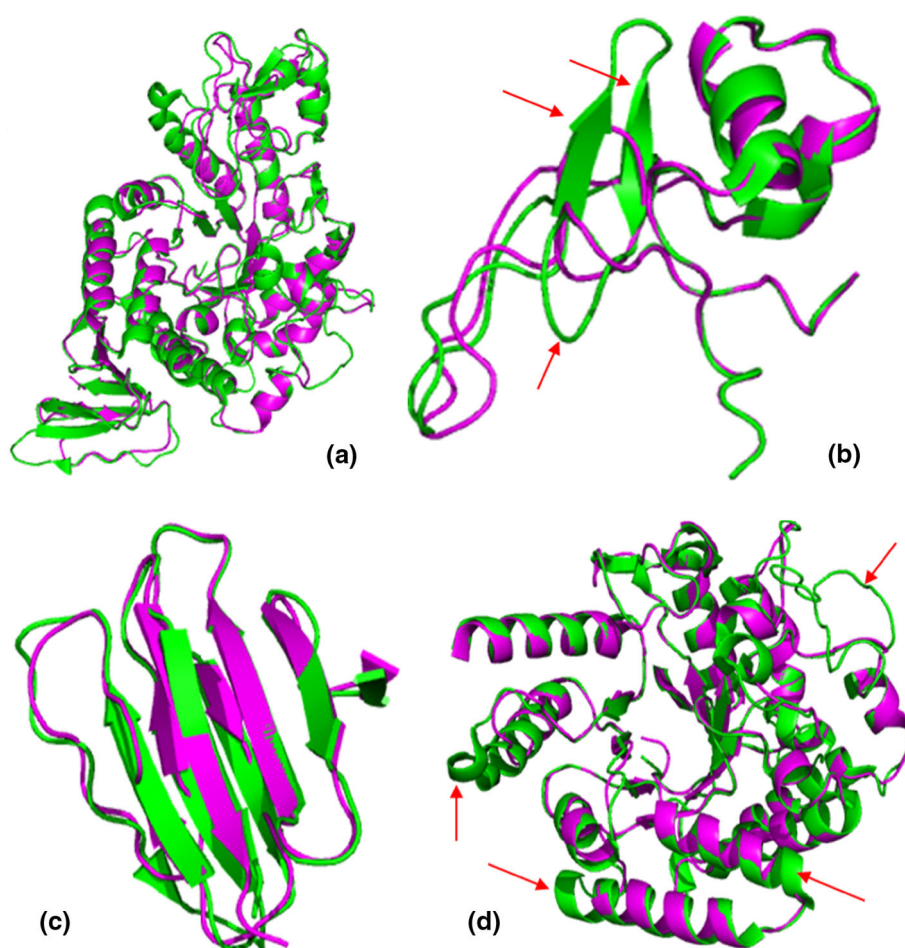


Fig. 2 Sequence similarity analysis of α -amylases. The multiple sequence alignment of the haloarchaeal α -amylases and their respective templates was performed by ESPript v3.0 [22]. On the top of the sequence alignment is represented the secondary structure alignment of the templates. Conserved residues are white colored in

red squares. Similar residues are red colored in yellow squares. Domains A, B and C are delimited by the underscore squares colored in red, blue and green respectively. Residues involved in the active site are indicated with black underscore squares. α : α -helix, β : β -strand, η : 3_{10} -helices, T: turn, TT: strict β -turns

Fig. 3 Structural analysis of the amyHm haloarchaeal model.

a Superposition of the selected haloarchaeal model of α -amylase from *H. marismortui* (in *magenta*, amyHm) and its respective template (in *green*, PDB ID: 1WZA), superimposed and visualized with Pymol [31]. **b–d** Zoom on domain B, C, and A respectively. The *red arrows* are pointed to the structural differences between the target and the template



Building of α -amylase models

The haloarchaeal α -amylase models were generated by homology modeling. The model with the lowest *Z score* was selected for further structural analysis. Model number 19 (amyHh_m19) was selected for amyHh, amyHj_m17 for amyHj, and amyHm_m12 for amyHm (see supplementary material, Tables S1, S2, and S3).

Structural superposition of the three selected models with their respective templates, shows structural resemblance highlighted with very low *rmsd* values (Table 2). The overall structural resemblance (target/template) covers domains A, B, and C. This result should enhance the reliability of our selected models, on structural level, to be used for further structural analysis. The Fig. 3 shows the structural resemblance between the selected α -amylase model for amyHm and its template (PDB ID: 1WZA). Interestingly, structural analysis of domain A, B, and C separately show while there is a marked conservation of the rigid structures (α -helix and β -strand), there is structural differences located especially on the coil-forming regions (Fig. 3b–d). In domain B of the template (Fig. 3b), two β -strands were replaced by coils in the same domain of the

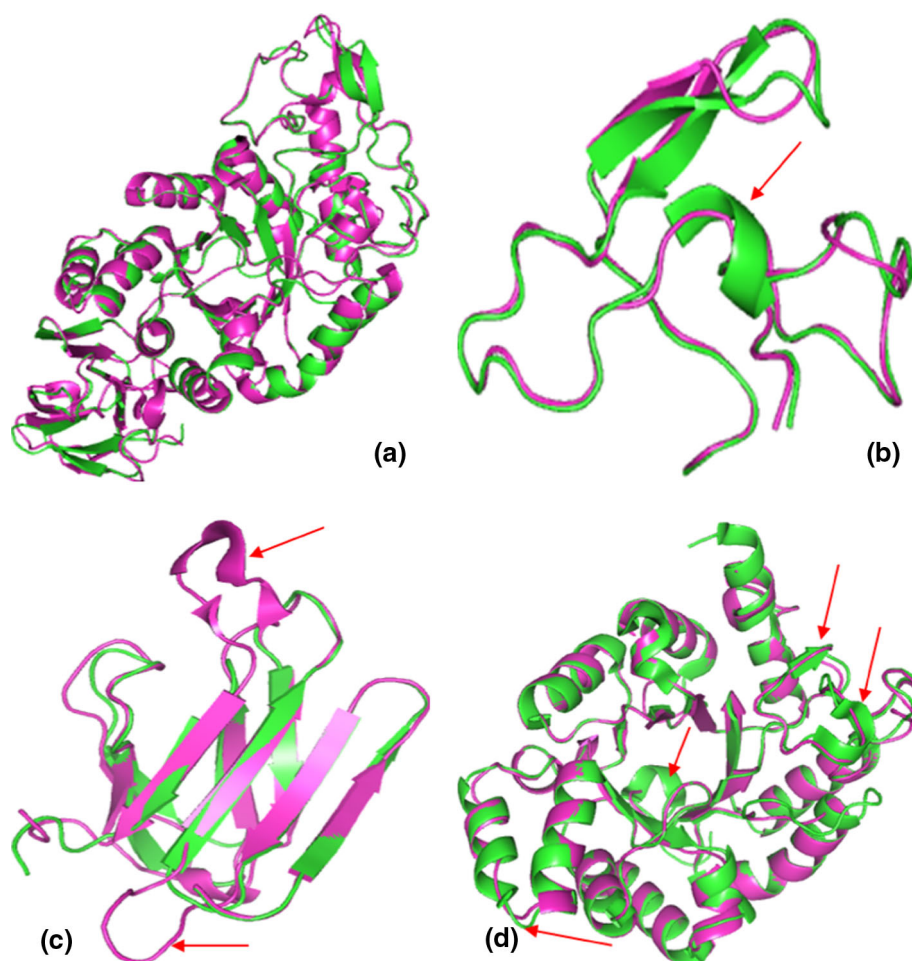
target. Whereas, in domain A (Fig. 3d), parts of the α -helices were also replaced by coil-forming regions. The same structural trends were observed when the haloarchaeal models of amyHh and amyHj were analyzed and compared to their respective templates 1MXG and 1UOK (Figs. 4, 5). This unique structural organization of the haloarchaeal α -amylase models could highlight the fact that these coils regions could be specific characteristics of haloarchaeal α -amylase.

Comparison with known protein structures

In order to deeply analyze the previous structural differences between the haloarchaeal α -amylases models and their non-halophilic counterparts, we have performed structural comparisons taking into account the various forms of protein folding (α -helix, β -strand and coil-forming regions). Compared to non-halophilic structures, the haloarchaeal models present an increase in coil-forming regions (52.7 % for amyHh, 55.1 % for amyHm, and 52.3 % for amyHj) and a decrease in α -helix and β -strands-forming regions (Table 3). Note that, the halophilic enzymes must assume a balance between being

Fig. 4 Structural analysis of the amyHh haloarchaeal model.

a Superposition of the selected haloarchaeal model of α -amylase from *H. hispanica* (in magenta, amyHh) and its respective template (in green, PDB ID: 1MXG), superimposed and visualized with Pymol [31]. **b–d** Zoom on domain B, C, and A respectively. The red arrows are pointed to the structural differences between the target and the template



flexible enough to enable catalytic activities and rigid enough to avoid unfolding. Presumably, the coil regions provide a flexible conformation to the haloarchaeal α -amylases, which could enhance their stability and activity in environments with high ionic strength. Non-halophilic α -amylases, in other hand, present an increase in α -helix and β -strand-forming regions (Table 3). In this case, these rigid conformations could be one of the reasons causing lowering in dynamic motions of the enzyme in high salt concentration, known to be required for efficient catalysis and activity in salty environments [39].

Moreover, the structural inspection of the haloarchaeal models shows that the amino acid Asp is mostly in coil conformation, whereas the amino acid Glu is mostly in α -helix or β -strand conformations (Fig. 6b). From a structural viewpoint, Asp is recognized as an α -helix breaker, whereas Glu is favorable for α -helix formation [40]. In general, the coil regions of proteins are known to prefer Asp over Glu [41]. This could be probably the reason why halophilic proteins use more prevalently Asp than Glu residues.

As it will be discussed further below, residues having lower propensities for α -helical and β -strand regions and

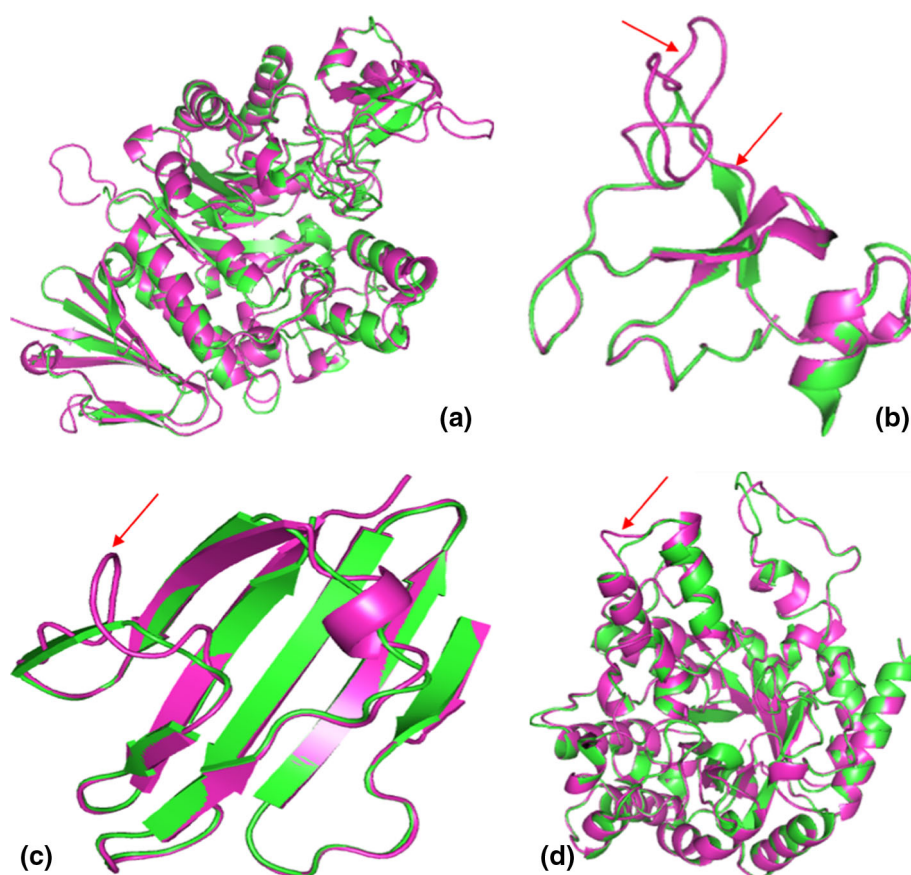
Table 3 Comparison of the average number of secondary structure between the α -amylase models of haloarchaeal species (in bold), the templates (underlined), and the structures of non-halophilic species

α -amylase origin	α -Helix (%)	β -Strand (%)	Coil (%)
<i>Haloarcula hispanica</i>	27.48	19.84	52.67
<i>Haloarcula marismortui</i>	28.14	16.78	55.08
<i>Halalkalicoccus jeotgali</i>	31.11	16.58	52.30
<u><i>Halothermothrix orenii</i></u>	<u>31.35</u>	<u>18.03</u>	<u>50.61</u>
<u><i>Bacillus cereus</i></u>	<u>33.51</u>	<u>17.38</u>	<u>49.10</u>
<u><i>Pyrococcus woesei</i></u>	<u>30.66</u>	<u>21.88</u>	<u>47.46</u>
<i>Bacillus stearothermophilus</i>	26.51	25.25	48.24
<i>Lactobacillus plantarum</i>	35.80	19.66	44.54

higher propensities for forming coil regions, such as Asp and Val, are preferred more in haloarchaeal α -amylases than in non-halophilic counterparts. This particular flexible structural organization among the haloarchaeal α -amylases, could be one of features conferring stability of the 3D structure, to help the enzyme in avoiding the loss of function in extreme salt concentrations by staying folded in the right conformation.

Fig. 5 Structural analysis of the amyHj haloarchaeal model.

a Superposition of the selected haloarchaeal model of α -amylase from *H. jeotgali* (in magenta, amyHj) and its respective template (in green, PDB ID: 1UOK), superimposed and visualized with Pymol [31]. **b–d** Zoom on domain B, C, and A respectively. The red arrows are pointed to the structural differences between the target and the template



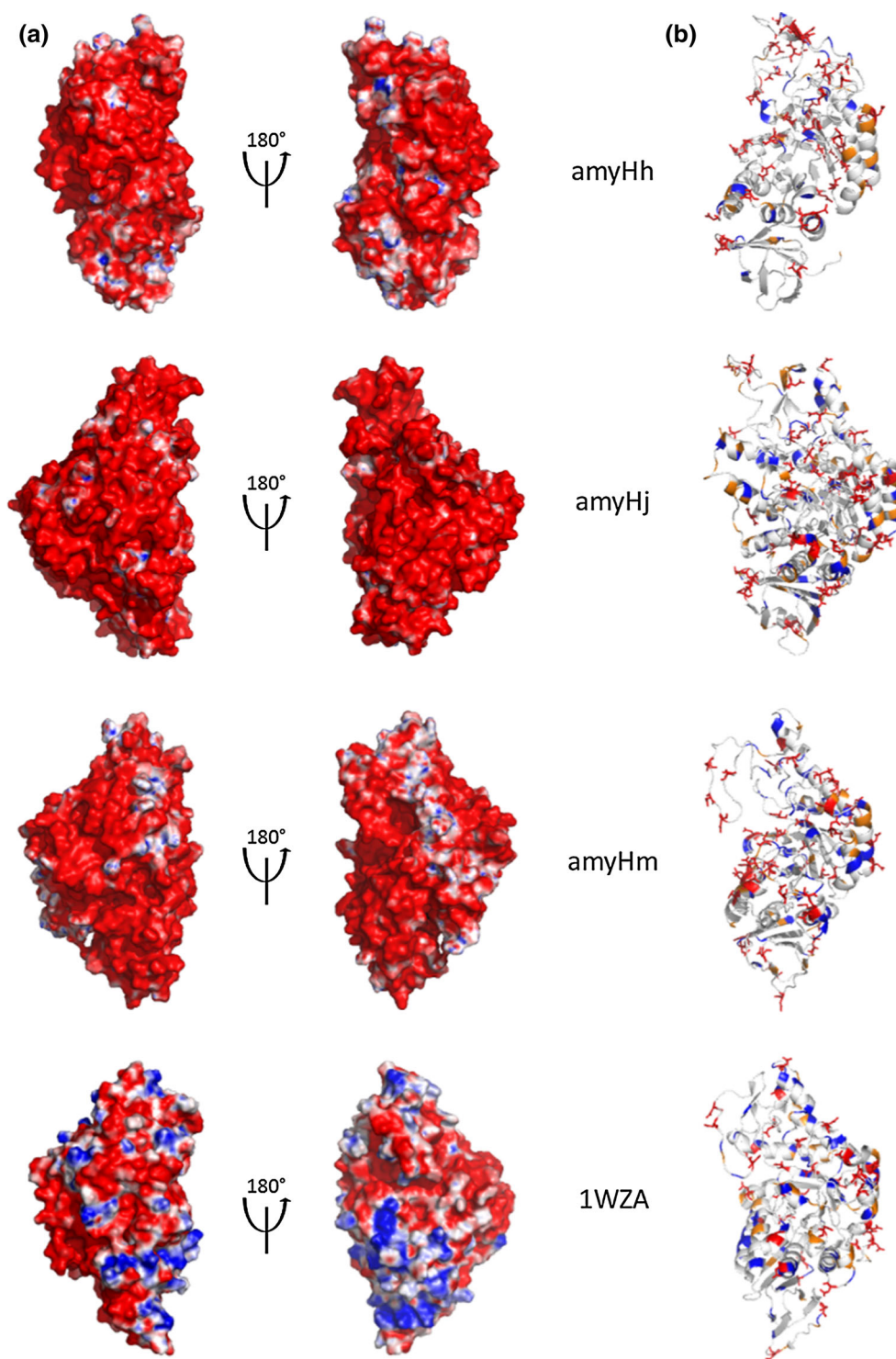
Surface electrostatic potential and intraprotein interactions

The stability of a protein is limited by the nature of the surrounding environment. Halophilic archaea lives in a continuous contact with high salt concentrations. In such harsh conditions, the solvent exposed regions of proteins play an important role for their stability. The calculation of surface electrostatic potential reveals a highly acidic surface for the three haloarchaeal models (Fig. 6a). Structural insights gathered from all known halophilic proteins structures suggest that the conserved acidic surface enhances the stability of the protein through an increased water binding capacity [42], for instance the esterase protein from *H. marismortui* [43]. Unlike all the other haloarchaeal models, the template used for the homology modeling of amyHm (1WZA in Fig. 6a) lacks the conserved acidic surface and both positively and negatively charged residues are distributed equally [24]. It should be mentioned that the α -amylase from the thermo-halophilic *Halothermothrix orenii* with the structure 1WZA exhibit a different haloadaptation mechanism. While 1WZA does not have an acidic surface it does have other structural characteristics, such as, the presence of an equal number of

positively charge residues, binding of two calcium ions and surface exposed salt bridges that form ions network known to stabilize halophilic and thermophilic proteins [24]. Further inspection of the surface-amino acid composition reveals significant differences between the haloarchaeal α -amylase and the 19 non-halophilic counterparts structures (Table 4). The haloarchaeal models present significantly more Asp residues on the surface of the protein than the non-halophilic homologues, and significantly less Lys residues. The higher usage of negatively charged amino acid residues, leads to the organization of a hydrated salt ion network at the surface of the protein and formation of salt bridges with strategically positioned basic residues. Otherwise, by accumulating a large number of negatively charged residues in the surface, mainly Asp residues, the tendency to aggregation will be reduced and the solubility at physiological pH will be increased due to the lowering in the isoelectric point.

Interestingly, the calculation of the number of salt bridges on the haloarchaeal models did not show any significant trends, compared to non-halophilic homologues (Table 5). Nevertheless, the balance between the coil-forming regions and the salt bridges on haloarchaeal α -amylases may confer the necessary rigidity to the enzyme

Fig. 6 Electrostatic properties and surface charge of the of haloarchaeal α -amylase models. **a** The electrostatic surface potential. Electrostatic drawings were produced using the program ABPS (Adaptive Poisson-Boltzmann Solver) implemented in Pymol [32]. Surface colors represent the potential from $-10k_B T^{-1}$ (red) to $+10k_B T^{-1}$ (blue). **b** Structural representation of the 3D structures showing the amino acids proportion and localization. Negatively charged residues Asp and Glu are shown in red sticks and orange, respectively. Positively charged residues His, Arg, and Lys are shown in blue and neutral residues are in white



to avoid unfolding in high ionic strength. At high salt concentrations, the negative free energy of ion binding to the sites formed by the salt bridge networks will counter-balance screening of protein-charged groups by solvent ions [44]. With a stabilizing effect and taking in account the balance between the flexibility provided by the coil-

forming regions and the rigidity provided by the salt bridges, this features constitutes an important structural signature of haloarchaeal α -amylase that could contribute to their stability.

The number of predicted intraprotein hydrophobic interactions shows a significant decrease in the average

Table 4 Average number of charged residues lying on the surface of the haloarchaeal α -amylase models and the non-halophilic homologues

α -Amylase origin	Negatively charged residues		Positively charged residues		
	Δ Asp ^a	Δ Glu	Δ His	Δ Lys	Δ Arg
Halophilic archaea					
Non-halophilic archaea	13.04	−5.55	1.12	−6.10	4.38
Bacteria	13.74	7.83	0.06	−4.61	5.29
Fungi	3.15	13.67	1.81	−5.63	8.81
Metazoan group	8.41	11.84	1.81	−7.19	7.98
Total ^b	38.35	27.79	4.80	−23.52	26.46
Average ^c	9.59	6.95	1.20	−5.88	6.61
<i>t</i> test ^d	0.001	0.110	0.240	0.022	0.003

^a Number of Asp residues lying on the surface of the haloarchaeal α -amylase models minus the number of Asp lying on the surface of the non-halophilic homologues

^b Total Δ residue (Asp, Glu, His, Lys or Arg)

^c Average Δ residue (Asp, Glu, His, Lys or Arg)

^d Boldfaced numbers indicate significant *p* values

number of hydrophobic interactions in haloarchaeal models, highlighted by an increased number of hydrophobic interactions in case of non-halophilic structures (Table 5). Indeed, maintaining stability and native structure through a correct protein folding in hypersaline environment may require less hydrophobic interaction to avoid collapse. As it was mentioned in a previous study, halophilic protein present 2-fold reduction in the proportion of lysine residues in the sequence leading to a 4-fold reduction in the exposed hydrophobic accessible surface compared to non-halophilic homologues [20]. For halophilic proteins, the lowering in the strength of hydrophobic interactions is commonly viewed as the main reason for preventing aggregation of proteins at high salt concentration [19]. Indeed, all structural analyses for the haloarchaeal models, mainly intra-protein interactions, were conducted on structures obtained from homology modeling. However, it is worth mentioning that before calculating these intraprotein interactions, the models were conducted to energy minimization and evaluation which makes the resulting models more reliable for such structural analysis.

Table 5 Intraprotein salt bridges and hydrophobic interactions

Table 5 Intraprotein salt bridges and hydrophobic interactions	α -Amylase PDB ID	Sequence length	Salt bridges (SB)			Hydrophobic interactions (HI)		
			Number of SB	SB (%) ^a	Δ SB ^b	Number of HI	HI (%) ^c	Δ HI ^d
	Haloarchaeal α -amylases							
	1WM0	435	55	6.32	−0.85	487	55.98	−20.32
	1EH9	558	95	8.51	−3.04	546	48.92	−13.26
^a Percentage of salt bridges predicted on the structures of non-halophilic α -amylases	1UA7	425	53	6.24	−0.77	362	42.59	−6.93
	3BH4	484	68	7.02	−1.55	400	41.32	−5.66
^b Percentage of salt bridges predicted on the structures of the haloarchaeal α -amylases minus the percentage of salt bridges predicted on the structures of non-halophilic homologues	2JGP	485	52	5.36	0.11	502	51.75	−16.09
	3DCO	425	73	8.59	−3.12	597	70.24	−34.58
	1HVX	483	58	6.00	−0.53	440	45.55	−9.89
	1OB0	483	57	5.90	−0.43	409	42.34	−6.68
	1UD2	480	57	5.94	−0.47	414	43.13	−7.47
	3DHU	427	45	5.27	0.20	436	51.05	−15.39
^c Percentage of hydrophobic interactions predicted on the structures of non-halophilic α -amylases	2BHU	602	90	7.48	−2.01	532	44.19	−8.53
	2D0F	637	46	3.61	1.86	565	44.35	−8.69
	2B5D	528	95	9.00	−3.53	524	49.62	−13.96
^d Percentage of hydrophobic interactions predicted on the structures of the haloarchaeal α -amylases minus the percentage of hydrophobic interactions predicted on the structures of non-halophilic homologues	2AAA	476	29	3.05	2.42	457	48.00	−12.34
	3KWX	476	37	3.89	1.58	467	49.05	−13.39
	3OLD	496	59	5.95	−0.48	424	42.74	−7.08
	3BSG	404	44	5.45	0.02	395	48.89	−13.23
	3L2L	496	60	6.05	−0.58	423	42.64	−6.98
	1CLV	471	43	4.56	0.91	403	42.78	−7.12
^e Total Δ SB and Δ HI	Total ^e				−10.25			−227.59
^f Average Δ SB and Δ HI	Average ^f				−0.54			−11.98
^g Boldfaced numbers indicate significant <i>p</i> values	<i>t</i> test ^g				0.30			0.004

^a Percentage of salt bridges predicted on the structures of non-halophilic α -amylases

^b Percentage of salt bridges predicted on the structures of the haloarchaeal α -amylases minus the percentage of salt bridges predicted on the structures of non-halophilic homologues

^c Percentage of hydrophobic interactions predicted on the structures of non-halophilic α -amylases

^d Percentage of hydrophobic interactions predicted on the structures of the haloarchaeal α -amylases minus the percentage of hydrophobic interactions predicted on the structures of non-halophilic homologues

^e Total Δ SB and Δ HI

^f Average Δ SB and Δ HI

^g Boldfaced numbers indicate significant *p* values

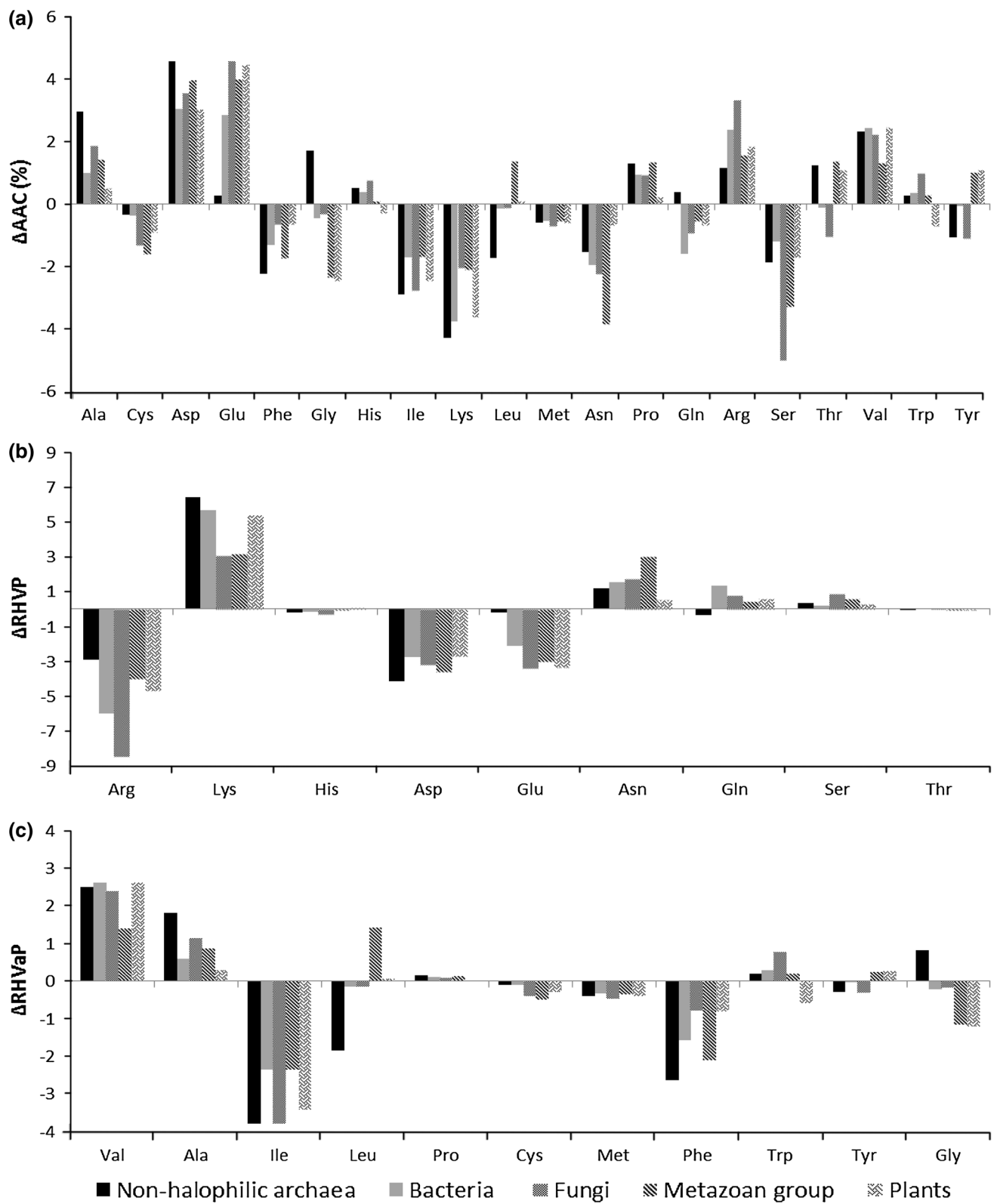


Fig. 7 Amino acid signature of α -amylases. **a** Column diagram reporting the Δ AAC of α -amylases, calculated by MEGA v5.05. Δ AAC report the differences between the amino acid compositions calculated within haloarchaeal α -amylases and non-halophilic homologues (see methods for calculation). Table under the column diagram report the p values of the AAC. *Underlined* digits report significant p value <0.05 . **Boldfaced** digits report very significant p value <0.0005 . *Gray square* digits report very high significant p value <0.0005 . **b, c** Column diagrams reporting the Δ RHV of polar and apolar residues of α -amylases, respectively. Δ RHV report the differences between the relative hydrophobic values of polar or apolar residues calculated within haloarchaeal α -amylases and non-halophilic homologues (see “Methods” for calculation). The relative hydrophobic values were calculated according to the normalized consensus hydrophobic scale of Eisenberg [49]. Labels of the amino acids are reported on the horizontal axes. AAC: amino acid composition. RHV: relative hydrophobic values. RHVP: relative hydrophobic values of polar residues. RHVaP: relative hydrophobic values of apolar residues

Amino acid composition of haloarchaeal α -amylases: A signature of haloadaptation

Genome sequence analyses of halophilic proteins suggest that they have a unique amino acid composition [45, 46]. To investigate the later characteristic, AAC of the haloarchaeal α -amylase was calculated and compared to a set of 806 non-halophilic homologues. The results report a very significant increase in Asp and Glu residues with an average Δ AAC 3.63 % and 3.23 %, respectively (Fig. 7a). Whereas, non-halophilic homologues present an overall decrease in the average number of acidic residues for all phyla included in this study, except from non-halophilic archaeal α -amylases, where it is clearly observed that there is an increased usage of Glu (8.08 %). It should be noted, that Asp residue is the most dominant negatively charged residue in haloarchaeal α -amylases compared to his homologues Glu. As pointed by [47], flexible but ordered proteins are characterized by higher average hydrophilicity and higher occurrence of negatively charged residues, especially Asp.

With highly hydrophobic side chain, Lys residue was significantly less present in haloarchaeal α -amylases then non-halophilic homologues (Fig. 7a). It was previously pointed that low occurrence of Lys combined with an increase in Arg content is another feature of amino acid composition of halophilic proteins [35]. In succeeding in reducing the interaction surface between the protein and the solvent, which is beneficial in an environment where water availability is limited (hydration of ion salts at high concentration) [48], amino acids with short side chains are more prevalent in halophilic proteins when compared to amino acids such as Lys [20].

According to the Eisenberg normalized consensus hydrophobicity scale [49], except the low significant frequency of Lys residues, hydrophilic residues are dominant

in the sequences of haloarchaeal α -amylase in comparison to non-halophilic homologues from other phyla (Fig. 7b). Moreover, there is a significant decreased percentage in large hydrophobic residues such as Ile and Phe with slight increase in small hydrophobic residues such as Val and Ala, in the sequences of haloarchaeal α -amylases compared to non-halophilic counterpart (Fig. 7c). Increased usage of Val in haloarchaeal α -amylases may make them more flexible, because the strong hydrophobic residue Val has a lower helix formation propensity than other strong hydrophobic residues such as Leu, Ile and Met, or positively charged such as Lys [41]. From a structural view point, it has been demonstrated that halophilic proteins present a significant decrease of apolar surface and an increase in polar surface area [20]. These observations are correlated to the increased frequency of Asp and Glu combined to decreased frequency of Lys on the proteins surface.

Conclusion

In order to investigate the haloadaptation mechanism, the present study constitutes the first report focused on structural features of haloarchaeal α -amylases, as well as comparison of the amino acid composition between a large set of homologues protein. Compared to non-halophilic ones, haloarchaeal α -amylases present a different 3D structure organization, to maintain their stability and activity in high salt concentrations. It here appeared that their structure is characterized by (1) an increase in strategically positioned coil regions, (2) an under-representation of α -helix and β -strand-forming regions, (3) highly acidic and negatively charged surface, and (4) the presence of intraprotein interactions such as salt bridges and significantly lower proportion of hydrophobic interactions. In a salty environment, the hydrophobic residues of newly synthesized proteins are exposed to high salt concentrations, leading to non-specific inter- or intramolecular interactions of their side chains, which may compete with proper intramolecular burial within the correct conformation. Those features might contribute to avoid aggregation by enabling flexible but stable conformation changes at high salt concentrations. While performing identical enzymatic functions as their non-halophilic counterparts, the use of haloarchaeal α -amylases could be advantageous as biocatalysts in application requiring low water activity. Eventually, the results of this study could lead to the development of halostable biomolecules.

References

1. Madigan MT, Marrs BL (1997) Extremophiles. *Sci Am* 276(4):82–87

2. Paul S, Bag SK, Das S, Harvill ET, Dutta C (2008) Molecular signature of hypersaline adaptation: insights from genome and proteome composition of halophilic prokaryotes. *Genome Biol* 9(4):R70.
3. Eisenberg H (1995) Life in unusual environments: progress in understanding the structure and function of enzymes from extreme halophilic bacteria. *Arch Biochem Biophys* 318(1):1–5
4. Pieper U, Kapadia G, Mevarech M, Herzberg O (1998) Structural features of halophilicity derived from the crystal structure of dihydrofolate reductase from the Dead Sea halophilic archaeon, *Haloferax volcanii*. *Structure* 6(1):75–88
5. Jaenicke R, Bohm G (1998) The stability of proteins in extreme environments. *Curr Opin Struct Biol* 8(6):738–748
6. Joo WA, Kim CW (2005) Proteomics of halophilic archaea. *J Chromatogr B Anal Technol Biomed Life Sci* 815(1–2):237–250
7. Rasiah IA, Rehm BH (2009) One-step production of immobilized alpha-amylase in recombinant *Escherichia coli*. *Appl Environ Microbiol* 75(7):2012–2016
8. Ghollasi M, Khajeh K, Naderi-Manesh H, Ghasemi A (2010) Engineering of a *Bacillus* alpha-amylase with improved thermostability and calcium independency. *Appl Biochem Biotechnol* 162(2):444–459
9. Hutcheon GW, Vasisht N, Bolhuis A (2005) Characterisation of a highly stable alpha-amylase from the halophilic archaeon *Haloarcula hispanica*. *Extremophiles* 9(6):487–495
10. Prakash O, Jaiswal N (2010) Alpha-amylase: an ideal representative of thermostable enzymes. *Appl Biochem Biotechnol* 160(8):2401–2414
11. Qian M, Haser R, Buisson G, Duee E, Payan F (1994) The active center of a mammalian alpha-amylase. Structure of the complex of a pancreatic alpha-amylase with a carbohydrate inhibitor refined to 2.2-A resolution. *Biochemistry* 33(20):6284–6294
12. van der Maarel M, van der Veen B, Uitdehaag J, Leemhuis H, Dijkhuizen L (2002) Properties and applications of starch-converting enzymes of the alpha-amylase family. *J Biotechnol* 94(2):137–155
13. Aghajari N, Feller G, Gerday C, Haser R (1998) Structures of the psychrophilic *Alteromonas haloplanctis* alpha-amylase give insights into cold adaptation at a molecular level. *Structure* 6(12):1503–1516
14. Gilles C, Astier JP, Marchis-Mouren G, Cambillau C, Payan F (1996) Crystal structure of pig pancreatic alpha-amylase isoenzyme II, in complex with the carbohydrate inhibitor acarbose. *Eur J Biochem* 238(2):561–569
15. Swift HJ, Brady L, Derewenda ZS, Dodson EJ, Dodson GG, Turkenburg JP, Wilkinson AJ (1991) Structure and molecular model refinement of *Aspergillus oryzae* (TAKA) alpha-amylase: an application of the simulated-annealing method. *Acta Crystallogr B* 47(Pt 4):535–544
16. Kagawa M, Fujimoto Z, Momma M, Takase K, Mizuno H (2003) Crystal structure of *Bacillus subtilis* alpha-amylase in complex with acarbose. *J Bacteriol* 185(23):6981–6984
17. Linden A, Mayans O, Meyer-Klaucke W, Antranikian G, Wilmanns M (2003) Differential regulation of a hyperthermophilic alpha-amylase with a novel (Ca, Zn) two-metal center by zinc. *J Biol Chem* 278(11):9875–9884
18. Elcock AH, McCammon JA (1998) Electrostatic contributions to the stability of halophilic proteins. *J Mol Biol* 280(4):731–748
19. Mevarech M, Frolov F, Gloss LM (2000) Halophilic enzymes: proteins with a grain of salt. *Biophys Chem* 86(2–3):155–164
20. Siglioccolo A, Paiardini A, Piscitelli M, Pascarella S (2011) Structural adaptation of extreme halophilic proteins through decrease of conserved hydrophobic contact surface. *BMC Struct Biol* 11:50
21. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410
22. Gouet P, Robert X, Courcelle E (2003) ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res* 31(13):3320–3323
23. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234(3):779–815
24. Sivakumar N, Li N, Tang JW, Patel BK, Swaminathan K (2006) Crystal structure of AmyA lacks acidic surface and provide insights into protein stability at poly-extreme condition. *FEBS Lett* 580(11):2646–2652
25. Watanabe K, Hata Y, Kizaki H, Katsube Y, Suzuki Y (1997) The refined crystal structure of *Bacillus cereus* oligo-1,6-glucosidase at 2.0 Å resolution: structural characterization of proline-substitution sites for protein thermostabilization. *J Mol Biol* 269(1):142–153
26. Bekker H, Berendsen HJC, Dijkstra EJ, Achterop S, Van Drunen R, Van Der Spoel D, Sijbers A, Keegstra H, Reitsma B, Renardus MKR (1993) Gromacs: a parallel computer for molecular dynamics simulations. *Phys Comput* 92:252–256
27. Melo F, Devos D, Depiereux E, Feytmans E (1997) ANOLEA: a www server to assess protein structures. *Proc Int Conf Intell Syst Mol Biol* 5:187–190
28. Khemili S, Kwasigroch JM, Hamadouche T, Gilis D (2012) Modelling and bioinformatics analysis of the dimeric structure of house dust mite allergens from families 5 and 21: Der f 5 could dimerize as Der p 5. *J Biomol Struct Dyn* 29(4):663–675
29. Dietmann S, Holm L (2001) Identification of homology in protein structure classification. *Nat Struct Biol* 8(11):953–957
30. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22(12):2577–2637
31. Delano WL (2002) The pymol molecular graphics system. DeLano Scientific
32. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98(18):10037–10041
33. Tina KG, Bhadra R, Srinivasan N (2007) PIC: protein interactions calculator. *Nucleic Acids Res* 35(Web Server issue):W473–W476
34. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24(8):1596–1599
35. Kastiris PL, Papandreou NC, Hamodrakas SJ (2007) Haloadaptation: insights from comparative modeling studies of halophilic archaeal DHFRs. *Int J Biol Macromol* 41(4):447–453
36. Chagnot C, Zorgani MA, Astruc T, Desvaux M (2013) Proteinaceous determinants of surface colonization in bacteria: bacterial adhesion and biofilm formation from a protein secretion perspective. *Front Microbiol* 4:303
37. Desvaux M, Hebraud M, Talon R, Henderson IR (2009) Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue. *Trends Microbiol* 17(4):139–145
38. Desvaux M, Parham NJ, Scott-Tucker A, Henderson IR (2004) The general secretory pathway: a general misnomer? *Trends Microbiol* 12(7):306–309
39. Sawaya MR, Kraut J (1997) Loop and subdomain movements in the mechanism of *Escherichia coli* dihydrofolate reductase: crystallographic evidence. *Biochemistry* 36(3):586–603
40. Karlin S, Brocchieri L, Bergman A, Mrazek J, Gentles AJ (2002) Amino acid runs in eukaryotic proteomes and disease associations. *Proc Natl Acad Sci USA* 99(1):333–338
41. Costantini S, Colonna G, Facchiano AM (2006) Amino acid propensities for secondary structures are influenced by the protein structural class. *Biochem Biophys Res Commun* 342(2):441–451
42. Bieger B, Essen LO, Oesterheld D (2003) Crystal structure of halophilic dodecin: a novel, dodecameric flavin binding protein from *Halobacterium salinarum*. *Structure* 11(4):375–385
43. Muller-Santos M, de Souza EM, Pedrosa Fde O, Mitchell DA, Longhi S, Carriere F, Canaan S, Krieger N (2009) First evidence

- for the salt-dependent folding and activity of an esterase from the halophilic archaea *Haloarcula marismortui*. *Biochim Biophys Acta* 1791(8):719–729
44. Ebrahimie E, Ebrahimi M, Sarvestani NR (2011) Protein attributes contribute to halo-stability, bioinformatics approach. *Saline Syst* 7(1):1
 45. Dym O, Mevarech M, Sussman JL (1995) Structural features that stabilize halophilic malate dehydrogenase from an archaeobacterium. *Science* 267(5202):1344–1346
 46. Fukuchi S, Yoshimune K, Wakayama M, Moriguchi M, Nishikawa K (2003) Unique amino acid composition of proteins in halophilic bacteria. *J Mol Biol* 327(2):347–357
 47. Radivojac P, Obradovic Z, Smith DK, Zhu G, Vucetic S, Brown CJ, Lawson JD, Dunker AK (2004) Protein flexibility and intrinsic disorder. *Protein Sci* 13(1):71–80
 48. Tadeo X, Lopez-Mendez B, Trigueros T, Lain A, Castano D, Millet O (2009) Structural basis for the aminoacid composition of proteins from halophilic archaea. *PLoS Biol* 7(12):e1000257. doi:[10.1371/journal.pbio.1000257](https://doi.org/10.1371/journal.pbio.1000257)
 49. Eisenberg D, Schwarz E, Komaromy M, Wall R (1984) Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J Mol Biol* 179(1):125–142

Proteinaceous determinants of surface colonization in bacteria: bacterial adhesion and biofilm formation from a protein secretion perspective

Caroline Chagnot, Mohamed A. Zorgani, Thierry Astruc and Mickaël Desvaux



Proteinaceous determinants of surface colonization in bacteria: bacterial adhesion and biofilm formation from a protein secretion perspective

Caroline Chagnot^{1,2†}, Mohamed A. Zorgani^{1†}, Thierry Astruc² and Mickaël Desvaux^{1*}

¹ UR454 Microbiologie, INRA, Saint-Genès Champanelle, France

² UR370 Qualité des Produits Animaux, INRA, Saint-Genès Champanelle, France

Edited by:

Biswarup Mukhopadhyay, Virginia
Bioinformatics Institute, USA

Reviewed by:

Dirk Linke, Max Planck Society,
Germany
Karl G. Wooldridge, University of
Nottingham, UK

*Correspondence:

Mickaël Desvaux, UR454
Microbiology, INRA
Clermont-Ferrand Research Centre,
Site of Theix, F-63122 Saint-Genès
Champanelle, France
e-mail: mickael.desvaux@
clermont.inra.fr

[†] These authors have contributed
equally to this work and share first
authorship.

Bacterial colonization of biotic or abiotic surfaces results from two quite distinct physiological processes, namely bacterial adhesion and biofilm formation. Broadly speaking, a biofilm is defined as the sessile development of microbial cells. Biofilm formation arises following bacterial adhesion but not all single bacterial cells adhering reversibly or irreversibly engage inexorably into a sessile mode of growth. Among molecular determinants promoting bacterial colonization, surface proteins are the most functionally diverse active components. To be present on the bacterial cell surface, though, a protein must be secreted in the first place. Considering the close association of secreted proteins with their cognate secretion systems, the secretome (which refers both to the secretion systems and their protein substrates) is a key concept to apprehend the protein secretion and related physiological functions. The protein secretion systems are here considered in light of the differences in the cell-envelope architecture between diderm-LPS (archetypal Gram-negative), monoderm (archetypal Gram-positive) and diderm-mycolate (archetypal acid-fast) bacteria. Besides, their cognate secreted proteins engaged in the bacterial colonization process are regarded from single protein to supramolecular protein structure as well as the non-classical protein secretion. This state-of-the-art on the complement of the secretome (the secretion systems and their cognate effectors) involved in the surface colonization process in diderm-LPS and monoderm bacteria paves the way for future research directions in the field.

Keywords: secretome, adhesin, pili/fimbriae/curli, cell surface, aggregation, secreted protein, MSCRAMM, protein secretion system

INTRODUCTION

Relative to the bacterial cell, three major classes of interactions can be distinguished: (i) the symbiotic relationships with others biological entities from eukaryotic cells, bacterial cells to viruses (bacteriophages), e.g., mutualism, amensalism, competition, etc..., (ii) the sensing of products or stimuli (cell-cell communication, mechano-physico-chemico sensitivity), and (iii) the direct contacts with surfaces or interfaces. Bacterial colonization corresponds to the presence of microorganisms in a particular environment. Colonization of biotic or abiotic surfaces results from two quite distinct microbiological processes, namely bacterial adhesion or biofilm formation. Initial bacterial adhesion can either be reversible or irreversible. Broadly speaking, a biofilm is defined as the sessile development of microbial cells. The engagement into a sessile mode of growth arises following irreversible bacterial adhesion (Zhao et al., 2013); though, not all single adhered bacterial cells necessarily engage into sessile development. Considering the differences in the molecular physiology of the bacterial cells, the distinction between single adhered and biofilm cells is of importance (Buncic et al., 2013). While initial attachment depends on the physiological state of the cell

prior to adhesion, significant changes in gene/protein expression occur upon irreversible adhesion and further sessile cell division to colonize the site of adhesion. In all cases, the bacterial cell envelope plays a critical role in the primary interactions between the bacterial cell and its environment.

Numerous factors promote bacterial cell adhesion to surface and interfaces (An and Friedman, 1998; Shirtliff et al., 2002). They can be categorized into (i) general physico-chemical surface properties, (ii) the exopolymeric matrix, and (iii) the cell surface biochemical components. The physico-chemistry of the interactions (such as van der Waals attraction, gravitational force, electrostatic charge, or hydrophobic interaction) (Gottenbos et al., 2000) have been theorized into models by the thermodynamic (Morra and Cassinelli, 1996) and the DLVO (Derjaguin-Landau-Verwey-Overbeek) theories, as well as its extended version (Jucker et al., 1998; Hermansson, 1999). While those models can help explaining some experimental observations, neither of them can fully describe bacterial adhesion as they fail in taking into account a fundamental properties of a biological system that is its adaptability and variability (Katsikogianni and Missirlis, 2004). Indeed, in the course of adhesion and/or biofilm formation, bacterial

cells change their physiology at different regulatory levels, e.g., gene/protein expression. This can induce modifications of cell morphology, general surface properties and/or express specific determinants for adhesion. Among those, the exopolymers can be prominent components synthesized in the course of bacterial adhesion and biofilm formation where they play the role of molecular glue (Flemming and Wingender, 2010). A large variety of exopolymers can be involved in bacterial colonization process namely various exopolysaccharides (EPS), e.g., alginate, cellulose, or poly-*N*-acetylglucosamine (PNAG) (Ryder et al., 2007; Vu et al., 2009; Bazaka et al., 2011), extracellular DNA (eDNA) and polyglutamate (Candela and Fouet, 2006). In some cases, those exopolymers are closely associated with the bacterial cell envelope, such as some polyglutamate covalently linked to peptidoglycan (Candela et al., 2012). Some other components of the cell envelope can participate in biofilm formation such as lipopolysaccharides (LPS) (Nakao et al., 2012), (lipo)teichoic acids (Gross et al., 2001; Fabretti et al., 2006), or mycolic acids (Ojha et al., 2005; Zambrano and Kolter, 2005). Among cell envelope components, proteins are undoubtedly the most functionally diverse active components. To be present on the bacterial cell surface, though, a protein must be secreted in the first place (or released by non-active translocation process, e.g., cell lysis or membrane budding).

Protein secretion is a key event for the presence of effectors at the interface between the bacterial cell and its immediate environment (Henderson et al., 2004). Those effectors can be displayed on the bacterial cell surface following anchoring to the cell envelope, released into the extracellular milieu or even beyond, i.e., within a host cell. Secreted proteins feature the lifestyle of a bacterium, its interaction within the ecosystems, microbiota, and ecological niches; for instance, virulence factors in pathogenic bacteria or degradative enzymes in saprophytic bacteria. As such, protein secretion is a key player in bacterial cell physiology and interactions with their environment. To understand protein secretion systems in bacteria, it is crucial to consider the cell envelope architecture. In recent years, it clearly appeared the grouping of bacteria into Gram-positive and Gram-negative bacteria was not satisfactory but ambiguous to describe and categorize accurately the protein secretion systems. Considering the inherent ambiguities of the Gram-negative and Gram-positive terminology, which can refer to three different and sometimes completely unrelated aspects (i.e., Gram staining, cell envelope architecture, and taxonomic grouping) (Desvaux et al., 2009b), the description of monoderm and diderm bacteria is much more appropriate in the field of protein secretion, at least. Indeed, monoderm bacteria (monodermata) refers specifically to species exhibiting only one biological membrane, i.e., the cytoplasmic membrane (CM), whereas diderm bacteria (didermata) refers specifically to species exhibiting two biological membranes, i.e., a CM, then also called inner membrane (IM) and an outer membrane (OM). Diderm bacteria can be further discriminated into (i) diderm-LPS bacterial cells, which possess an archetypal and asymmetrical OM containing lipopolysaccharide (LPS) on the external side, (ii) simple-diderm bacterial cells, which possess an OM lacking LPS, and (iii) diderm-mycolate bacterial cells, where an outer lipid layer resembling an OM is composed of mycolic acid molecules

arranged in a highly ordered form (Brennan and Nikaido, 1995; Sutcliffe, 2010; Gupta, 2011).

Considering the close association of secreted proteins with their cognate secretion systems, the secretome is a key concept to apprehend the protein secretion and related physiological functions. The secretome refers both to the secretion/translocation systems and the protein substrates of these transport systems (Tjalsma et al., 2000; Antelmann et al., 2001, 2006; Van Dijk et al., 2001; Economou, 2002; Sarvas et al., 2004; Buist et al., 2006; Desvaux et al., 2009b). The secretome concept provides an integrated and global view of protein secretion by considering protein routing, transport mechanisms, post-translational modifications, and protein subcellular location. Confusion between the secretome and the secreted/extracellular proteins is a common misunderstanding promulgated in part of the scientific literature but it must be stressed again the secretome is neither the exoproteome (extracellular proteome) nor an “omics” approach *per se* (Desvaux et al., 2009b). The exoproteome refers specifically to the subset of proteins present in the supernatant (which are not necessarily secreted). Secreted proteins are not necessarily free soluble extracellular proteins (exoproteins) since they can have different final subcellular locations (membranes, cell wall, extracellular milieu) or be subunits of supramolecular protein complexes (e.g., flagellum, pilus, cellulosome). Secreted proteins can even have multiple final subcellular locations, which are described following the gene ontology (GO) for “Cellular component” (Figure 1). All extracytoplasmic proteins are not systematically secreted since some exoproteins can be released upon molecular events that are not active translocation process and thus not secretion *per se*, e.g., cell lysis (autolysis, allolysis, bacteriophage lysis), GTA (gene transfer agent) or membrane budding (vesicles). The translocation corresponds to the active transport across a biological membrane (Desvaux et al., 2009b); the secretion refers to active transport from the interior to the exterior of the cell and the export to active transport across the CM (Economou et al., 2006). While in monoderm bacteria secretion and export are synonymous, in diderm bacteria the secretion is completed only upon translocation across the OM. For the sake of clarity, these key definitions in the field of protein secretion are reminded in Table 1.

With these different concepts in hand, it becomes clear a comprehensive understanding of protein determinants involved in bacterial adhesion and/or biofilm formation necessitates a consideration of the cell envelope architecture, i.e., respective to the diderm-LPS, monoderm and diderm-mycolate bacteria trichotomy, as well as their respective protein secretion systems.

THE PROTEIN SECRETION SYSTEMS IN DIDERM-LPS, MONODERM AND DIDERM-MYCOLATE BACTERIA

In diderm-LPS bacteria, nine protein secretion systems have been unravelled so far. For better or worse, these systems have been numerically classified from the Type I (T1SS) to Type IX secretion systems (T9SS) (Desvaux et al., 2009b; McBride and Zhu, 2013) (Figure 2). The T1SS refers to a three-component complex, i.e., a pore-forming OM protein (OMP) of the TolC family, a membrane fusion protein (MFP), and an IM ATP-binding cassette (ABC) exporter (Delepelaire, 2004; Holland et al., 2005;

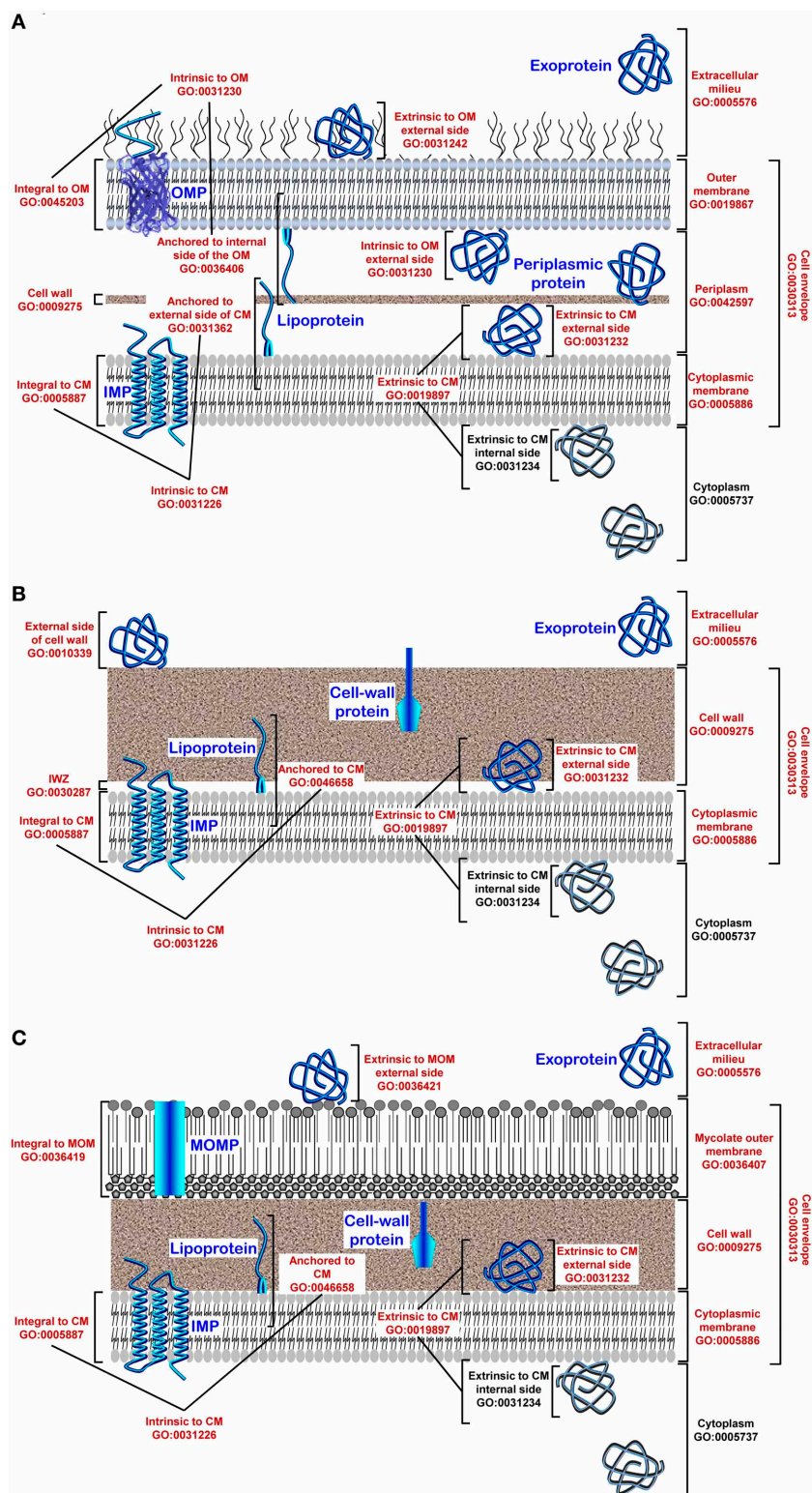


FIGURE 1 | Gene Ontology (GO) for cellular components and protein subcellular location in bacteria. (A) In diderm-LPS bacteria, five clearly defined compartments are considered (i) the cytoplasm (CP; GO:0005737), (ii) the cytoplasmic membrane (CM; GO:0005886), (iii) the periplasm (PP; GO:0042597), (iv) the outer membrane (OM; GO:0019867), and (v) the extracellular milieu (EM; GO:0005576). The cell envelope (GO:0010339) is

constituted of the OM and CM, also called inner membrane (IM), as well as a thin cell wall (CW) located in-between. The situation at the CM can be discriminated even further between locations either intrinsic (GO:0031226) or extrinsic (GO:0019897) to the CM. The former refers to gene products with covalently attached moieties embedded in the CM, which splits into

(Continued)

FIGURE 1 | Continued

locations (i) integral to CM (GO:0005887) where some part of the peptide sequence spans all or part of the CM, i.e., the integral membrane proteins (IMPs), and (ii) anchored to the external side of CM (GO: 0031362) corresponding to proteins tethered to the CM by non-polypeptidic covalently attached anchor, i.e., the lipoproteins. The latter refers to proteins extrinsic to the CM, i.e., neither anchored by covalent bonds to any moiety nor directly embedded in the CM. Peripheral proteins are loosely bound to the surface components of the CM on the internal (GO:0031234) or external side (GO:0031232). Some proteins localized at the OM can also be subunits of supramolecular protein complexes (GO:0043234). The situation at the OM can be discriminated even further between locations intrinsic (GO:003230) or extrinsic (GO:0031242) to the OM. The former refers to gene products with a covalently attached moiety embedded in the OM, which splits into locations (i) integral to OM (GO:00045203) where some part of the peptide sequence spans all or part of the OM, i.e., the outer membrane proteins (OMPs), and (ii) anchored to internal side of OM (GO:0036406) corresponding to proteins tethered to the OM by non-polypeptidic covalently attached anchor, i.e., some lipoproteins. **(B)** In monoderm bacteria, four clearly defined compartments are considered (i) the cytoplasm (CP; GO:0005737), (ii) the cytoplasmic membrane (CM; GO:0005886), (iii) the cell wall (CW; GO:0009275), and (iv) the extracellular milieu (EM; GO:0005576). An inner wall zone (IWZ) (Matias and Beveridge, 2005) has been identified (GO:0030287); importantly, it should not be considered *sensu stricto* as a periplasm since the CW is porous and therefore it is not bordered (bounded) contrary to the situation in diderm-LPS bacteria where the periplasmic space is strictly delimited by two biological membranes. The CM and CW constitute the cell envelope (GO:0010339). The

situation at the CM is similar to what is described for diderm-LPS bacteria. Some proteins localized at the CM or CW can also be subunits of protein complex (GO:0043234) or be extrinsic to the CW (GO:0010339). **(C)** In diderm-mycolate bacteria, five clearly defined compartments are considered (i) the cytoplasm (CP; GO:0005737), (ii) the cytoplasmic membrane (CM; GO:0005886), (iii) the cell wall (CW; GO:0009275), (iv) the mycolate outer membrane (MOM) or mycomembrane (GO:0036407), and (v) the extracellular milieu (EM; GO:0005576). The cell envelope (GO:0010339) is constituted of the MOM, CW, and CM; a pseudo-periplasm might exist but remains to be evidenced. The situation at the CM is similar to what is described for diderm-LPS or diderm bacteria. The situation at the MOM (GO:0036407) can be discriminated further between location integral to the mycomembrane (GO:0036419), i.e., the MOM proteins (MOMPs), or extrinsic to the MOM (GO:0036420). Location at the cell surface (GO:0009986) refers (i) in diderm-LPS, to the OM and/or external side of the OM and is intended to proteins exposed externally (GO:0031242) or intrinsic to the OM (GO:0031230) (similar remarks apply to diderm-mycolate bacteria and the MOM), and (ii) in monoderm bacteria, to the CW and/or external side of the CM and is intended to proteins exposed externally or attached to the CW (GO:0009275) or the CM, i.e., integrated (GO:0005887), anchored (GO:0046658) or loosely bound (GO:0031232). As the CW is not a permeability barrier in monoderm bacteria but porous, the surface proteins do not necessarily have domains protruding from the confine of the cell envelope to interact with the external environment. Altogether with the subset of proteins localized extracellularly (GO:0005576), i.e., the exoproteome, these gene products (GO numbers in red) correspond to the extracytoplasmic proteins, i.e., the extracytoproteome (proteins depicted in blue).

Lee et al., 2012); a common misunderstanding is to make the T1SS synonymous to an ABC transporter (Desvaux et al., 2009b). The T2SS, also called the secretion-depend pathway (SDP) is a protein complex composed of around a dozen of proteins bridging the IM and OM to allow secretion of proteins translocated in the first place by the Sec or Tat export system (Sandkvist, 2001; Voulhoux et al., 2001; Cianciotto, 2005; Douzi et al., 2012; McLaughlin et al., 2012). It can be stressed again that referring to the general secretory pathway (GSP) or to the main terminal branch (MTB) for this system is nowadays obsolete and misleading (Desvaux et al., 2004b); as referred to in Pfam (Bateman et al., 2004), the naming of the different T2SS subunits as T2SE for instance is much more preferable than GspE (Peabody et al., 2003; Desvaux et al., 2004b). The T3SS is a highly complex molecular machine composed of at least 20 proteins and also one of the most extensively investigated protein secretion system (Ghosh, 2004; Cornelis, 2006, 2010; Minamino et al., 2008; Büttner, 2012). The T4SS is composed of around a dozen of proteins subunits forming a protein-conducting channel spanning the entire bacterial cell envelope (Christie et al., 2005; Waksman and Fronzes, 2010; Zechner et al., 2012); there is still much misunderstanding and confusion in part of the scientific literature about the T4SS and T2SS, as well as Type 4 pili (T4P), but it must be stressed again they can be clearly phylogenetically differentiated (Nunn, 1999; Planet et al., 2001; Mattick, 2002; Peabody et al., 2003; Hazes and Frost, 2008). T5SS refers to proteins depending on the Sec machinery for IM transit and then transported across the OM *via* a translocation unit formed by a β -barrel to complete secretion (Henderson et al., 2004; Leo et al., 2012); the BAM (β -barrel assembly machinery) complex as well as several periplasmic chaperones (namely SurA, Skp, DegP, and FkpA) are taking part to the secretion process across the OM (Desvaux

et al., 2004a; Knowles et al., 2009; Ruiz-Perez et al., 2010; Rossiter et al., 2011b; Leyton et al., 2012). The T6SS is a composite system of at least 13 protein subunits from various hypothetical phylogenetic origins, with essentially two subassemblies, i.e., one dynamic structure related the contractile bacteriophage tail-like structure and one cell-envelope-spanning membrane-associated complex (Cascales and Cambillau, 2012; Silverman et al., 2012). The T7SS corresponds to the chaperone-usher pathway (CUP) used for pilus assembly (Desvaux et al., 2009b; Waksman and Hultgren, 2009; Busch and Waksman, 2012); as further explained below, this system for diderm-LPS bacteria must not be mistaken with the diderm-mycolate bacterial “Type VII secretion system,” which is in fact the ESX (ESAT-6 system). The T8SS corresponds to the extracellular nucleation-precipitation pathway (ENP) (Barnhart and Chapman, 2006; Desvaux et al., 2009b; Blanco et al., 2012; Dueholm et al., 2012). The T9SS corresponds to the Por (porphyrin accumulation on the cell surface) secretion system (Sato et al., 2010, 2013; Shoji et al., 2011; McBride and Zhu, 2013). In diderm-LPS bacteria, the complement of the secretome potentially involved in bacterial colonization process gathers some secreted proteins and their associated secretion systems, which can be either the T1SS, T2SS, T3SS, T4SS, T5SS, T7SS, T8SS, or T9SS (**Figure 2**). In general, the secreted proteins involved in bacterial colonization are either cell-surface exposed single proteins or subunits of cell-surface supramolecular complexes, such as pili or flagella. In the rest of the manuscript, pili will be used as a generic term synonymous with fimbriae or curli (which are just some particular types of pili).

There are no counterparts to the molecular machineries required for transport across the OM of diderm bacteria in monoderm bacteria as this membrane is not present in the latter organisms. Consequently, the use of the numerical classification

Table 1 | Some key definitions in the field of bacterial protein secretion.

Terminology	Definition	Note ^a	References
Monoderm bacteria	Bacterial species exhibiting only one biological membrane, i.e. the CM.	Corresponds to the archetypal Gram-positive bacteria (i.e., with a CW) but also includes the <i>Mycoplasma</i> (devoided of a CW) since they also possess only one biological membrane. Monoderm bacteria are also called monodermata.	Shatalkin, 2004; Desvaux et al., 2009b; Sutcliffe, 2010; Gupta, 2011
Diderm-LPS bacteria	Bacterial species exhibiting two biological membranes and where the asymmetrical OM contains lipopolysaccharide (LPS) on the external side.	Corresponds to the archetypal Gram-negative bacteria. Some diderm bacteria lack LPS in their OM and are called simple-diderm bacteria. Diderm bacteria are also called didermata.	Shatalkin, 2004; Desvaux et al., 2009b; Sutcliffe, 2010; Gupta, 2011
Diderm-mycolate bacteria	Bacterial species exhibiting two biological membranes and where an outer lipid layer called MOM is composed of mycolic acid molecules arranged in a highly ordered form.	Corresponds to the archetypal acid-fast bacteria, e.g., <i>Mycobacterium</i> and <i>Corynebacterium</i> .	Shatalkin, 2004; Desvaux et al., 2009b; Sutcliffe, 2010; Gupta, 2011
Secretome	Concept for an integrated and global view of the protein secretion by considering protein routing, transport mechanisms, post-translational modifications, and protein subcellular location.	The secretome considers both secreted proteins and proteins constituting the secretion machinery (and associated maturation pathways). This original definition of the secretome has been somehow usurped, misused, and misunderstood by some authors in the literature. The secretome can be investigated by different “omics” approaches (i.e., proteogenomics, transcriptomics, proteomics, and meta-omics counterparts) but is not a proteome <i>per se</i> . The secretome is not the exoproteome, which is the most commonly investigated but only one of the complement of the secretome. Other complements of the secretome can be (i) the protein secretion systems, (ii) the cell-surface proteins (including single and supramolecular protein structure), or (iii) the lipoproteome, etc... Since it is not secretion <i>per se</i> , the secretome do not cover protein release upon molecular events that are not active translocation process, e.g., cell lysis (autolysis, allolysis, bacteriophage lysis), or membrane budding (vesicles).	Tjalsma et al., 2000; Antelmann et al., 2001, 2006; Van Dijk et al., 2001; Economou, 2002; Sarvas et al., 2004; Buist et al., 2006; Desvaux et al., 2009b
Secretion	Active transport from the interior to the exterior of the cell.	Applies to protein entirely outside of the outer-most lipid bilayer, including exoproteins, surface proteins, and cell-surface appendages (e.g., pili and flagella, cellulosomes). In diderm bacteria, secretion is mediated by specific translocon for transport across the OM (or MOM in diderm-mycolate bacteria) and cannot be defined by the translocons located at IM (CM). <i>Sensu stricto</i> , it does not cover molecular events that are not active translocation process and thus not secretion <i>per se</i> , e.g., cell lysis (autolysis, allolysis, bacteriophage lysis) or membrane budding (vesicles).	Desvaux et al., 2004b, 2009b; Economou et al., 2006
Export	Active transport across the CM	In monoderm bacteria, export and secretion are synonymous but not in diderm bacteria. In diderm bacteria, the Sec and Tat translocon directs proteins to the CM (IM) or periplasm but cannot ensure their secretion <i>sensu stricto</i> .	Desvaux et al., 2004b, 2009b; Economou et al., 2006

(Continued)

Table 1 | Continued

Terminology	Definition	Note ^a	References
Translocation	Active transport across a biological membrane.	Relative to the monoderm/diderm-LPS/diderm-mycolate bacteria trichotomy, it can occur at the CM (IM), OM, and/or MOM. Translocation event results in a change of subcellular location.	Economou et al., 2006; Desvaux et al., 2009b
Extracytoplasmic protein	Protein found outside the cytoplasm.	Corresponds to a protein either located elsewhere than the cytoplasm, e.g., intrinsic to a membrane (IMP, OMP, MOMP, lipoproteins), extrinsic to a membrane but on the side away from the cytoplasm, within the periplasm, associated to CW, at the cell surface or into the extracellular milieu. In other words, every protein except a cytoplasmic protein. This subset of proteins corresponds to the extracytoplasmic proteome.	Dinh et al., 1994; Skorko-Glonek and Sobiecka-Szkatula, 2008
Exoprotein	Protein present in the extracellular milieu, i.e., an extracellular protein.	Corresponds to free soluble protein found in the extracellular milieu or within a host cell. Some exoproteins are not necessarily secreted <i>sensu stricto</i> , since it can occur by not active translocation process e.g., cell lysis (autolysis, allolysis, bacteriophage lysis), or membrane budding (vesicles).	Pugsley and Francetic, 1998; Desvaux et al., 2009b
Cytoplasmic protein	Protein present in the cytoplasm, i.e., a cytoplasmic protein.	Corresponds to a cytosoluble protein, protein extrinsic to the CM located on the cytoplasmic side or a protein subunit of a CM protein complex but which subunit is entirely outside of the CM.	
Exoproteome	The subset of proteins present in the extracellular milieu (the exoproteins), i.e., the extracellular proteome.	It corresponds to one of the complements of the secretome.	Tjalsma, 2007; Desvaux et al., 2009b
Surface proteome	The subset of proteins present on the bacterial cell surface (the cell-surface proteins).	It corresponds to one of the complements of the secretome. The terminology “surface proteome” should be preferred to “surfaceome” (or “surfacome”) because it stresses it focuses on the protein content and not all the components on the bacterial cell surface (LPS, teichoic acids, exopolysaccharides, polyglutamate, etc. . .) as misleadingly suggested by the term “surfaceome.”	Cullen et al., 2005; Desvaux et al., 2006a, 2009b; Dreisbach et al., 2011; Voigt et al., 2012

^a IM, inner membrane; CM, cytoplasmic membrane; CW, cell wall; OM, outer membrane; MOM: mycolate outer membrane; IMP, integral membrane protein; OMP, outer membrane protein; MOMP, mycolate outer membrane protein.

for systems dedicated to protein secretion in diderm-LPS bacteria (i.e., protein transport from inside to outside the cell across the IM and OM) does not make any sense and cannot be applied to monoderm bacteria. However, this does not prohibit phylogenetic relationships between the protein translocation systems in monoderm and diderm-LPS bacteria. Indeed, they both possess a cytoplasmic membrane (also called IM in didermata) with some common protein transport systems allowing secretion in monoderm bacteria and export in diderm-LPS bacteria, respectively. The protein secretion system present in monoderm bacteria are (i) the Sec (secretion), (ii) the Tat (twin-arginine translocation), (iii) ABC protein exporter, (iv) the FPE (fimbriin-protein exporter), (v) the holin (hole forming), (vi) the Tra (transfer), misleadingly called the “Type IV-like secretion system” in

monoderm bacteria), (vii) the FEA (flagella export apparatus), and (viii) the Wss (WXG100 secretion system) (Desvaux et al., 2009b; Desvaux, 2012) (**Figure 3**). From the current knowledge in the field, the complement of the secretome potentially involved in the colonization process in monoderm bacteria gathers the Sec, FPE and FEA as well as some of their respective substrates (**Figure 3**). These systems secrete either cell-surface exposed single proteins or subunits of cell-surface supramolecular complexes, such as pili, cellulosome, or flagella (Renier et al., 2012).

In diderm-mycolate (archetypal acid-fast) bacteria, the term “Type VII secretion system” has also been coined to describe a Wss-like machinery (Economou et al., 2006; Abdallah et al., 2007; Schneewind and Missiakas, 2012). It must be stressed, this “Type VII secretion system” is restricted to diderm-mycolate bacteria

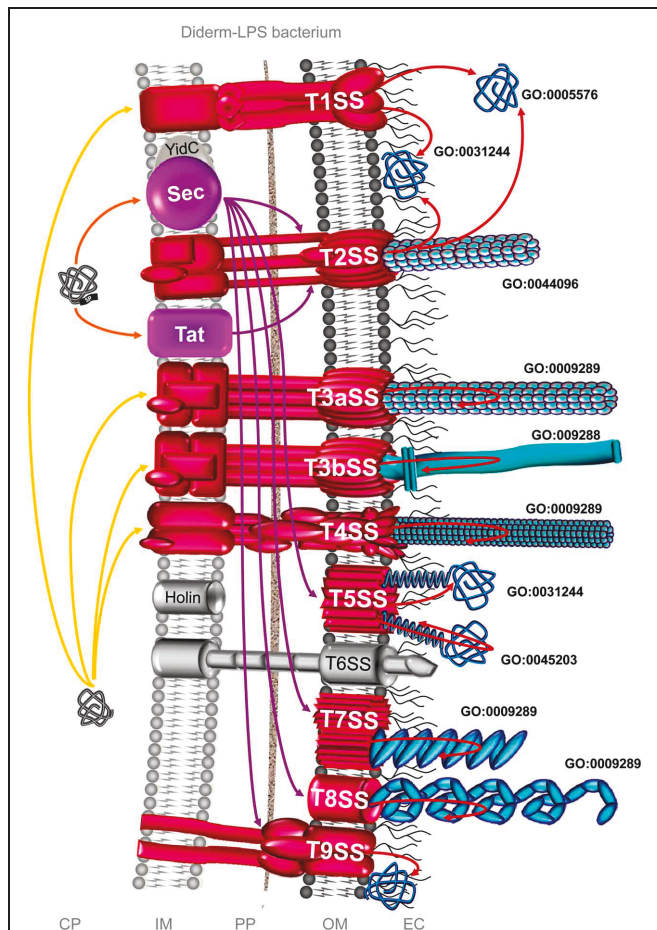


FIGURE 2 | The complement of the secretome involved in colonization process in diderm-LPS bacteria. Among the 9 distinct secretion systems

through which a secreted protein can be translocated across the OM in diderm-LPS bacteria, the T1SS, T2SS, T3SS, T4SS, T5SS, T7SS, T8SS, and T9SS can be involved in colonization process (depicted in red). Translocation machineries depicted in violet are protein export pathway participating to the protein transport of effectors involved in colonization process. The T1SS can secrete adhesins, which the release into the extracellular milieu (GO:0005576) and/or the association to the external side of the OM (GO:0031244) remain to be clarified. Besides the secretion of adhesion factors, the T2SS is involved in the formation of Type 4 pilus (GO:0044096), i.e., the T2SS subfamily c (T2cSS). The subfamily a of the T3SS (T3aSS) can be involved in the formation of pilus structure (GO:0009289), i.e., either the injectisome or the Hrp (hypersensitive response and pathogenicity) pilus, whereas the subfamily b of the T3SS (T3bSS) is involved in flagellum assembly (GO:009288). The T4SS is involved in the formation of pili (GO:0009289), either pilus T (T4aSS) or pilus F (T4bSS). The T5SS is involved in the secretion of adhesion either integral (GO:0045203) or extrinsic (GO:0031244) to the OM. The T7SS (CUP; chaperone-usher pathway) is involved in the formation of Type 1 pilus, and the T8SS (ENP; extracellular nucleation-precipitation pathway) in the formation of pilus of the type curli. The T9SS (Por secretion system) can secrete adhesins involved in gliding motility. In addition, some surface proteins could use systems as yet uncovered, the so-called non-classical (NC) secretion. Only branches corresponding to the complement of the secretome involved in bacterial colonization are colored. Extracytoplasmic proteins, i.e., single proteins and supramolecular protein structures, potentially involved in surface colonization are depicted in blue. Orange and yellow arrows indicate the routes for proteins targeted to the CM possessing or lacking an N-terminal SP, respectively. Violet arrows indicate the routes for exported proteins and red arrows for secreted proteins. CP, cytoplasm; IM, inner membrane; PP, periplasm; OM, outer membrane; EC, extracellular milieu; SP, signal peptide.

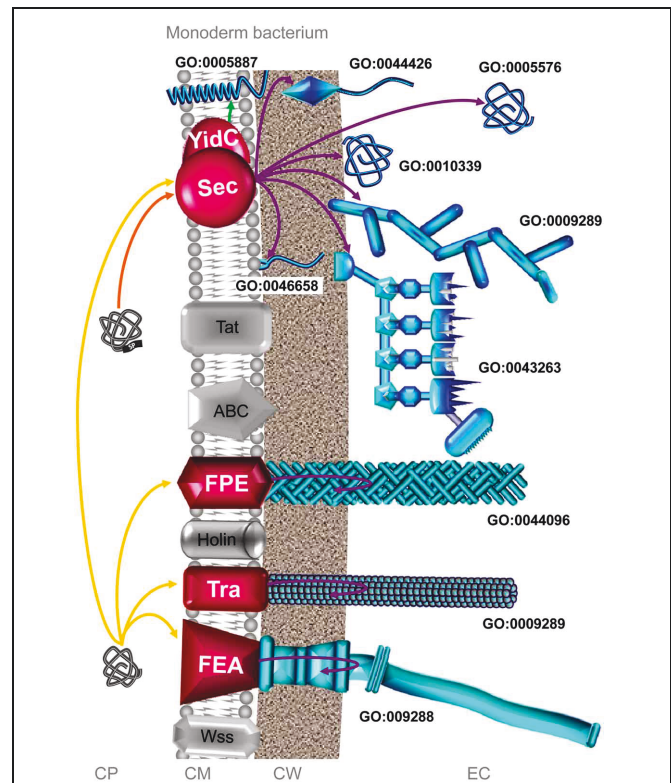


FIGURE 3 | The complement of the secretome involved in colonization process in monoderm bacteria. Among the 8 distinct secretion systems

through which a secreted protein can be translocated across the CM, the Sec, FPE, Tra, and FEA pathways can potentially be involved in colonization process in monoderm bacteria (depicted in red). The Sec pathway covers (i) integration of membrane protein (GO:0005887) via YidC, (ii) the anchoring to CM (GO:0046658) via the lipoprotein maturation pathway, (iii) the anchoring to the CW (GO:0044426) in a covalent or non-covalent manner, (iv) the association on the external side of the CW (GO:0010339), (v) the formation of cell surface supramolecular structure, namely pilus (GO:0009289) and cellulosome (GO:0043263), and (vi) protein secretion in the extracellular milieu (GO:0005576). It is worth noting that some proteins with no N-terminal SP can be translocated via Sec in a SecA2-dependent manner in monodermata (Rigel and Braunstein, 2008; Renier et al., 2013). The FPE is involved the formation of Type 4 pilus (GO:0044096). The Tra system (misleadingly called “Type IV-like secretion system” in monoderm bacteria) is involved in the formation of conjugative pili (GO:0009289). The FEA is involved in the secretion and assembly of the flagellum protein subunits (GO:009288). In addition, some surface proteins could use systems as yet uncovered, the so-called non-classical (NC) secretion. Extracytoplasmic proteins, i.e., single proteins and supramolecular protein structures, potentially involved in surface colonization are depicted in blue. Only branches corresponding to the complement of the secretome involved in bacterial colonization are colored. Orange and yellow arrows indicate the routes for proteins targeted to the CM possessing or lacking an N-terminal SP. Violet arrows indicate the routes for exported/secreted proteins (export and secretion are synonymous in monoderm bacteria). Green arrow indicates proteins integrated into the CM. CP: cytoplasm; CM: cytoplasmic membrane; CW: cell wall; EC, extracellular milieu; SP: signal peptide.

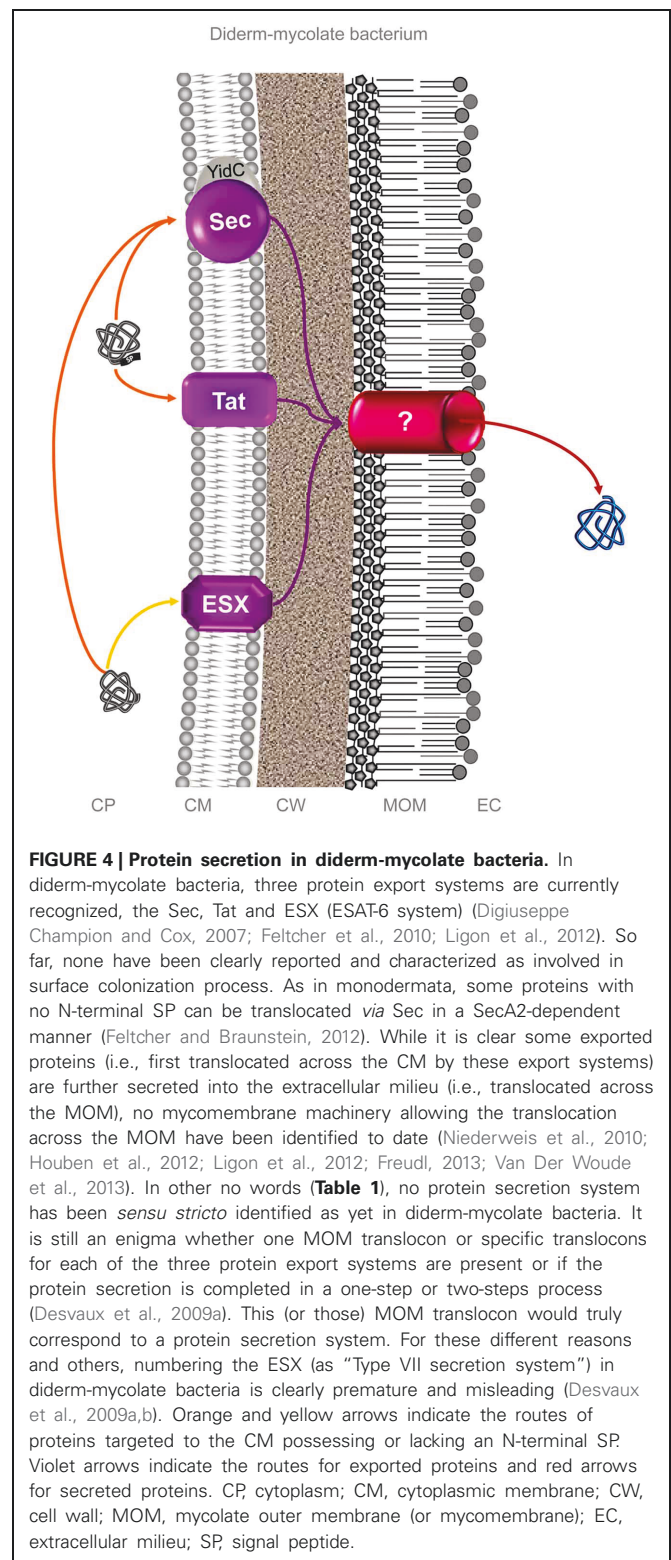
only since this system is absolutely not found in any archetypal diderm-LPS bacteria and, consequently, does not fit with standard numerical classification of protein secretion systems specifically designated for diderm-LPS (archetypal Gram-negative) bacteria (Salmond and Reeves, 1993; Sutcliffe, 2010). This “Type VII secretion system” further stands alone in diderm-mycolate

bacteria since there are neither T1SS, T2SS, T3SS, T4SS, T5SS, nor T6SS but only Sec and Tat systems (which are not included in the standard numerical classification) (Digiuseppe Champion and Cox, 2007) (**Figure 4**). The molecular machinery described for this diderm-mycolate bacterial “Type VII secretion system” is an export pathway (protein transport across the cytoplasmic membrane) just as the Sec and Tat systems are (Economou et al., 2006; Desvaux et al., 2009a; Houben et al., 2012). In fact, no translocon in the mycolate outer membrane (MOM), which would truly enable protein secretion and thus form a complete secretion pathway, has been uncovered as yet in diderm-mycolate bacteria [the secretion of proteins exported in the first instance by the Wss, Sec, and Tat could then be completed by the very same MOM translocon, or different MOM translocons specific to each of these export systems (Desvaux et al., 2009a)] (Niederweis, 2003; Converse and Cox, 2005; Ize and Palmer, 2006; Digiuseppe Champion and Cox, 2007; Song et al., 2008; Desvaux et al., 2009a; Niederweis et al., 2010; Stoop et al., 2012; Freudl, 2013; Van Der Woude et al., 2013). This “Type VII secretion system” nomenclature in diderm-mycolate bacteria is clearly not compatible with the numerical classification basically designed to describe OM translocation systems in diderm-LPS bacteria (Salmond and Reeves, 1993; Economou et al., 2006; Desvaux et al., 2009a). All-in-all, using the “Type VII secretion system” denomination for the phylogenetically related secretion systems in monoderm bacteria (i.e., the Wss) is very much confusing since it does not align with the other secretion systems present such as Sec and Tat (which do not withstand the numerical classification terminology) (Desvaux et al., 2004b, 2009a). When designating this “Type VII secretion system” it is then highly advisable to clearly specify it relates to diderm-mycolate bacteria only to prevent any confusion with the unrelated T7SS in diderm-LPS bacteria. Ultimately, its use should be refrained in favor of the “ESX (ESAT-6 system)” designation in diderm-mycolate bacteria (archetypal acid-fast bacteria) and/or the generic “WXG100 secretion system (Wss)” designation especially relevant to monodermata (archetypal Gram-positive bacteria). At the moment only protein export systems have been reported in diderm-mycolate bacteria (Digiuseppe Champion and Cox, 2007; Feltcher et al., 2010; Ligon et al., 2012) but *sensu stricto* no protein secretion system has been identified (**Figure 4**). As for the T1SS to T9SS in diderm-LPS bacteria, only the identification of translocon components at the MOM would truly permit to define a protein secretion system (Desvaux et al., 2009a), which at worse could be the unique terminal branch for all the three export systems (Sec, Tat, and ESX) in diderm-mycolate bacteria (**Figure 4**).

For effectors involved in bacterial colonization, this review will especially focus on protein secretion systems in diderm-LPS (archetypal Gram-negative) and monoderm (archetypal Gram-positive) bacteria since none has been characterized as yet in diderm-mycolate (archetypal acid-fast) bacteria.

SECRETED PROTEINS INVOLVED IN SURFACE COLONIZATION IN DIDERM-LPS BACTERIA

Out of the 9 distinct protein secretion systems through which a secreted protein can be translocated across the OM, only the T6SS



has never been reported so far to be involved in bacterial adhesion and/or biofilm formation in diderm-LPS bacteria (**Figure 2**). The T2SS, T3SS, T4SS, and T5SS are further divided into different subtypes.

ADHESINS SECRETED VIA THE T1SS

The T1SS has been demonstrated to allow the secretion of large bacterial adhesins belonging to the Bap family (Delepeleire, 2004; Lasa and Penadés, 2006; Latasa et al., 2006). In *Salmonella enterica*, SiiE (*Salmonella* intestinal infection E) and BapA (Biofilm-associated protein A) have been characterized. SiiE is a very large adhesin of 600 kDa of *Salmonella enteric* required for adhesion to epithelial cells. SiiE would essentially exist as an exoprotein that may only be loosely associated the OM but upon host cell contact it is retained on the bacterial cell surface (Wong et al., 1998; Morgan et al., 2004; Gerlach et al., 2007). The ability of SiiE to bind Ca^{2+} ions would confer a rigid rod-like habitus that is required for reach out beyond the LPS and initiates bacterial adhesion to polarized host cells (Wagner et al., 2011; Griessl et al., 2013). BapA is a 386 kDa protein allowing homotypic adhesion during *Salmonella* biofilm formation and is involved in pathogenesis, especially internalization and invasion of intestinal epithelium (Latasa et al., 2005; Jonas et al., 2007; Biswas et al., 2011; Suez et al., 2013). BapA is secreted extracellularly where it remains in loose association with the bacterial cell surface. In *Pseudomonas fluorescens*, LapA (large adhesion protein A) is the largest T1SS-dependent adhesin uncovered so far, with a molecular weight estimated at 888 kDa (Hinsa et al., 2003; Hinsa and O'toole, 2006). LapA enables irreversible adhesion and biofilm formation on abiotic surfaces but also adhesion to corn seeds in *Pseudomonas putida* (Mus20 or Mus24; mutant unattached to seeds) (Espinosa-Urgel et al., 2000; Huber et al., 2002; Hinsa et al., 2003; Ivanov et al., 2012; Zhang et al., 2013). LapA is found both in the extracellular milieu and in a loose association with the bacterial cell surface but not in the OM (Hinsa et al., 2003). In *Burkholderia cepacia*, the loss of Bap resulted in decreased surface hydrophobicity and in colony conversion to a rough morphotype (Huber et al., 2002). Additional adhesins of the Bap family, presumably secreted via a T1SS as supported by the genomic context, have been characterized in diderm-LPS bacteria, namely VPA1445 (*Vibrio parahaemolyticus* locus 1445) (Enos-Berlage et al., 2005) and YeeJ (systematic nomenclature) in *Escherichia coli* (Roux et al., 2005). In all cases, the mediation of adhesion via an extracellular protein is difficult to conceptualize and still demands experimental investigations for validation (Gerlach and Hensel, 2007).

TYPE 4 PILUS (T4P) AND COLONIZATION FACTOR VIA THE T2SS

Based on phylogenetic analysis, T2SSs were further subdivided into the subfamilies T2aSS and T2bSS, corresponding, respectively, to Xcp (*Xanthomonas campestris* general secretion pathway) and Hxc (homolog to Xcp) systems (Filloux, 2004; Michel and Voulhoux, 2009; Durand et al., 2011). At the same period, it was proposed to include the Type 4 piliation (T4P) system (Desvaux et al., 2009b) within the T2SS since the T4P is homologous to the Xcp and falls into discrete phylogenetic cluster (Planet et al., 2001; Peabody et al., 2003; Tomich et al., 2007; Ayers et al., 2010). To make the nomenclature coherent, we propose to classify those systems as T2aSS for the classical SDP extensively investigated in *Pseudomonas aeruginosa*, i.e., Xcp type system (Voulhoux et al., 2001), T2bSS for the Hxc type system (Durand et al., 2011), and T2cSS for the T4P system (Mattick, 2002). The Type 4 pili

promote initial bacterial attachment and are of importance for the early stages of biofilm development, i.e., microcolony formation (Hahn, 1997; Giltner et al., 2006; Burrows, 2012). The Type 4 pili are subdivided into two main classes assembled by the T2cSS (Strom and Lory, 1993; Kachlany et al., 2001; Pelicic, 2008), i.e., (i) the T4a pilins, and (ii) the T4b pilins (Strom and Lory, 1993; Skerker and Shapiro, 2000; Craig et al., 2004; Craig and Li, 2008; Giltner et al., 2012). Besides adherence, the T4aP are responsible for a number of related processes associated with bacterial motility, including twitching, swarming, crawling, walking, and slingshotting, which participate to biofilm development (Desvaux et al., 2005a; Burrows, 2012; Conrad, 2012).

The colonization factor GbpA (N-acetylglucosamine-binding protein A) from *Vibrio cholerae* is secreted in a T2SS-dependent manner in the extracellular milieu (Kirn et al., 2005). This protein is required for efficient environmental colonization by organisms such as zooplankton, and also intestinal colonization, especially of human epithelial cells, by binding to a sugar present on both surfaces (Stauder et al., 2012). GbpA has a modular architecture with chitin-binding and mucin-binding domains enabling attachment to different host surfaces and other domains binding to the bacterial cell surface (Wong et al., 2012). Despite its secretion into the extracellular milieu, GbpA thus has a bridging function between *V. cholerae* and its host allowing efficient colonization of chitinous exoskeletons of arthropods or the intestinal epithelium.

T3SS: INJECTISOME, HRP (HYPERSENSITIVE RESPONSE AND PATHOGENICITY) PILUS AND FLAGELLUM

The T3SSs can be subdivided into (i) the non-flagellar T3SS, i.e., the T3aSS, involved in the assembly of the injectisome or Hrp (hypersensitive response and pathogenicity) pilus, and (ii) the flagellar T3SS, i.e., T3bSS, responsible for assembly of the flagellum (Tampakaki et al., 2004; Journet et al., 2005; Pallen et al., 2005; Desvaux et al., 2006b).

While acting primarily as a molecular syringe for injecting protein effectors directly into the cytosol of a host cell (Cornelis, 2006, 2010), the injectisome (T3aSS) mediates intimate bacterial adhesion and colonization of cells such as gut epithelial cells (Garmendia et al., 2005). In enterohemorrhagic *Escherichia coli* (EHEC), it was further demonstrated to play a role in the adhesion to lettuce leaves with a marked tropism for the stomata, and thus to contribute to the transmission of this food-borne pathogen to humans (Shaw et al., 2008; Berger et al., 2010). While the injectisome is found in animal-pathogenic bacteria, the Hrp pilus (T3aSS) is found in plant-pathogenic bacteria (Tang et al., 2006; Tampakaki et al., 2010). By injecting T3SS effectors into the host cell, the Hrp pilus allows the phytopathogens to subdue the vegetable-cell response (Zhou and Chai, 2008) but also participate to plant colonization through initial bacterial adhesion (Darsonval et al., 2008; Correa et al., 2012).

The flagella (T3bSS) are involved in cell motility through swimming and/or swarming, where it can contribute to bacterial colonization by bringing the cell in contact with biotic or abiotic surfaces by chemotaxis for instance, or by translocating the bacterial cells over a surface, respectively (Pratt and Kolter, 1998; Harshey, 2003; Verstraeten et al., 2008). Besides, the flagella mediate bacterial adhesion directly to biotic and abiotic surfaces, e.g.,

to plasticware, human mucus and mucin, epithelial cells, enterocytes, vegetables (Grant et al., 1993; Ramphal et al., 1996; Lillehoj et al., 2002; Kirov et al., 2004; Berger et al., 2009b; Mahajan et al., 2009; Shaw et al., 2011; Tran et al., 2011; Bucior et al., 2012; Troge et al., 2012). Interestingly in *Salmonella enterica*, adhesion to salad leaves mediated by the flagella was shown to be strain-specific (Berger et al., 2009a). *Salmonella enterica* can further internalize into vegetable leaves through flagella-chemotaxis via open stomata (Kroupitski et al., 2009a). Although the internalization is variable in leafy vegetables and fresh herbs (Golberg et al., 2011; Kroupitski et al., 2011), the flagella is clearly an important determinant for colonization of some vegetables and can thus contribute to the transmission of this food-borne pathogen to humans even after washing (Berger et al., 2009a; Kroupitski et al., 2009a,b).

PILI ASSEMBLED BY THE T4SS

The T4SSs are broadly subdivided into (i) T4aSS corresponding to the prototypical VirB/D4 complex extensively investigated in *Agrobacterium tumefaciens*, and (ii) T4bSS corresponding to the prototypical F-conjugal transfer system of the self-transmissible IncI plasmid composed of the widely conserved Tra (transfer) proteins also present in monoderm bacteria (Christie and Vogel, 2000; Sexton and Vogel, 2002; Lawley et al., 2003; Harris and Silverman, 2004; Christie et al., 2005; Hazes and Frost, 2008; Voth et al., 2012). The T4SS allows transport of proteins into prokaryotic or eukaryotic cells. This system is also involved and ancestrally related to bacterial conjugation with the transport of DNA as nucleo-protein complex (Lawley et al., 2003). Rather than focusing on protein secretion system only, the nomenclature is quite confusing at the moment since it mixes up homologous (phylogenetic) and analogous (functional) based-classifications on aspects related to DNA transport as well as the systems present in monoderm and diderm-LPS bacteria (Alvarez-Martinez and Christie, 2009; Wallden et al., 2010); in addition, some phylogenetic relatedness between the T2SS assembling Type 4 pili (T4P) and T4SS assembling conjugative pili of Type 2 (Ottow, 1975) accentuates the complexity of the situation (Hazes and Frost, 2008). As it has been done previously for other protein secretion systems (Henderson et al., 2000; Desvaux et al., 2004b, 2006b, 2009b), this stresses the need to clarify the classification within the T4SS respective to the current terminology and ontology issues in the field of bacterial protein secretion, which should be rather based on homology and phylogeny rather than analogy (Desvaux et al., 2006b, 2009b). Regarding bacterial colonization, the F episome (natural conjugative plasmid) was demonstrated to induce the formation of a thick biofilm in *E. coli* and that conjugative pilus synthesis from the *tra* operon of the F plasmid was required for biofilm formation (Ghigo, 2001). In fact, the bacterial cells harboring the F episome had increased adhesion ability thanks to the expression of conjugative pili. It appeared these plasmids can be further transferred by conjugation from these donor cells to recipient cells. Then, the increasing proportion of transconjugant cells within the biofilm further improved the colonization ability of the bacterial population. While the positive influence of conjugative plasmid on biofilm formation was reported in other species (Bahl et al., 2007; Burmolle et al., 2008), it can also

have adverse effects in complex multi-species biofilms and would depend on the composition of the bacterial community (Roder et al., 2013).

Phylogenetic analyses revealed the tight-adherence (Tad) system originally uncovered in *Aggregatibacter* (formerly *Actinobacillus*) *actinomycetemcomitans* was actually related to the T4SS and would constitute a major subfamily (Tomich et al., 2007). The Tad system is involved in the piliation of the Fli (fimbrial low-molecular-weight protein). Those pili are composed of Type 4 pilin of the subfamily b (Giltner et al., 2012). Fli pili are highly adhesive and essential for biofilm formation (Kachlany et al., 2001). Rather than binding to specific receptor(s), the bacterial-bacterial and bacterial-host interactions lead to strong attachment. Bacterial aggregation is mediated by the association of pilus fibers, which leads to pilus bundling. This system is widespread among diderm-LPS bacterial species, including the genera *Haemophilus*, *Pasteurella*, *Pseudomonas*, *Yersinia*, or *Caulobacter* (Kachlany et al., 2000, 2001; Bernard et al., 2009).

CELL-SURFACE ADHESINS OF THE T5SS

T5SSs are subdivided into five subtypes (Henderson et al., 2000, 2004; Desvaux et al., 2003, 2004a; Henderson and Desvaux, 2004; Salacha et al., 2010; Leo et al., 2012), (i) the T5aSS corresponding to the classical autotransporter or autotransporter of Type 1 (AT-1), (ii) the T5bSS corresponding to the two-partner secretion (TPS), (iii) the T5cSS corresponding to the trimeric autotransporter or autotransporter of Type 2 (AT-2), (iv) the T5dSS corresponding to hybrid autotransporter between AT-1 and TPS, or autotransporter of Type 3 (AT-3), and (v) the T5eSS corresponding to invasins/intimin family of inverted autotransporter or autotransporter of Type 4 (AT-4). No function in bacterial colonization has been reported as yet for the hybrid autotransporters (T5dSS) (Salacha et al., 2010).

Adhesins of the classical autotransporter pathway (T5aSS)

Among the classical autotransporters (T5aSS), the SAATs (Self-associating autotransporters) actively participate in biofilm development and encompass the Ag43 (antigen 43), AIDA (adhesin involved in diffuse adherence), and TibA (enterotoxigenic invasion locus b protein A) (Klemm et al., 2006; Van Der Woude and Henderson, 2008). The SAATs promote bacterial auto-aggregation and interaction with other SAAT partners, for instance between AIDA and Ag43, leading to the formation of mixed bacterial cell aggregates (Sherlock et al., 2004). Self-recognition is sensitive to environmental conditions, such as pH or the presence of bile salt (Klemm et al., 2004; Sherlock et al., 2004, 2005; Girard et al., 2010) and intercellular aggregation can be abolished by pili or exopolymers, which compromise SAATs interaction (Klemm et al., 2006). SAATs further mediate and enhance biofilm formation (Klemm et al., 2006). Ag43, however, would not contribute significantly to intestinal colonization (Parham et al., 2005; De Luna et al., 2008). Besides SAATs, several other AT-1s act as adhesins (Henderson et al., 2004; Girard and Mourez, 2006). YfaL, YpjA, and YcgV (*E. coli* systematic nomenclature) are involved in initial bacterial adhesion and biofilm formation (Roux et al., 2005). BrkA (*Bordetella* resistance to killing protein A) (Ewanowich et al., 1989; Fernandez

and Weiss, 1994) and Pertactin (Leininger et al., 1991) promote adhesion to different mammalian cells. Aae (*A. actinomycetem-comitans* epithelial cell binding) (Rose et al., 2003; Fine et al., 2005), App (adhesion and penetration protein) (Serruto et al., 2003), CapA (*Campylobacter* adhesion protein A) (Ashgar et al., 2007), EspP (extracellular serine protease P) (Dziva et al., 2007; Puttamreddy et al., 2010), McaP (*Moraxella catarrhalis* adherence protein) (Timpe et al., 2003; Lipski et al., 2007), PmpD (Polymorphic membrane protein D) (Wehrl et al., 2004), rOmpB (Rickettsia outer membrane protein B) (Uchiyama et al., 2006), and Sab (STEC autotransporter mediating biofilm formation) (Herold et al., 2009), enable adherence to epithelial cells; for some of them it was further demonstrated they participated to intestinal colonization and/or biofilm formation, namely CapA, EspP and Sab. Hap (*Haemophilus* autotransporter protein) further mediates microcolony formation as well as binding to fibronectin, laminin, and collagen (Fink et al., 2003), whereas MisL (membrane insertion and secretion protein L) and ShdA (shedding protein A) bind to collagen and fibronectin (Kingsley et al., 2000, 2002, 2004; Dorsey et al., 2005). EhaA (EHEC autotransporter A) (Wells et al., 2008), EhaB (Wells et al., 2009), and EhaJ (Easton et al., 2011) are MSCRAMM (microbial surface components recognizing adhesive matrix molecules) proteins binding differentially to various extracellular matrix (ECM) proteins and contributing to biofilm formation (Chagnot et al., 2012, 2013). Tsh (temperature-sensitive haemagglutinin) also binds to ECM proteins and further adheres to red blood cells and hemoglobin (Kostakioti and Stathopoulos, 2004). AlpA (adherence-associated lipoprotein A) (Odenbreit et al., 1999), BabA (blood group antigen-binding adhesin) (Ilver et al., 1998), and SabA (sialic acid-binding adhesin) (Mahdavi et al., 2002) are involved in adhesion to human gastric epithelia by the fucosylated Lewis b histo-blood group antigen for BabA and by sialyl-dimeric-Lewis x glycosphingolipid for SabA.

Of major importance in the colonization process, the expression of many autotransporters (such as Ag43) is subjected to phase variation (Henderson et al., 1999; Van Der Woude and Henderson, 2008; Rossiter et al., 2011a). In addition, proteolytic processing of the passenger domain of the autotransporter on the bacterial cell surface can have a regulatory function on bacterial adhesion (Leyton et al., 2012). In fact, the passenger domain of AT-1s can either remain attached or cleaved off the translocation unit upon OM translocation. This proteolytic cleavage can arise following different scenarios (Leyton et al., 2012) but basically the cleavage can either be intramolecular and autocatalytic or intermolecular resulting in the release of passenger domain into the extracellular milieu, for example the SPATEs (serine protease autotransporters of *Enterobacteriaceae*). Some processed AT-1s, though, remain strongly associated with their cognate translocation unit, e.g., the adhesins Hap, pertactin, AIDA, or Ag43. It also appeared, the protein function is not modified when the cleavage is abolished (Charbonneau et al., 2009), which questions the purpose or benefit of this processing. For Hap, however, it was demonstrated the cleavage influenced the adherence to epithelial cells as well as adhesion to ECM proteins (Fink et al., 2002, 2003). Interestingly, NalP (*Neisseria* autotransporter lipoprotein), which is subjected to phase-variable expression by slipped-strand

mispairing (Saunders et al., 2000) and modifies the cleavage patterns of other surface-associated proteins (including some other autotransporters and surface-exposed lipoproteins) (Van Ulsen et al., 2003, 2006; Roussel-Jazede et al., 2010; Serruto et al., 2010), affects in turn adhesion and biofilm formation (Arenas et al., 2013). Similar proteins are found in other bacterial pathogens.

Adhesins of the TPS pathway (T5bSS)

In the TPS (T5bSS), the passenger domain or exoprotein (TpsA) is translated separately from its cognate translocation unit (TpsB) (Jacob-Dubuisson et al., 2004). The exoprotein FHA (filamentous haemagglutinin) is a multifaceted adhesin and major attachment factor of *Bordetella* spp required for colonization of the lower respiratory tract (Locht et al., 1993). While part of the exoprotein is found to be surface-associated, another part is released into the extracellular milieu. Contrary to our previous understanding, it appeared that the C-terminus of FHA is oriented away from the bacterial cell surface and is required for adherence to epithelial cells (Mazar and Cotter, 2006). The exoproteins HMW1 (high-molecular-weight protein 1) and HMW2 of *Haemophilus influenza* are crucial colonization factors involved in bacterial adherence to a variety of respiratory epithelial cell types (St Geme et al., 1993; St Geme and Yeo, 2009). HrpA (hemagglutinin/hemolysin-related protein A) from *Neisseria meningitidis* and MhaB1 (*M. catarrhalis* FhaB-like protein 1) from *Moraxella catarrhalis* mediate binding to human epithelial cells (Schmitt et al., 2007). In the phytopathogen *Xanthomonas axonopodis*, XacFhaB (*X. axonopodis* pv. *citri* filamentous haemagglutinin-like protein B) is required for leaf tissue colonization, especially adhesion and biofilm formation (Gottig et al., 2009). In *Pseudomonas putida*, HlpA (haemolysin-like protein A) would play a direct role in the bacterial cell-root surface interaction (Molina et al., 2006). HecA (hemolysin-like *E. chrysanthemi* protein A) from *Erwinia chrysanthemi* contributes to the attachment and cell aggregation on leaves (Rojas et al., 2002). Interestingly, the exoprotein EtpA (ETEC two-partner secretion protein A) mediates bacterial adhesion by bridging the flagella with the host cells (Roy et al., 2009).

Adhesins of the trimeric autotransporter pathway (T5cSS)

A cardinal feature of the T5SS is a functional OM translocation unit, e.g., located at the C-terminus in monomeric classical autotransporter (T5aSS) or the TspB for the T5bSS. In the T5cSS, the C-terminal translocation unit is formed upon trimerization and serves as an OM anchor for cell-surface exposure of the passenger domains (Cotter et al., 2005; Lyskowski et al., 2011). YadA (*Yersinia* adhesin A) is the prototypical member of this T5SS subfamily (Hoiczky et al., 2000; El Tahir and Skurnik, 2001; Eitel and Dersch, 2002; Nummelin et al., 2004; Leo et al., 2010). The high-molecular weight trimer formed by the passenger domains protrudes from the OM and consists of three domains, the stalk, necks and head (Linke et al., 2006; Hartmann et al., 2012). Differences in the size of the protruding AT-2s result from the length of their stalks (Linke et al., 2006). The head is involved in binding to some ECM proteins, including collagen, and makes YadA a major adhesion factor. All trimeric

autotransporters (T5cSS) characterized to date are involved in bacterial adhesion (Linke et al., 2006). Besides YadA, several other AT-2s have been characterized to date in this expanding protein family (Cotter et al., 2005), e.g., Hia (*Haemophilus influenzae* adhesin) (Hoiczky et al., 2000; St Geme and Cutter, 2000), UspA (ubiquitous surface protein A) of *Moraxella catharralis* (Connors et al., 2008), Hag (hemagglutinin) (Pearson et al., 2002; Bullard et al., 2007), NadA (*Neisseria* adhesin A) (Comanducci et al., 2002), BadA (*Bartonella henselae* adhesin) (Riess et al., 2004), NcaA (necessary for collagen adhesion A) (Fulcher et al., 2006), NhhA (*Neisseria* Hsf homologue protein A) (Scarselli et al., 2006), DsrA (*ducreyi* serum resistance A) (Leduc et al., 2008), EhaG (EHEC adhesin G) (Valle et al., 2008), SadA (*Salmonella* adhesin A) (Raghunathan et al., 2011), or UpaG (UPEC adhesin G) (Totsika et al., 2012). Each of these trimeric autotransporters are quite multifunctional and in terms of bacterial colonization they can be involved in autoagglutination, hemagglutination, ECM-binding, and/or adhesion to epithelial cells (El Tahir and Skurnik, 2001; Linke et al., 2006).

Intimins/invasins of the inverted autotransporter pathway (T5eSS)

Intimins/invasins (T5eSS) belong to a large and novel subfamily of the T5SS (Tsai et al., 2010; Leo et al., 2012; Oberhettinger et al., 2012). Compared to the classical autotransporter, the secretion mechanism is inverted in the sense, the translocation unit is located at the N-terminal instead of the C-terminal end of the monomeric autotransporter. The intimin/invasin family regroups adhesins that mediate bacterial adhesion and/or invasion of their host cells. In *Yersinia* spp., invasin binds to mammalian cell receptors of the integrin family (Isberg et al., 2000; Palumbo and Wang, 2006). In enteropathogenic *E. coli*, intimin mediates intimate adherence between the bacteria and the host cells by interacting with Tir (translocated intimin receptor) (Frankel and Phillips, 2008). Tir is a T3aSS-secreted protein, which integrates into the plasma membrane of the host cell (Devinney et al., 1999); upon binding to intestinal epithelial cells, the intimin-Tir interaction leads to the formation of actin pedestals beneath bound bacteria (Campellone and Leong, 2003; Brady et al., 2011). Interestingly, intimin can further promote intestinal colonization in a Tir-independent way (Mallick et al., 2012).

PILI ASSEMBLED BY THE T7SS

In diderm-LPS bacteria, the T7SS corresponds to the chaperone-usher pathway (CUP) (Desvaux et al., 2009b); as explained above it has nothing to do with the “Type VII secretion system” of diderm-mycolate (archetypal acid-fast) bacteria, i.e., the ESX. The T7SS is involved in the OM secretion and assembly of pili, which have been named very differently from one bacterial species to another (Sauer et al., 2004; Zav'yalov et al., 2010; Busch and Waksman, 2012; Thanassi et al., 2012), e.g., Afa/Dr (diffuse adherence fibrillar adhesin/Dr blood group antigen), AF/R1 (adhesive fimbriae on RDEC-1), AafD (aggregative adherence fimbriae), P pili, Type 3 fimbriae, Lpf (long polar fimbriae) or CS (coli surface) pili. More rationally, those pili can be categorized into (i) monoadhesive pili (represent the majority of cases), which display either thick rigid or thin flexible morphology and present a single adhesive domain at the tip, and (ii) polyadhesive

pili, which display non-pilar, amorphous or capsule-like morphology and present two independent binding sites specific to different host-cell receptors for each of the subunits composing the organelle (Zav'yalov et al., 2010). Monoadhesive pili are also named FGS (F1-G1 short loop short) chaperone-assembled monoadhesin (e.g., Lpf) (Hung et al., 1996), and polyadhesive pili can either be FGL (F1-G1 long loop) chaperone-assembled polyadhesin (e.g., Afa/Dr) or FGS chaperone-assembled polyadhesin (e.g., AF/R1) (Zav'yalov et al., 2007). Of note, the so-called “alternate chaperone-usher pathway” (Soto and Hultgren, 1999) is obsolete and now included within the T7SS (Nuccio and Baumber, 2007; Poole et al., 2007; Zav'yalov et al., 2010; Thanassi et al., 2012). Pili secreted and assembled by the T7SS are involved bacterial adhesion, interbacterial interactions, aggregation, thereby promoting biofilm formation (Zav'yalov et al., 2010; Thanassi et al., 2012). Those pili can also initiate contact with host-cell receptors and mediate colonization of host cell surfaces. Interestingly, the adhesion of FimH pili (subfamily 2.5 FGS chaperone-assembled monoadhesins-5-1) is modulated by the shear force (Nilsson et al., 2006; Yakovenko et al., 2008; Le Trong et al., 2010).

PILI ASSEMBLED BY THE T8SS

The T8SS corresponds to the extracellular nucleation-precipitation (ENP) pathway (Desvaux et al., 2009b) involved in the OM secretion and assembly of thin and aggregative pili called curli (Hammar et al., 1996; Barnhart and Chapman, 2006; Dueholm et al., 2012; Hammer et al., 2012). Curli are functional amyloid fibers (Epstein and Chapman, 2008). They are involved in cell aggregation, bacterial adhesion and the formation of mature biofilms (Fronzes et al., 2008). They can actually change bacterial surface properties thereby enhancing adherence and attachment to surfaces. In the course of sessile development they constitute a significant part of the proteinaceous component of the biofilm matrix (Blanco et al., 2012). Curli also mediate host cell–bacteria interactions during infection (Wurpel et al., 2013).

ADHESINS SECRETED VIA THE T9SS

The T9SS corresponds to the Por (porphyrin accumulation on the cell surface) secretion system (Sato et al., 2010, 2013; Shoji et al., 2011; McBride and Zhu, 2013). This secretion system has been essentially investigated in *Porphyromonas gingivalis* and *Flavobacterium johnsoniae* and seems restricted to members of the phylum *Bacteroidetes* (McBride and Zhu, 2013). Respective to surface colonization process, the T9SS is required for the secretion of cell-surface motility adhesins, namely SprB (colony-spreading protein B) and RemA (redundant motility protein A), but also some hemin-binding proteins (Nelson et al., 2007; Shoji et al., 2011; Shrivastava et al., 2012; Sato et al., 2013). The cell-surface adhesin SprB allows attachment to the substratum but is also required for efficient gliding as it is propelled along a closed helical loop track, generating rotation and translation of the bacterial cell (Nelson et al., 2007; Nakane et al., 2013a,b). While SprB is involved in movement over agar, RemA is involved in movement over surfaces coated with *F. johnsoniae* polysaccharide (Shrivastava et al., 2012). Gliding involves the rapid movement of the semiredundant motility adhesins SprB and RemA along

the cell surface (Shrivastava et al., 2013). T9SS-secreted proteins exhibit C-terminal domains (CTDs) considered essential for attachment to the bacterial cell surface by an A-LPS anchor containing anionic polysaccharide repeating units (Kondo et al., 2010; Shoji et al., 2011; Slakeski et al., 2011; Shrivastava et al., 2013). CTD region function as a recognition signal for the T9SS and glycosylation occurs after removal of the CTD region (Shoji et al., 2011; Glew et al., 2012).

SECRETED PROTEINS INVOLVED IN SURFACE COLONIZATION IN MONODERM BACTERIA

In monoderm bacteria, no protein secreted *via* the Tat, ABC exporter, holin, or Wss pathways has been reported as yet to be involved in bacterial adhesion and/or biofilm formation (Figure 3) but some proteins secreted by Sec, FPE, Tra and FEA have.

SEC-SECRETED PROTEINS INVOLVED IN BACTERIAL COLONIZATION

Once translocated by Sec, secreted proteins can have radically different locations in monoderm bacteria as they can either be (i) integrated or anchored to the CM, i.e., IMPs (inner membrane proteins) and lipoproteins, respectively, (ii) associated with the CW, i.e., parietal proteins (CW-proteins), (iii) subunits of pili, or (iv) released into the extracellular milieu and beyond (e.g., into a host cell) (Desvaux et al., 2006a, 2009b; Renier et al., 2012) (Figures 1, 3). Sec-secreted CW-proteins can either be anchored (i) covalently to the CW by a sortase if they exhibit an LPXTG domain (Ton-That et al., 2004; Schneewind and Missiakas, 2012), or (ii) non-covalently if they exhibit a CWBD (cell-wall binding domain), i.e., SLHD (S-layer homology domain), CWBD1 (CWBD of Type 1), CWBD2, LysM (lysine motif), WXL or GW motifs (Desvaux et al., 2006a). While no protein located at the CM has been reported to be involved in bacterial colonization in monodermata, several proteins located at the CW or forming cell-surface supramolecular complexes were reported to participate to this process.

Colonization factors located at the cell wall

In monoderm bacteria, proteins of the Bap family exhibit a C-terminal LPXTG domain enabling their covalent anchoring to the CW by sortases (Lasa and Penadés, 2006; Latasa et al., 2006). While Bap is involved in biofilm formation in *S. aureus* (Cucarella et al., 2001), the *bap* gene has never been found in *S. aureus* human isolates but only in isolates associated with ruminant mastitis (Lasa and Penadés, 2006). Strains encoding *bap* show lower adherence to some ECM proteins (fibrinogen and fibronectin) but also epithelial cell cultures, suggesting Bap might act as an anti-attachment factor preventing initial attachment to host tissues (Cucarella et al., 2002). At the same time, Bap facilitates the colonization of host tissues and the establishment of persistent infections by *S. aureus* (Cucarella et al., 2004; Valle et al., 2012). The interaction of Bap with Gp96/GRP94/Hsp90 provokes a significant reduction of epithelial cell invasion by interfering with the fibronectin binding protein invasion pathway. While Bap is often considered as a key determinant for biofilm formation (Lasa, 2006), a recent investigation revealed that there is no direct correlation with biofilm formation and that a single gene

or subset of genes cannot be utilized as a biofilm indicator for morphology in *S. aureus* (Tang et al., 2013). Interestingly, Bap is subjected to phase variation and the switch between Bap ON/Bap OFF states might regulate sessile development (Henderson et al., 1999; Tormo et al., 2007). In *Enterococcus faecalis*, Esp (enterococcal surface protein) also leads to a significant increase in biofilm formation (Tendolkar et al., 2004) and contributes to persistence in the host (Sava et al., 2010). In *S. epidermidis* and *Listeria monocytogenes*, however, Bhp (Bap homologue protein) and BapL, respectively, are not clearly involved in biofilm formation (Tormo et al., 2005; Lasa and Penadés, 2006; Jordan et al., 2008; Renier et al., 2011).

As recently reviewed (Chagnot et al., 2012), different MSCRAMM surface proteins have been investigated in monoderm bacteria (Vengadesan and Narayana, 2011). Among the main ECM fibrillar proteins, namely collagen, fibronectin, laminin and elastin, eight fibronectin-binding domains (FBD1 to FBD8) and only one collagen-binding domain (CBD) have been characterized to date and much remains to be learned about specific binding to other ECM components (Chagnot et al., 2012). Very interestingly, domains involved in binding to ECM proteins such as fibronectin and collagen could also bind polystyrene as demonstrated with RspA (*Rhusiopathiae* surface protein A) (Shimoji et al., 2003).

In *B. subtilis*, the S-layer protein BslA (*Bacillus* S-layer protein A; formerly YuaB) is important for pellicle biofilms (Ostrowski et al., 2011; Vlamakis et al., 2013). BslA has amphiphilic properties and forms a hydrophobic layer on the surface of the biofilm which can contribute to colonization of the air-surface interface (Kovacs and Kuipers, 2011; Kobayashi and Iwano, 2012).

Pili: Type 3 (T3P) and amyloid fibers

A little bit more than a decade ago, pili were considered to be absent from monoderm bacteria. The first pathway recognized in pilus biogenesis in monodermata involves polymerization of LPXTG-pilins by different types of transpeptidase sortases, (Shimoji et al., 2003; Ton-That and Schneewind, 2003; Ton-That et al., 2004; Telford et al., 2006; Kang and Baker, 2012). Those pili of Type 3 (T3P) are involved in adhesion, biofilm formation and host colonization (Swierczynski and Ton-That, 2006; Danne and Dramsi, 2012). LipA (light-inducible pilin A) from *Arthrobacter photogonimoides* was proposed as the major pilin of an alternative pilus formation pathway in monoderm bacteria (Yang and Hooper, 1998, 1999; Ton-That and Schneewind, 2004; Swierczynski and Ton-That, 2006).

In *B. subtilis*, pili formation of the amyloid-fiber type has been recently demonstrated to be essential in colony and pellicle biofilms but not for surface-adhered biofilms (Romero et al., 2010; Vlamakis et al., 2013). The CW-protein TapA (TasA anchoring and assembly protein A; formerly YqxM) allows the CW anchoring and assembly of TasA (translocation-dependent antimicrobial spore component A) into long amyloid fibers (Stover and Driks, 1999; Romero et al., 2011), these proteins are not required in biofilms. In *Peptostreptococcus micros*, auto-aggregative pili are formed at least from a CWBD1 protein called FibA (fibril-like structure subunit A) (Kremer et al., 1999).

Cellulosome

Cellulosomes are supramolecular protein complexes found on the cell surface of some cellulolytic anaerobic monoderm bacteria, such as the thermophilic *Clostridium thermocellum* or mesophilic *C. cellulolyticum*, and dedicated to specific adhesion, colonization and degradation of insoluble lignocellulosic substrates (Bayer et al., 2004, 2008; Gilbert, 2007; Fontes and Gilbert, 2010). It is also responsible for regulating the entering carbon metabolic flux in the bacterial cell, which in turn can influence the bacterial colonization of the insoluble lignocellulosic substratum (Guedon et al., 1999, 2000; Payot et al., 1999; Desvaux et al., 2001; Desvaux and Petitdemange, 2001, 2002; Desvaux, 2004). The cellulosome is composed of a non-catalytic protein called scaffoldin, which exhibits domains called cohesin of Type I (Coh I) domains allowing the specific binding of a dockerin domain of Type I (Doc) borne by cellulosomal enzymes. Actually, a single Doc I can simultaneously bind two Coh I and thus allow the linking of two scaffoldins resulting in the formation of polycellulosomes (Carvalho et al., 2003). In *C. thermocellum*, a cellulosome has a size ranging from 2.0 to 6.5 MDa depending on the bacterial strain (Béguin and Lemaire, 1996) and forms polycellulosomal protuberances of up to 100 MDa (Shoham et al., 1999). The cellulosomal enzymes and/or the scaffoldin contain carbohydrate-binding modules (CBMs) involved in the tight binding to a carbohydrate-polymer such as cellulose (Shoseyov et al., 2006; Guillen et al., 2010). The scaffoldin can also exhibit a Doc II domain that can interact with Coh II present in SLHD cell-wall proteins, thus permitting cell-surface display of the cellulosome. In *C. cellulolyticum* for instance, the molecular mechanisms for attachment of the cellulosome to the bacterial cell surface remains unknown (Desvaux, 2005a,b). The architecture of the cellulosome can be even more complex with the assembly of scaffoldins on other scaffoldins like in *Acetivibrio cellulolyticus* where one supramolecular complex can potentially bind 84 different cellulosomal enzymes (Rincon et al., 2003; Xu et al., 2003, 2004). However, the molecular mechanisms permitting the coordinated assembly of the different protein subunits on the bacterial cell-surface remain unclear (Desvaux, 2005b, 2006).

TYPE 4 PILUS (T4P) ASSEMBLED BY THE FPE

The Com (Competence development) pathway involves both the bacterial competence-related DNA transformation transporter and the FPE system (Dubnau, 1997, 1999; Dubnau and Provvedi, 2000; Chen and Dubnau, 2004). In *B. subtilis*, the FPE system is involved in the formation of a pseudopilus from Type 4 prepilins (Chen and Dubnau, 2004; Chen et al., 2006) but do not form a complete Type 4 pilus (T4P). While this was considered as a paradigm for monoderm bacteria, it was recently demonstrated the FPE could form complete Type 4 pilus in *Streptococcus pneumoniae* (Laurenceau et al., 2013). While this T4P was only considered in the context of DNA transformation, its implication in bacterial colonization as reported for T4P in diderm-LPS bacteria is an intriguing possibility that would require further investigations in other species such as *L. monocytogenes* (Rabinovich et al., 2012). GP25 (glycoprotein of 25kDa) from *Ruminococcus albus* was characterized as the major protein subunit of T4P present at the bacterial cell surface (Pegden et al., 1998; Mosoni

and Gaillard-Martinie, 2001; Rakotoarivonina et al., 2002, 2005). Those pili could be involved in cellulose colonization and cellulose degradation. T4P involved in biofilm formation and gliding motility was also uncovered in *Clostridium perfringens* (Varga et al., 2006; Mendez et al., 2008; Varga et al., 2008).

Protein components of the FPE machinery are homologous to some proteins found in T2SS, T4SS, and T4P assembly apparatus in diderm-LPS bacteria and have been collectively called PSTC (Pilus/Secretion/Twitching motility/Competence) (Fussenegger et al., 1997; Dubnau, 1999; Peabody et al., 2003). In *Bifidobacterium breve*, a functional Tad system involved in the formation of T4bP has recently been uncovered (O'connell Motherway et al., 2011). These pili are essential for intestinal colonization of the host. Besides the genus *Actinobacteria*, a Tad system was also identified in *Clostridia*, namely *Cl. acetobutylicum* (Desvaux et al., 2005b). Again, there is much confusion between the pili assembly by FPE, Tad and the misleading called "Type IV-like secretion system" in monoderm bacteria corresponding more adequately to the Tra system. This would necessitate in-depth phylogenetic and functional-genetic analyses to clarify the situation.

FLAGELLUM ASSEMBLED BY THE FEA

In monoderm bacteria, the flagellar subunits are secreted and assembled by the FEA but the involvement and exact contribution of flagella in biofilm formation remain difficult to establish. Taking *L. monocytogenes* as a case study, it is clear that the regulation of flagella expression is quite complex with the intervention of several transcription regulators but basically that flagella are not expressed at temperatures higher than 30°C and bacterial cells are then non-motile (Renier et al., 2011). Regarding listerial colonization, it was shown flagella were not essential for biofilm development but facilitated initial attachment to surfaces (Vatanyoopaisarn et al., 2000) and were even crucial for initial adhesion (Tresse et al., 2006). During infection, though, the flagella would act as mediators of motility rather than adhesins to enhance the adhesion of *L. monocytogenes* to targeted host cells (O'neil and Marquis, 2006). By contrast with early findings, it was further shown the role of the flagella as mediators of motility rather than adhesins was critical for initial surface attachment and subsequent biofilm formation (Lemon et al., 2007). In dynamic culture conditions for sessile development (i.e., using flow cells), the loss of the flagella resulted in the formation of a very dense biofilm (hyperbiofilm) (Todhanakasem and Young, 2008). These conflicting findings could arise from the use of different *L. monocytogenes* strains as some of them can be more or less laboratory-adapted (domesticated) and have different regulations of flagella expression (Grundling et al., 2004). It could also result from the use of different culture conditions, such as the growth media (e.g., complex undefined or chemically defined), temperature (e.g., related to infection or environmental conditions), support (e.g., plasticware, stainless steel or glass) and/or culture systems (e.g., static or dynamic conditions) (Renier et al., 2011).

In *Bacillus cereus*, flagella motility (i) is necessary for biofilm formation at air-liquid interface as it allows bacteria to reach it, (ii) promotes recruitment of planktonic cells within the biofilm by allowing motile bacteria to invade the whole biofilm, and (iii)

allows spreading of the biofilm on surfaces (Houry et al., 2010). In *B. subtilis*, however, motility and biofilm formation are quite clearly coordinately and oppositely controlled (Kearns et al., 2005; Newman et al., 2013; Vlamakis et al., 2013). It was further demonstrated a protein clutch arrested flagellar motility by disengaging motor force-generating elements in cells engaged in sessile development (Blair et al., 2008). Nonetheless, and while the number of motile cells decreases in the course of sessile development, a sub-population of motile cells remains in mature biofilms (Vlamakis et al., 2008) and a biofilm exclusively composed of unflagellated cells is delayed in forming pellicle biofilms (Kobayashi, 2007) and defective in the formation of surface-adhered biofilms (Chagneau and Saier, 2004). Interestingly, motile *Bacillus thuringiensis* can tunnel deep within a biofilm structure (Houry et al., 2012). Those swimming cells create transient pores that increase irrigation within the biofilm. This can either improve biofilm bacterial fitness by increasing nutrient flow in the matrix or facilitating the penetration of toxic substances detrimental for biofilm formation.

NON-CLASSICAL PROTEIN SECRETION AND BACTERIAL COLONIZATION

Analysis of the protein content of the extracellular milieu and/or the cell surface, revealed the presence of unexpected proteins primarily predicted or known to be localized in the cytoplasm, with no N-terminal signal peptide (SP), and no dedicated protein secretion system. While early investigations suggested it could result from bacterial cell lysis, it later appeared some of those proteins were actually secreted by previously unrecognized pathways, e.g., some SecA2-dependent secreted proteins (Desvaux and Hébraud, 2006, 2009; Rigel and Braunstein, 2008; Renier et al., 2013). Non-classical (NC) protein secretion refers to the uncharacterized mechanisms responsible for secretion of those proteins (prior to characterization such apparent secretion could be the result of phenomenon unrelated to secretion *per se*, such as membrane budding, allolysis, phage-mediated lysis, GTA, etc. . .) (Bendtsen et al., 2005; Bendtsen and Wooldridge, 2009). Another unexpected finding was the involvement of some of those proteins in an additional function completely different from the primary one. Indeed, the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is a well-known and key glycolytic enzyme but when present on the bacterial cell surface it exhibits very significant binding activity against plasmin(ogen), fibrinogen, and human cells (C1q component of the classical component pathway) (Bergmann et al., 2004; Jin et al., 2005; Egea et al., 2007; Matta et al., 2010; Tunio et al., 2010; Terrasse et al., 2012). Several of those so-called moonlighting proteins have been identified in bacteria and play a significant role in bacterial colonization, especially in an infection context (Jeffery, 2003, 2009; Huberts and Van Der Klei, 2010; Copley, 2012; Henderson and Martin, 2013). Most enzymes of the EMP (Embden-Meyerhof-Parnas) pathway can moonlight and a role as an adhesin is quite common, e.g., for enolase, aldolase, or phosphoglucose mutase (PGM) (Henderson and Martin, 2011, 2013). However, the molecular mechanisms responsible for their bacterial cell-surface localization remain unknown.

CONCLUSIONS AND PERSPECTIVES

While extremely useful, the secretome concept has not been fully exploited and is still in its infancy. In the context of bacterial colonization, it allows a rational approach for comprehending the secreted proteins directly involved in bacterial adhesion and/or biofilm formation. Considering the protein routing and secretion mechanisms embrace the idea that some of those adhesins and/or adhesion supramolecular complexes share common post-translational and post-translocational maturation mechanisms that may lead to the identification of novel targets for therapeutic strategies. While the majority of investigations are focused on colonization factors in the context of infection, their involvement in the interactions of commensal bacterial, as well as bacteria in an environmental context, are also important areas and should not be overlooked. Bacterial colonization is not the prerogative of pathogens and there is much to gain from the study of colonization of host species by commensal bacteria, as well as environmentally orientated studies. Clearly, there are numerous proteins and supramolecular structures involved in bacterial adhesion and biofilm formation but their functional characterization still demands extensive investigations, e.g., alternative pili in monoderm bacteria or collagen-/laminin-/elastin-binding domains. The understanding of the global and coordinated regulation of the secretome in the course of colonization at the single cell, homogenous population, and multi-species microbial community levels is a major challenge in different ecological niches. Phase variation for instance is a well-known phenomenon especially for factors involved in colonization process but is rarely considered and is poorly understood at a global scale. In the last couple of years, new secretion system such as the T6SS or T9SS have been unexpectedly uncovered and there is much to bet that other protein secretion system(s) are yet to be discovered. Besides, the exact contribution of NC secretion and other protein trafficking mechanisms in bacterial colonization (e.g., allolysis, phage-mediated lysis, membrane budding or GTA) remains to be determined. Regarding the different secretion systems present in a single bacterial cell and numerous systems within microbiota as well as their associated secreted proteins involved in colonization, they promise to be the subject of a lot of research work ahead as well as exciting new developments in the field.

ACKNOWLEDGMENTS

This work was supported by INRA (French National Institute for Agronomical Research) with the MICEL project funded by the inter-department CEPIA/MICA (Science and Process Engineering of Agricultural Products/Microbiology and Food Chain) AIP (“Action Incitative Programmée”) and COST (European Cooperation in Science and Technology) Action 1202 BacFoodNet (A European Network For Mitigating Bacterial Colonization and Persistence On Foods and Food Processing Environments). Mohamed Amine Zorgani is a PhD research fellow granted by the “Ministère de l’Enseignement Supérieur et de la Recherche.” Caroline Chagnot is a PhD Research Fellow granted by the “Région Auvergne—FEDER (Fonds Européen de Développement Régional).” The authors are thankful to Reviewers for their very constructive comments.

REFERENCES

- Abdallah, A. M., Gey Van Pittius, N. C., Champion, P. A., Cox, J., Luirink, J., Vandenbroucke-Grauls, C. M., et al. (2007). Type VII secretion—mycobacteria show the way. *Nat. Rev. Microbiol.* 5, 883–891. doi: 10.1038/nrmicro1773
- Alvarez-Martinez, C. E., and Christie, P. J. (2009). Biological diversity of prokaryotic Type IV secretion systems. *Microbiol. Mol. Biol. Rev.* 73, 775–808. doi: 10.1128/MMBR.00023-09
- An, Y. H., and Friedman, R. J. (1998). Concise review of mechanisms of bacterial adhesion to biomaterial surfaces. *J. Biomed. Mater. Res.* 43, 338–348.
- Antelmann, H., Tjalsma, H., Voigt, B., Ohlmeier, S., Bron, S., Van Dijk, J. M., et al. (2001). A proteomic view on genome-based signal peptide predictions. *Genome Res.* 11, 1484–1502. doi: 10.1101/gr.182801
- Antelmann, H., Van Dijk, J. M., Bron, S., and Hecker, M. (2006). Proteomic survey through secretome of *Bacillus subtilis*. *Methods Biochem. Anal.* 49, 179–208.
- Arenas, J., Nijland, R., Rodriguez, F. J., Bosma, T. N., and Tommasen, J. (2013). Involvement of three meningococcal surface-exposed proteins, the heparin-binding protein Nhba, the α -peptide of IgA protease and the auto-transporter protease NalP, in initiation of biofilm formation. *Mol. Microbiol.* 87, 254–268. doi: 10.1111/mmi.12097
- Ashgar, S. S., Oldfield, N. J., Wooldridge, K. G., Jones, M. A., Irving, G. J., Turner, D. P., et al. (2007). CapA, an autotransporter protein of *Campylobacter jejuni*, mediates association with human epithelial cells and colonization of the chicken gut. *J. Bacteriol.* 189, 1856–1865. doi: 10.1128/JB.01427-06
- Ayers, M., Howell, P. L., and Burrows, L. L. (2010). Architecture of the Type II secretion and Type 4 pilus machineries. *Future Microbiol.* 5, 1203–1218. doi: 10.2217/fmb.10.76
- Bahl, M. I., Hansen, L. H., and Sorensen, S. J. (2007). Impact of conjugal transfer on the stability of IncP-1 plasmid pJKK5 in bacterial populations. *FEMS Microbiol. Lett.* 266, 250–256. doi: 10.1111/j.1574-6968.2006.00536.x
- Barnhart, M. M., and Chapman, M. R. (2006). Curli biogenesis and function. *Annu. Rev. Microbiol.* 60, 131–147. doi: 10.1146/annurev.micro.60.080805.142106
- Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., et al. (2004). The Pfam protein families database. *Nucleic Acids Res.* 32, D138–D141. doi: 10.1093/nar/gkh121
- Bayer, E. A., Belaich, J. P., Shoham, Y., and Lamed, R. (2004). The cellulosome: multienzyme machines for degradation of plant cell wall polysaccharides. *Annu. Rev. Microbiol.* 58, 521–554. doi: 10.1146/annurev.micro.57.030502.091022
- Bayer, E. A., Lamed, R., White, B. A., and Flint, H. J. (2008). From cellulosomes to cellulosomes. *Chem. Rev.* 8, 364–377. doi: 10.1002/tcr.20160
- Bazaka, K., Crawford, R. J., Nazarenko, E. L., and Ivanova, E. P. (2011). Bacterial extracellular polysaccharides. *Adv. Exp. Med. Biol.* 715, 213–226. doi: 10.1007/978-94-007-0940-9_13
- Béguin, P., and Lemaire, M. (1996). The cellulosome: an exocellular, multiprotein complex specialized in cellulose degradation. *Crit. Rev. Biochem. Mol. Biol.* 31, 201–236. doi: 10.3109/10409239609106584
- Bendtsen, J. D., Kiemer, L., Fausboll, A., and Brunak, S. (2005). Non-classical protein secretion in bacteria. *BMC Microbiol.* 5:58. doi: 10.1186/1471-2180-5-58
- Bendtsen, J. D., and Wooldridge, K. G. (2009). “Chapter 10: non-classical secretion,” in *Bacterial Secreted Proteins: Secretory Mechanisms and Role in Pathogenesis*, ed K. Wooldridge (Norwich: Caister Academic Press), 226–235.
- Berger, C. N., Shaw, R. K., Brown, D. J., Mather, H., Clare, S., Dougan, G., et al. (2009a). Interaction of *Salmonella enterica* with basil and other salad leaves. *ISME J.* 3, 261–265. doi: 10.1038/ismej.2008.95
- Berger, C. N., Shaw, R. K., Ruiz-Perez, F., Nataro, J. P., Henderson, I. R., Pallen, M. J., et al. (2009b). Interaction of enteroaggregative *Escherichia coli* with salad leaves. *Environ. Microbiol. Rep.* 1, 234–239. doi: 10.1111/j.1758-2229.2009.00037.x
- Berger, C. N., Sodha, S. V., Shaw, R. K., Griffin, P. M., Pink, D., Hand, P., et al. (2010). Fresh fruit and vegetables as vehicles for the transmission of human pathogens. *Environ. Microbiol.* 12, 2385–2397. doi: 10.1111/j.1462-2920.2010.02297.x
- Bergmann, S., Rohde, M., and Hammerschmidt, S. (2004). Glyceraldehyde-3-phosphate dehydrogenase of *Streptococcus pneumoniae* is a surface-displayed plasminogen-binding protein. *Infect. Immun.* 72, 2416–2419. doi: 10.1128/IAI.72.4.2416-2419.2004
- Bernard, C. S., Bordi, C., Termine, E., Filloux, A., and De Bentzmann, S. (2009). Organization and PprB-dependent control of the *Pseudomonas aeruginosa* tad Locus, involved in Flp pilus biology. *J. Bacteriol.* 191, 1961–1973. doi: 10.1128/JB.01330-08
- Biswas, R., Agarwal, R. K., Bhilegaonkar, K. N., Kumar, A., Nambiar, P., Rawat, S., et al. (2011). Cloning and sequencing of biofilm-associated protein (*bapA*) gene and its occurrence in different serotypes of *Salmonella*. *Lett. Appl. Microbiol.* 52, 138–143. doi: 10.1111/j.1472-765X.2010.02975.x
- Blair, K. M., Turner, L., Winkelman, J. T., Berg, H. C., and Kearns, D. B. (2008). A molecular clutch disables flagella in the *Bacillus subtilis* biofilm. *Science* 320, 1636–1638. doi: 10.1126/science.1157877
- Blanco, L. P., Evans, M. L., Smith, D. R., Badtke, M. P., and Chapman, M. R. (2012). Diversity, biogenesis and function of microbial amyloids. *Trends Microbiol.* 20, 66–73. doi: 10.1016/j.tim.2011.11.005
- Brady, M. J., Radhakrishnan, P., Liu, H., Magoun, L., Murphy, K. C., Mukherjee, J., et al. (2011). Enhanced actin pedestal formation by enterohemorrhagic *Escherichia coli* O157:H7 Adapted to the Mammalian Host. *Front. Microbiol.* 2:226. doi: 10.3389/fmicb.2011.00226
- Brennan, P. J., and Nikaido, H. (1995). The envelope of mycobacteria. *Annu. Rev. Biochem.* 64, 29–63. doi: 10.1146/annurev.bi.64.070195.000333
- Bucior, I., Pielage, J. F., and Engel, J. N. (2012). *Pseudomonas aeruginosa* pili and flagella mediate distinct binding and signaling events at the apical and basolateral surface of airway epithelium. *PLoS Pathog.* 8:e1002616. doi: 10.1371/journal.ppat.1002616
- Buist, G., Ridder, A. N. J. A., Kok, J., and Kuipers, O. P. (2006). Different subcellular locations of secretome components of Gram-positive bacteria. *Microbiology* 152, 2867–2874. doi: 10.1099/mic.0.29113-0
- Bullard, B., Lipski, S., and Lafontaine, E. R. (2007). Regions important for the adhesin activity of *Moraxella catarrhalis* Hag. *BMC Microbiol.* 7:65. doi: 10.1186/1471-2180-7-65
- Buncic, S., Nychas, G. J., Lee, M. R., Koutsoumanis, K., Hebraud, M., Desvaux, M., et al. (2013). Microbial pathogen control in the beef chain: recent research advances. *Meat Sci.* doi: 10.1016/j.meatsci.2013.04.040. [Epub ahead of print]
- Burmolle, M., Bahl, M. I., Jensen, L. B., Sorensen, S. J., and Hansen, L. H. (2008). Type 3 fimbriae, encoded by the conjugative plasmid pOLA52, enhance biofilm formation and transfer frequencies in *Enterobacteriaceae* strains. *Microbiology* 154, 187–195. doi: 10.1099/mic.0.2007/010454-0
- Burrows, L. L. (2012). *Pseudomonas aeruginosa* twitching motility: Type 4 pili in action. *Annu. Rev. Microbiol.* 66, 493–520. doi: 10.1146/annurev-micro-092611-150055
- Busch, A., and Waksman, G. (2012). Chaperone-usher pathways: diversity and pilus assembly mechanism. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 1112–1122. doi: 10.1098/rstb.2011.0206
- Büttner, D. (2012). Protein export according to schedule: architecture, assembly, and regulation of type iii secretion systems from plant- and animal-pathogenic bacteria. *Microbiol. Mol. Biol. Rev.* 76, 262–310. doi: 10.1128/MMBR.05017-11
- Campellone, K. G., and Leong, J. M. (2003). Tails of two Tirs: actin pedestal formation by enteropathogenic *E. coli* and enterohemorrhagic *E. coli* O157:H7. *Curr. Opin. Microbiol.* 6, 82–90. doi: 10.1016/S1369-5274(03)00005-5
- Candela, T., Dumetz, F., Tosi-Couture, E., Mock, M., Goossens, P. L., and Fouet, A. (2012). Cell-wall preparation containing poly-g-D-glutamate covalently linked to peptidoglycan, a straightforward extractable molecule, protects mice against experimental anthrax infection. *Vaccine* 31, 171–175. doi: 10.1016/j.vaccine.2012.10.071
- Candela, T., and Fouet, A. (2006). Poly-gamma-glutamate in bacteria. *Mol. Microbiol.* 60, 1091–1098. doi: 10.1111/j.1365-2958.2006.05179.x
- Carvalho, A. L., Dias, F. M. V., Prates, J. A. M., Nagy, T., Gilbert, H. J., Davies, G. J., et al. (2003). Cellulosome assembly revealed by the crystal structure of the cohesin-dockerin complex. *Proc. Natl. Acad. Sci. U.S.A.* 100, 13809–13814. doi: 10.1073/pnas.1936124100
- Cascales, E., and Cambillau, C. (2012). Structural biology of Type VI secretion systems. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 1102–1111. doi: 10.1098/rstb.2011.0209
- Chagneau, C., and Saier, M. H. Jr. (2004). Biofilm-defective

- mutants of *Bacillus subtilis*. *J. Mol. Microbiol. Biotechnol.* 8, 177–188. doi: 10.1159/000085790
- Chagnot, C., Agus, A., Renier, S., Peyrin, F., Talon, R., Astruc, T., et al. (2013). *In vitro* colonization of the muscle extracellular matrix components by *Escherichia coli* O157:H7: the influence of growth medium, temperature and pH on initial adhesion and induction of biofilm formation by collagens I and III. *PLoS ONE* 8:e59386. doi: 10.1371/journal.pone.0059386
- Chagnot, C., Listrat, A., Astruc, T., and Desvaux, M. (2012). Bacterial adhesion to animal tissues: protein determinants for recognition of extracellular matrix components. *Cell Microbiol.* 14, 1687–1696. doi: 10.1111/cmi.12002
- Charbonneau, M. E., Janvore, J., and Mourez, M. (2009). Autoprocessing of the *Escherichia coli* AIDA-I autotransporter: a new mechanism involving acidic residues in the junction region. *J. Biol. Chem.* 284, 17340–17351. doi: 10.1074/jbc.M109.010108
- Chen, I., and Dubnau, D. (2004). DNA uptake during bacterial transformation. *Nat. Rev. Microbiol.* 2, 241–249. doi: 10.1038/nrmicro844
- Chen, I., Provvedi, R., and Dubnau, D. (2006). A macromolecular complex formed by a pilin-like protein in competent *Bacillus subtilis*. *J. Biol. Chem.* 281, 21720–21727. doi: 10.1074/jbc.M604071200
- Christie, P. J., Atmakuri, K., Krishnamoorthy, V., Jakubowski, S., and Cascales, E. (2005). Biogenesis, architecture, and function of bacterial Type IV secretion systems. *Ann. Rev. Microbiol.* 59, 451–485. doi: 10.1146/annurev.micro.58.030603.123630
- Christie, P. J., and Vogel, J. P. (2000). Bacterial Type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. *Trends Microbiol.* 8, 354–360. doi: 10.1016/S0966-842X(00)01792-3
- Cianciotto, N. P. (2005). Type II secretion: a protein secretion system for all seasons. *Trends Microbiol.* 13, 581–588. doi: 10.1016/j.tim.2005.09.005
- Comanducci, M., Bambini, S., Brunelli, B., Adu-Bobie, J., Arico, B., Capocchi, B., et al. (2002). NadA, a novel vaccine candidate of *Neisseria meningitidis*. *J. Exp. Med.* 195, 1445–1454. doi: 10.1084/jem.20020407
- Connors, R., Hill, D. J., Borodina, E., Agnew, C., Daniell, S. J., Burton, N. M., et al. (2008). The *Moraxella* adhesin UspA1 binds to its human CEACAM1 receptor by a deformable trimeric coiled-coil. *EMBO J.* 27, 1779–1789. doi: 10.1038/emboj.2008.101
- Conrad, J. C. (2012). Physics of bacterial near-surface motility using flagella and Type 4 pili: implications for biofilm formation. *Res. Microbiol.* 163, 619–629. doi: 10.1016/j.resmic.2012.10.016
- Converse, S. E., and Cox, J. S. (2005). A protein secretion pathway critical for *Mycobacterium tuberculosis* virulence is conserved and functional in *Mycobacterium smegmatis*. *J. Bacteriol.* 187, 1238–1245. doi: 10.1128/JB.187.4.1238-1245.2005
- Copley, S. D. (2012). Moonlighting is mainstream: paradigm adjustment required. *Bioessays* 34, 578–588. doi: 10.1002/bies.201100191
- Cornelis, G. R. (2006). The Type III secretion injectisome. *Nat. Rev. Microbiol.* 4, 811–825. doi: 10.1038/nrmicro1526
- Cornelis, G. R. (2010). The Type III secretion injectisome, a complex nanomachine for intracellular ‘toxin’ delivery. *Biol. Chem.* 391, 745–751. doi: 10.1515/bc.2010.079
- Correa, V. R., Majerczak, D. R., Ammar, E. D., Merighi, M., Pratt, R. C., Hogenhout, S. A., et al. (2012). The bacterium *Pantoea stewartii* uses two different Type III secretion systems to colonize its plant host and insect vector. *Appl. Environ. Microbiol.* 78, 6327–6336. doi: 10.1128/AEM.00892-12
- Cotter, S. E., Surana, N. K., and St Geme, J. W. 3rd. (2005). Trimeric autotransporters: a distinct subfamily of autotransporter proteins. *Trends Microbiol.* 13, 199–205. doi: 10.1016/j.tim.2005.03.004
- Craig, L., and Li, J. (2008). Type 4 pili: paradoxes in form and function. *Curr. Opin. Struct. Biol.* 18, 267–277. doi: 10.1016/j.sbi.2007.12.009
- Craig, L., Pique, M. E., and Tainer, J. A. (2004). Type 4 pilus structure and bacterial pathogenicity. *Nat. Rev. Microbiol.* 2, 363–378. doi: 10.1038/nrmicro885
- Cucarella, C., Solano, C., Valle, J., Amorena, B., Lasa, I., and Penades, J. R. (2001). Bap, a *Staphylococcus aureus* surface protein involved in biofilm formation. *J. Bacteriol.* 183, 2888–2896. doi: 10.1128/JB.183.9.2888-2896.2001
- Cucarella, C., Tormo, M. A., Knecht, E., Amorena, B., Lasa, I., Foster, T. J., et al. (2002). Expression of the biofilm-associated protein interferes with host protein receptors of *Staphylococcus aureus* and alters the infective process. *Infect. Immun.* 70, 3180–3186. doi: 10.1128/IAI.70.6.3180-3186.2002
- Cucarella, C., Tormo, M. A., Ubeda, C., Trottonda, M. P., Monzon, M., Peris, C., et al. (2004). Role of biofilm-associated protein bap in the pathogenesis of bovine *Staphylococcus aureus*. *Infect. Immun.* 72, 2177–2185. doi: 10.1128/IAI.72.4.2177-2185.2004
- Cullen, P. A., Xu, X., Matsunaga, J., Sanchez, Y., Ko, A. I., Haake, D. A., et al. (2005). Surfaceome of *Leptospira* spp. *Infect. Immun.* 73, 4853–4863. doi: 10.1128/IAI.73.8.4853-4863.2005
- Danne, C., and Dramsi, S. (2012). Pili of Gram-positive bacteria: roles in host colonization. *Res. Microbiol.* 163, 645–658. doi: 10.1016/j.resmic.2012.10.012
- Darsonval, A., Darrasse, A., Meyer, D., Demarty, M., Durand, K., Bureau, C., et al. (2008). The Type III secretion system of *Xanthomonas fuscans* subsp. *fuscans* is involved in the phyllosphere colonization process and in transmission to seeds of susceptible beans. *Appl. Environ. Microbiol.* 74, 2669–2678. doi: 10.1128/AEM.02906-07
- De Luna, M. G., Scott-Tucker, A., Desvaux, M., Ferguson, P., Morin, N. P., Dudley, E. G., et al. (2008). The *Escherichia coli* biofilm-promoting protein Antigen 43 does not contribute to intestinal colonization. *FEMS Microbiol. Lett.* 284, 237–246. doi: 10.1111/j.1574-6968.2008.01207.x
- Delepelaire, P. (2004). Type I secretion in Gram-negative bacteria. *Biochim. Biophys. Acta Mol. Cell Res.* 1694, 149–161. doi: 10.1016/j.bbamcr.2004.05.001
- Desvaux, M. (2004). Mapping of carbon flow distribution in the central metabolic pathways of *Clostridium cellulolyticum*: direct comparison of bacterial metabolism with a soluble versus an insoluble carbon source. *J. Microbiol. Biotechnol.* 14, 1200–1210.
- Desvaux, M. (2005a). The cellulosome of *Clostridium cellulolyticum*. *Enz. Microb. Technol.* 37, 373–385. doi: 10.1016/j.enzmict.2004.04.025
- Desvaux, M. (2005b). *Clostridium cellulolyticum*: model organism of mesophilic cellulolytic clostridia. *FEMS Microbiol. Rev.* 29, 741–764. doi: 10.1016/j.femsre.2004.11.003
- Desvaux, M. (2006). Unravelling carbon metabolism in anaerobic cellulolytic bacteria. *Biotechnol. Prog.* 22, 1229–1238. doi: 10.1002/bp060016e
- Desvaux, M. (2012). Contribution of holins to protein trafficking: secretion, leakage or lysis? *Trends Microbiol.* 20, 259–261. doi: 10.1016/j.tim.2012.03.008
- Desvaux, M., Dumas, E., Chafsey, I., and Hébraud, M. (2006a). Protein cell surface display in Gram-positive bacteria: from single protein to macromolecular protein structure. *FEMS Microbiol. Lett.* 256, 1–15. doi: 10.1111/j.1574-6968.2006.00122.x
- Desvaux, M., Hébraud, M., Henderson, I. R., and Pallen, M. J. (2006b). Type III secretion: what’s in a name? *Trends Microbiol.* 14, 157–160. doi: 10.1016/j.tim.2006.02.009
- Desvaux, M., Guedon, E., and Petitdemange, H. (2001). Kinetics and metabolism of cellulose degradation at high substrate concentrations in steady-state continuous cultures of *Clostridium cellulolyticum* on a chemically defined medium. *Appl. Environ. Microbiol.* 67, 3837–3845. doi: 10.1128/AEM.67.9.3837-3845.2001
- Desvaux, M., and Hébraud, M. (2006). The protein secretion systems in *Listeria*: inside out bacterial virulence. *FEMS Microbiol. Rev.* 30, 774–805. doi: 10.1111/j.1574-6976.2006.00035.x
- Desvaux, M., and Hébraud, M. (2009). “Chapter 14: *Listeria* monocytogenes,” in *Bacterial Secreted Proteins: Secretory Mechanisms and Role in Pathogenesis*, ed K. Wooldridge (Norwich: Caister Academic Press), 313–345.
- Desvaux, M., Hébraud, M., Talon, R., and Henderson, I. R. (2009a). Outer membrane translocation: numerical protein secretion nomenclature in question in mycobacteria. *Trends Microbiol.* 17, 338–340. doi: 10.1016/j.tim.2009.05.008
- Desvaux, M., Hébraud, M., Talon, R., and Henderson, I. R. (2009b). Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue. *Trends Microbiol.* 17, 139–145. doi: 10.1016/j.tim.2009.01.004
- Desvaux, M., Khan, A., Beatson, S., Scott-Tucker, A., and Henderson, I. R. (2005a). Protein secretion systems in *Fusobacterium nucleatum*: genomic identification of Type 4 piliation and complete Type V pathways brings new insight in mechanisms of pathogenesis. *Biochim. Biophys. Acta Biomembr.* 1713, 92–112. doi: 10.1016/j.bbamem.2005.05.002
- Desvaux, M., Khan, A., Scott-Tucker, A., Chaudhuri, R. R., Pallen, M. J., and Henderson, I. R. (2005b). Genomic analysis of the protein secretion systems in *Clostridium acetobutylicum*

- ATCC824. *Biochim. Biophys. Acta Mol. Cell Res.* 1745, 223–253. doi: 10.1016/j.bbamcr.2005.04.006
- Desvaux, M., Parham, N. J., and Henderson, I. R. (2003). Type V secretion system in Gram-negative bacteria *Biofutur* 237, 34–37.
- Desvaux, M., Parham, N. J., and Henderson, I. R. (2004a). Type V protein secretion: simplicity gone away? *Curr. Issues Mol. Biol.* 6, 111–124.
- Desvaux, M., Parham, N. J., Scott-Tucker, A., and Henderson, I. R. (2004b). The general secretory pathway: a general misnomer? *Trends Microbiol.* 12, 306–309. doi: 10.1016/j.tim.2004.05.002
- Desvaux, M., and Petitdemange, H. (2001). Flux analysis of the metabolism of *Clostridium cellulolyticum* grown in cellulose-fed continuous culture on a chemically defined medium under ammonium-limited conditions. *Appl. Environ. Microbiol.* 67, 3846–3851. doi: 10.1128/AEM.67.9.3846-3851.2001
- Desvaux, M., and Petitdemange, H. (2002). Sporulation of *Clostridium cellulolyticum* while grown in cellulose-batch and cellulose-fed continuous cultures on a mineral-salt based medium. *Microb. Ecol.* 43, 271–279. doi: 10.1007/s00248-001-0043-7
- Devinney, R., Gauthier, A., Abe, A., and Finlay, B. B. (1999). Enteropathogenic *Escherichia coli*: a pathogen that inserts its own receptor into host cells. *Cell Mol. Life Sci.* 55, 961–976. doi: 10.1007/PL00013202
- Digiuseppe Champion, P. A., and Cox, J. S. (2007). Protein secretion systems in *Mycobacteria*. *Cell Microbiol.* 9, 1376–1384. doi: 10.1111/j.1462-5822.2007.00943.x
- Dinh, T., Paulsen, I. T., and Saier, M. H. Jr. (1994). A family of extracytoplasmic proteins that allow transport of large molecules across the outer membranes of Gram-negative bacteria. *J. Bacteriol.* 176, 3825–3831.
- Dorsey, C. W., Laarakker, M. C., Humphries, A. D., Weening, E. H., and Bauml, A. J. (2005). *Salmonella enterica* serotype Typhimurium MisL is an intestinal colonization factor that binds fibronectin. *Mol. Microbiol.* 57, 196–211. doi: 10.1111/j.1365-2958.2005.04666.x
- Douzi, B., Filloux, A., and Voulhoux, R. (2012). On the path to uncover the bacterial Type II secretion system. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 1059–1072. doi: 10.1098/rstb.2011.0204
- Dreisbach, A., Van Dijk, J. M., and Buist, G. (2011). The cell surface proteome of *Staphylococcus aureus*. *Proteomics* 11, 3154–3168. doi: 10.1002/pmic.201000823
- Dubnau, D. (1997). Binding and transport of transforming DNA by *Bacillus subtilis*: the role of type-IV pilin-like proteins—a review. *Gene* 192, 191–198. doi: 10.1016/S0378-1119(96)00804-9
- Dubnau, D. (1999). DNA uptake in bacteria. *Annu. Rev. Microbiol.* 53, 217–244. doi: 10.1146/annurev.micro.53.1.217
- Dubnau, D., and Provvedi, R. (2000). Internalizing DNA. *Res. Microbiol.* 151, 475–480. doi: 10.1016/S0923-2508(00)00166-2
- Dueholm, M. S., Albertsen, M., Otzen, D., and Nielsen, P. H. (2012). Curli functional amyloid systems are phylogenetically widespread and display large diversity in operon and protein structure. *PLoS ONE* 7:e51274. doi: 10.1371/journal.pone.0051274
- Durand, E., Alphonse, S., Brochier-Armanet, C., Ball, G., Douzi, B., Filloux, A., et al. (2011). The assembly mode of the pseudopilus: a hallmark to distinguish a novel secretion system subtype. *J. Biol. Chem.* 286, 24407–24416. doi: 10.1074/jbc.M111.234278
- Dziva, F., Mahajan, A., Cameron, P., Currie, C., McKendrick, I. J., Wallis, T. S., et al. (2007). EspP, a Type V-secreted serine protease of enterohaemorrhagic *Escherichia coli* O157:H7, influences intestinal colonization of calves and adherence to bovine primary intestinal epithelial cells. *FEMS Microbiol. Lett.* 271, 258–264. doi: 10.1111/j.1574-6968.2007.00724.x
- Easton, D. M., Totsika, M., Allsopp, L. P., Phan, M. D., Idris, A., Worpel, D. J., et al. (2011). Characterization of EhaJ, a new autotransporter protein from enterohaemorrhagic *Escherichia coli*. *Front. Microbiol.* 2:120. doi: 10.3389/fmicb.2011.00120
- Economou, A. (2002). Bacterial secretome: the assembly manual and operating instructions. *Mol. Membr. Biol.* 19, 159–169. doi: 10.1080/09687680210152609
- Economou, A., Christie, P. J., Fernandez, R. C., Palmer, T., Plano, G. V., and Pugsley, A. P. (2006). Secretion by numbers: protein traffic in prokaryotes. *Mol. Microbiol.* 62, 308–319. doi: 10.1111/j.1365-2958.2006.05377.x
- Egea, L., Aguilera, L., Gimenez, R., Sorolla, M. A., Aguilar, J., Badia, J., et al. (2007). Role of secreted glyceraldehyde-3-phosphate dehydrogenase in the infection mechanism of enterohaemorrhagic and enteropathogenic *Escherichia coli*: interaction of the extracellular enzyme with human plasminogen and fibrinogen. *Int. J. Biochem. Cell Biol.* 39, 1190–1203. doi: 10.1016/j.biocel.2007.03.008
- Eitel, J., and Dersch, P. (2002). The YdaA protein of *Yersinia pseudotuberculosis* mediates high-efficiency uptake into human cells under environmental conditions in which invasion is repressed. *Infect. Immun.* 70, 4880–4891. doi: 10.1128/IAI.70.9.4880-4891.2002
- El Tahir, Y., and Skurnik, M. (2001). YdaA, the multifaceted *Yersinia* adhesin. *Int. J. Med. Microbiol.* 291, 209–218. doi: 10.1078/1438-4221-00119
- Enos-Berlage, J. L., Guvenier, Z. T., Keenan, C. E., and McCarter, L. L. (2005). Genetic determinants of biofilm development of opaque and translucent *Vibrio parahaemolyticus*. *Mol. Microbiol.* 55, 1160–1182. doi: 10.1111/j.1365-2958.2004.04453.x
- Epstein, E. A., and Chapman, M. R. (2008). Polymerizing the fibre between bacteria and host cells: the biogenesis of functional amyloid fibres. *Cell Microbiol.* 10, 1413–1420. doi: 10.1111/j.1462-5822.2008.01148.x
- Espinosa-Urgel, M., Salido, A., and Ramos, J. L. (2000). Genetic analysis of functions involved in adhesion of *Pseudomonas putida* to seeds. *J. Bacteriol.* 182, 2363–2369. doi: 10.1128/JB.182.9.2363-2369.2000
- Ewanowich, C. A., Melton, A. R., Weiss, A. A., Sherburne, R. K., and Peppler, M. S. (1989). Invasion of HeLa 229 cells by virulent *Bordetella pertussis*. *Infect. Immun.* 57, 2698–2704.
- Fabretti, F., Theilacker, C., Baldassarri, L., Kaczynski, Z., Kropec, A., Holst, O., et al. (2006). Alanine esters of enterococcal lipoteichoic acid play a role in biofilm formation and resistance to antimicrobial peptides. *Infect. Immun.* 74, 4164–4171. doi: 10.1128/IAI.00111-06
- Feltcher, M. E., and Braunstein, M. (2012). Emerging themes in SecA2-mediated protein export. *Nat. Rev. Microbiol.* 10, 779–789. doi: 10.1038/nrmicro2874
- Feltcher, M. E., Sullivan, J. T., and Braunstein, M. (2010). Protein export systems of *Mycobacterium tuberculosis*: novel targets for drug development? *Future Microbiol.* 5, 1581–1597. doi: 10.2217/fmb.10.112
- Fernandez, R. C., and Weiss, A. A. (1994). Cloning and sequencing of a *Bordetella pertussis* serum resistance locus. *Infect. Immun.* 62, 4727–4738.
- Filloux, A. (2004). The underlying mechanisms of Type II protein secretion. *Biochim. Biophys. Acta Mol. Cell Res.* 1694, 163–179. doi: 10.1016/j.bbamcr.2004.05.003
- Fine, D. H., Velliyagounder, K., Furgang, D., and Kaplan, J. B. (2005). The *Actinobacillus actinomycescomitans* autotransporter adhesin Aae exhibits specificity for buccal epithelial cells from humans and old world primates. *Infect. Immun.* 73, 1947–1953. doi: 10.1128/IAI.73.4.1947-1953.2005
- Fink, D. L., Buscher, A. Z., Green, B., Fernsten, P., and St Geme, J. W. 3rd. (2003). The *Haemophilus influenzae* Hap autotransporter mediates microcolony formation and adherence to epithelial cells and extracellular matrix via binding regions in the C-terminal end of the passenger domain. *Cell Microbiol.* 5, 175–186. doi: 10.1046/j.1462-5822.2003.00266.x
- Fink, D. L., Green, B. A., and St Geme, J. W. 3rd. (2002). The *Haemophilus influenzae* Hap autotransporter binds to fibronectin, laminin, and collagen IV. *Infect. Immun.* 70, 4902–4907. doi: 10.1128/IAI.70.9.4902-4907.2002
- Flemming, H. C., and Wingender, J. (2010). The biofilm matrix. *Nat. Rev. Microbiol.* 8, 623–633.
- Fontes, C. M., and Gilbert, H. J. (2010). Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates. *Annu. Rev. Biochem.* 79, 655–681. doi: 10.1146/annurev-biochem-091208-085603
- Frankel, G., and Phillips, A. D. (2008). Attaching effacing *Escherichia coli* and paradigms of Tir-triggered actin polymerization: getting off the pedestal. *Cell Microbiol.* 10, 549–556. doi: 10.1111/j.1462-5822.2007.01103.x
- Freudl, R. (2013). Leaving home ain't easy: protein export systems in Gram-positive bacteria. *Res. Microbiol.* 164, 664–674. doi: 10.1016/j.resmic.2013.03.014
- Fronzes, R., Remaut, H., and Waksman, G. (2008). Architectures and biogenesis of non-flagellar protein appendages in Gram-negative bacteria. *EMBO J.* 27, 2271–2280. doi: 10.1038/emboj.2008.155
- Fulcher, R. A., Cole, L. E., Janowicz, D. M., Toffer, K. L., Fortney, K. R., Katz, B. P., et al. (2006). Expression of *Haemophilus ducreyi* collagen binding outer membrane protein NcaA is required for virulence in swine and human

- challenge models of chancroid. *Infect. Immun.* 74, 2651–2658. doi: 10.1128/IAI.74.5.2651-2658.2006
- Fussenegger, M., Rudel, T., Barten, R., Ryll, R., and Meyer, T. F. (1997). Transformation competence and type-4 pilus biogenesis in *Neisseria gonorrhoeae*—a review. *Gene* 192, 125–134.
- Garmendia, J., Frankel, G., and Crepin, V. F. (2005). Enteropathogenic and enterohemorrhagic *Escherichia coli* infections: translocation, translocation, translocation. *Infect. Immun.* 73, 2573–2585. doi: 10.1128/IAI.73.5.2573-2585.2005
- Gerlach, R. G., and Hensel, M. (2007). Protein secretion systems and adhesins: the molecular armory of Gram-negative pathogens. *Int. J. Med. Microbiol.* 297, 401–415. doi: 10.1016/j.ijmm.2007.03.017
- Gerlach, R. G., Jackel, D., Stecher, B., Wagner, C., Lupas, A., Hardt, W. D., et al. (2007). *Salmonella* Pathogenicity Island 4 encodes a giant non-fimbrial adhesin and the cognate Type 1 secretion system. *Cell Microbiol.* 9, 1834–1850. doi: 10.1111/j.1462-5822.2007.00919.x
- Ghigo, J. M. (2001). Natural conjugative plasmids induce bacterial biofilm development. *Nature* 412, 442–445. doi: 10.1038/35086581
- Ghosh, P. (2004). Process of protein transport by the Type III secretion system. *Microbiol. Mol. Biol. Rev.* 68, 771–795. doi: 10.1128/MMBR.68.4.771-795.2004
- Gilbert, H. J. (2007). Cellulosomes: microbial nanomachines that display plasticity in quaternary structure. *Mol. Microbiol.* 63, 1568–1576. doi: 10.1111/j.1365-2958.2007.05640.x
- Giltner, C. L., Nguyen, Y., and Burrows, L. L. (2012). Type 4 pilin proteins: versatile molecular modules. *Microbiol. Mol. Biol. Rev.* 76, 740–772. doi: 10.1128/MMBR.00035-12
- Giltner, C. L., Van Schaik, E. J., Audette, G. F., Kao, D., Hodges, R. S., Hassett, D. J., et al. (2006). The *Pseudomonas aeruginosa* Type 4 pilin receptor binding domain functions as an adhesin for both biotic and abiotic surfaces. *Mol. Microbiol.* 59, 1083–1096. doi: 10.1111/j.1365-2958.2005.05002.x
- Girard, V., Cote, J. P., Charbonneau, M. E., Campos, M., Berthiaume, F., Hancock, M. A., et al. (2010). Conformation change in a self-recognizing autotransporter modulates bacterial cell-cell interaction. *J. Biol. Chem.* 285, 10616–10626. doi: 10.1074/jbc.M109.069070
- Girard, V., and Mourez, M. (2006). Adhesion mediated by autotransporters of Gram-negative bacteria: structural and functional features. *Res. Microbiol.* 157, 407–416. doi: 10.1016/j.resmic.2006.02.001
- Glew, M. D., Veith, P. D., Peng, B., Chen, Y. Y., Gorasia, D. G., Yang, Q., et al. (2012). PG0026 is the C-terminal signal peptidase of a novel secretion system of *Porphyromonas gingivalis*. *J. Biol. Chem.* 287, 24605–24617. doi: 10.1074/jbc.M112.369223
- Golberg, D., Kroupitski, Y., Belausov, E., Pinto, R., and Sela, S. (2011). *Salmonella* Typhimurium internalization is variable in leafy vegetables and fresh herbs. *Int. J. Food Microbiol.* 145, 250–257. doi: 10.1016/j.ijfoodmicro.2010.12.031
- Gottenbos, B., Van Der Mei, H. C., and Busscher, H. J. (2000). Initial adhesion and surface growth of *Staphylococcus epidermidis* and *Pseudomonas aeruginosa* on biomedical polymers. *J. Biomed. Mater. Res.* 50, 208–214.
- Gottig, N., Garavaglia, B. S., Garofalo, C. G., Orellano, E. G., and Ottado, J. (2009). A filamentous hemagglutinin-like protein of *Xanthomonas axonopodis* pv. *citri*, the phytopathogen responsible for citrus canker, is involved in bacterial virulence. *PLoS ONE* 4:e4358. doi: 10.1371/journal.pone.0004358
- Grant, C. C., Konkel, M. E., Cieplak, W. Jr., and Tompkins, L. S. (1993). Role of flagella in adherence, internalization, and translocation of *Campylobacter jejuni* in nonpolarized and polarized epithelial cell cultures. *Infect. Immun.* 61, 1764–1771.
- Griessl, M. H., Schmid, B., Kassler, K., Braunsman, C., Ritter, R., Barlag, B., et al. (2013). Structural insight into the giant Ca^{2+} -binding adhesin SiiE: implications for the adhesion of *Salmonella enterica* to polarized epithelial cells. *Structure* 21, 741–752. doi: 10.1016/j.str.2013.02.020
- Gross, M., Cramton, S. E., Gotz, F., and Peschel, A. (2001). Key role of teichoic acid net charge in *Staphylococcus aureus* colonization of artificial surfaces. *Infect. Immun.* 69, 3423–3426. doi: 10.1128/IAI.69.5.3423-3426.2001
- Grundling, A., Burrack, L. S., Bouwer, H. G., and Higgins, D. E. (2004). *Listeria monocytogenes* regulates flagellar motility gene expression through MogR, a transcriptional repressor required for virulence. *Proc. Natl. Acad. Sci. U.S.A.* 101, 12318–12323. doi: 10.1073/pnas.0404924101
- Guedon, E., Desvaux, M., Payot, S., and Petitdemange, H. (1999). Growth inhibition of *Clostridium cellulolyticum* by an inefficiently regulated carbon flow. *Microbiology* 145, 1831–1838. doi: 10.1099/13500872-145-8-1831
- Guedon, E., Payot, S., Desvaux, M., and Petitdemange, H. (2000). Relationships between cellobiose catabolism, enzyme levels, and metabolic intermediates in *Clostridium cellulolyticum* grown in a synthetic medium. *Biotechnol. Bioeng.* 67, 327–335.
- Guillen, D., Sanchez, S., and Rodriguez-Sanoja, R. (2010). Carbohydrate-binding domains: multiplicity of biological roles. *Appl. Microbiol. Biotechnol.* 85, 1241–1249. doi: 10.1007/s00253-009-2331-y
- Gupta, R. S. (2011). Origin of diderm (Gram-negative) bacteria: antibiotic selection pressure rather than endosymbiosis likely led to the evolution of bacterial cells with two membranes. *Antonie Van Leeuwenhoek* 100, 171–182. doi: 10.1007/s10482-011-9616-8
- Hahn, H. P. (1997). The Type 4 pilus is the major virulence-associated adhesin of *Pseudomonas aeruginosa*. *Gene* 192, 99–108. doi: 10.1016/S0378-1119(97)00116-9
- Hammar, M., Bian, Z., and Normark, S. (1996). Nucleator-dependent intercellular assembly of adhesive curli organelles in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 93, 6562–6566. doi: 10.1073/pnas.93.13.6562
- Hammer, N. D., McGuffee, B. A., Zhou, Y., Badtke, M. P., Reinke, A. A., Brannstrom, K., Gestwicki, J. E., Olofsson, A., Almqvist, F., and Chapman, M. R. (2012). The C-terminal repeating units of CsgB direct bacterial functional amyloid nucleation. *J. Mol. Biol.* 422, 376–389. doi: 10.1016/j.jmb.2012.05.043
- Harris, R. L., and Silverman, P. M. (2004). Tra proteins characteristic of F-like Type IV secretion systems constitute an interaction group by yeast two-hybrid analysis. *J. Bacteriol.* 186, 5480–5485. doi: 10.1128/JB.186.16.5480-5485.2004
- Harshey, R. M. (2003). Bacterial motility on a surface: many ways to a common goal. *Annu. Rev. Microbiol.* 57, 249–273. doi: 10.1146/annurev.micro.57.030502.091014
- Hartmann, M. D., Grin, I., Dunin-Horkawicz, S., Deiss, S., Linke, D., Lupas, A. N., et al. (2012). Complete fiber structures of complex trimeric autotransporter adhesins conserved in enterobacteria. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20907–20912. doi: 10.1073/pnas.1211872110
- Hazes, B., and Frost, L. (2008). Towards a systems biology approach to study Type II/IV secretion systems. *Biochim. Biophys. Acta Biomembr.* 1778, 1839–1850. doi: 10.1016/j.bbamem.2008.03.011
- Henderson, B., and Martin, A. (2011). Bacterial virulence in the moonlight: multitasking bacterial moonlighting proteins are virulence determinants in infectious disease. *Infect. Immun.* 79, 3476–3491. doi: 10.1128/IAI.00179-11
- Henderson, B., and Martin, A. (2013). Bacterial moonlighting proteins and bacterial virulence. *Curr. Top. Microbiol. Immunol.* 358, 155–213. doi: 10.1007/978_2011_188
- Henderson, I. R., and Desvaux, M. (2004). Type V secretion pathway: a premium source source of virulence factors? *Drug Discov. Today* 9, 241.
- Henderson, I. R., Nataro, J. P., Kaper, J. B., Meyer, T. F., Farrand, S. K., Burns, D. L., et al. (2000). Renaming protein secretion in the Gram-negative bacteria. *Trends Microbiol.* 8, 352. doi: 10.1016/S0966-842X(00)01814-X
- Henderson, I. R., Navarro-Garcia, F., Desvaux, M., Fernandez, R. C., and Ala'aldene, D. (2004). Type V protein secretion pathway: the autotransporter story. *Microbiol. Mol. Biol. Rev.* 68, 692–744.
- Henderson, I. R., Owen, P., and Nataro, J. P. (1999). Molecular switches: the ON and OFF of bacterial phase variation. *Mol. Microbiol.* 33, 919–932. doi: 10.1046/j.1365-2958.1999.01555.x
- Hermansson, M. (1999). The DLVO theory in microbial adhesion. *Colloids Surf. B. Biointerfaces* 14, 105–119. doi: 10.1016/S0927-7765(99)00029-6
- Herold, S., Paton, J. C., and Paton, A. W. (2009). Sab, a novel autotransporter of locus of enterocyte effacement-negative shiga-toxinogenic *Escherichia coli* O113:H21, contributes to adherence and biofilm formation. *Infect. Immun.* 77, 3234–3243. doi: 10.1128/IAI.00031-09
- Hinsa, S. M., Espinosa-Urgel, M., Ramos, J. L., and O'toole, G. A. (2003). Transition from reversible to irreversible attachment during biofilm formation by *Pseudomonas fluorescens* WCS365 requires an ABC transporter and a large secreted protein. *Mol. Microbiol.* 49, 905–918.

- Hinsa, S. M., and O'toole, G. A. (2006). Biofilm formation by *Pseudomonas fluorescens* WCS365: a role for LapD. *Microbiology* 152, 1375–1383.
- Hoiczky, E., Roggenkamp, A., Reichenbecher, M., Lupas, A., and Heesemann, J. (2000). Structure and sequence analysis of *Yersinia* YadA and *Moraxella* UspAs reveal a novel class of adhesins. *EMBO J.* 19, 5989–5999. doi: 10.1093/emboj/19.22.5989
- Holland, I. B., Schmitt, L., and Young, J. (2005). Type I protein secretion in bacteria, the ABC-transporter dependent pathway. *Mol. Membr. Biol.* 22, 29–39. doi: 10.1080/09687860500042013
- Houben, E. N., Besteboer, J., Ummels, R., Wilson, L., Piersma, S. R., Jimenez, C. R., et al. (2012). Composition of the Type VII secretion system membrane complex. *Mol. Microbiol.* 86, 472–484. doi: 10.1111/j.1365-2958.2012.08206.x
- Houry, A., Briandet, R., Aymerich, S., and Gohar, M. (2010). Involvement of motility and flagella in *Bacillus cereus* biofilm formation. *Microbiology* 156, 1009–1018. doi: 10.1099/mic.0.034827-0
- Houry, A., Gohar, M., Deschamps, J., Tischenko, E., Aymerich, S., Gruss, A., et al. (2012). Bacterial swimmers that infiltrate and take over the biofilm matrix. *Proc. Natl. Acad. Sci. U.S.A.* 109, 13088–13093. doi: 10.1073/pnas.1200791109
- Huber, B., Riedel, K., Kothe, M., Givskov, M., Molin, S., and Eberl, L. (2002). Genetic analysis of functions involved in the late stages of biofilm development in *Burkholderia cepacia* H111. *Mol. Microbiol.* 46, 411–426. doi: 10.1046/j.1365-2958.2002.03182.x
- Huberts, D. H., and Van Der Klei, I. J. (2010). Moonlighting proteins: an intriguing mode of multitasking. *Biochim. Biophys. Acta Mol. Cell. Res.* 1803, 520–525. doi: 10.1016/j.bbamcr.2010.01.022
- Hung, D. L., Knight, S. D., Woods, R. M., Pinkner, J. S., and Hultgren, S. J. (1996). Molecular basis of two subfamilies of immunoglobulin-like chaperones. *EMBO J.* 15, 3792–3805.
- Iiver, D., Arnqvist, A., Ogren, J., Frick, I. M., Kersulyte, D., Incecik, E. T., et al. (1998). *Helicobacter pylori* adhesin binding fucosylated histo-blood group antigens revealed by retagging. *Science* 279, 373–377. doi: 10.1126/science.279.5349.373
- Isberg, R. R., Hamburger, Z., and Dersch, P. (2000). Signaling and invasin-promoted uptake via integrin receptors. *Microbes Infect.* 2, 793–801. doi: 10.1016/S1286-4579(00)90364-2
- Ivanov, I. E., Boyd, C. D., Newell, P. D., Schwartz, M. E., Turnbull, L., Johnson, M. S., et al. (2012). Atomic force and super-resolution microscopy support a role for LapA as a cell-surface biofilm adhesin of *Pseudomonas fluorescens*. *Res. Microbiol.* 163, 685–691. doi: 10.1016/j.resmic.2012.10.001
- Ize, B., and Palmer, T. (2006). *Microbiology*. Mycobacteria's export strategy. *Science* 313, 1583–1584. doi: 10.1126/science.1132537
- Jacob-Dubuisson, F., Fernandez, R., and Coutte, L. (2004). Protein secretion through autotransporter and two-partner pathways. *Biochim. Biophys. Acta Mol. Cell. Res.* 1694, 235–257. doi: 10.1016/j.bbamcr.2004.03.008
- Jeffery, C. J. (2003). Moonlighting proteins: old proteins learning new tricks. *Trends Genet.* 19, 415–417. doi: 10.1016/S0168-9525(03)00167-7
- Jeffery, C. J. (2009). Moonlighting proteins—an update. *Mol. Biosyst.* 5, 345–350. doi: 10.1039/b900658n
- Jin, H., Song, Y. P., Boel, G., Kochar, J., and Pancholi, V. (2005). Group A streptococcal surface GAPDH, SDH, recognizes uPAR/CD87 as its receptor on the human pharyngeal cell and mediates bacterial adherence to host cells. *J. Mol. Biol.* 350, 27–41. doi: 10.1016/j.jmb.2005.04.063
- Jonas, K., Tomenius, R., Kader, A., Normark, S., Romling, U., Belova, L. M., et al. (2007). Roles of curli, cellulose and BapA in *Salmonella* biofilm morphology studied by atomic force microscopy. *BMC Microbiol.* 7:70. doi: 10.1186/1471-2180-7-70
- Jordan, S. J., Perni, S., Glenn, S., Fernandes, I., Barbosa, M., Sol, M., et al. (2008). *Listeria monocytogenes* biofilm-associated protein (BapL) may contribute to surface attachment of *L. monocytogenes* but is absent from many field isolates. *Appl. Environ. Microbiol.* 74, 5451–5456. doi: 10.1128/AEM.02419-07
- Journet, L., Hughes, K. T., and Cornelis, G. R. (2005). Type III secretion: a secretory pathway serving both motility and virulence. *Mol. Membr. Biol.* 22, 41–50. doi: 10.1080/09687860500041858
- Jucker, B. A., Zehnder, A. J. B., and Harms, H. (1998). Quantification of polymer interactions in bacterial adhesion. *Envir. Science & Technology* 32, 2909–2915. doi: 10.1021/es980211s
- Kachlany, S. C., Planet, P. J., Bhattacharjee, M. K., Kollia, E., Desalle, R., Fine, D. H., et al. (2000). Nonspecific adherence by *Actinobacillus actinomycetemcomitans* requires genes widespread in *Bacteria* and *Archaea*. *J. Bacteriol.* 182, 6169–6176. doi: 10.1128/JB.182.21.6169-6176.2000
- Kachlany, S. C., Planet, P. J., Desalle, R., Fine, D. H., and Figurski, D. H. (2001). Genes for tight adherence of *Actinobacillus actinomycetemcomitans*: from plaque to plague to pond scum. *Trends Microbiol.* 9, 429–437. doi: 10.1016/S0966-842X(01)02161-8
- Kang, H. J., and Baker, E. N. (2012). Structure and assembly of Gram-positive bacterial pili: unique covalent polymers. *Curr. Opin. Struct. Biol.* 22, 200–207. doi: 10.1016/j.sbi.2012.01.009
- Katsikogianni, M., and Missirlis, Y. F. (2004). Concise review of mechanisms of bacterial adhesion to biomaterials and of techniques used in estimating bacteria-material interactions. *Eur. Cell. Mater.* 8, 37–57.
- Kearns, D. B., Chu, E., Branda, S. S., Kolter, R., and Losick, R. (2005). A master regulator for biofilm formation by *Bacillus subtilis*. *Mol. Microbiol.* 55, 739–749. doi: 10.1111/j.1365-2958.2004.04440.x
- Kingsley, R. A., Abi Ghanem, D., Puebla-Orsorio, N., Keestra, A. M., Berghman, L., and Baumler, A. J. (2004). Fibronectin binding to the *Salmonella enterica* serotype Typhimurium ShdA autotransporter protein is inhibited by a monoclonal antibody recognizing the A3 repeat. *J. Bacteriol.* 186, 4931–4939. doi: 10.1128/JB.186.15.4931-4939.2004
- Kingsley, R. A., Santos, R. L., Keestra, A. M., Adams, L. G., and Baumler, A. J. (2002). *Salmonella enterica* serotype Typhimurium ShdA is an outer membrane fibronectin-binding protein that is expressed in the intestine. *Mol. Microbiol.* 43, 895–905. doi: 10.1046/j.1365-2958.2002.02805.x
- Kingsley, R. A., Van Amsterdam, K., Kramer, N., and Baumler, A. J. (2000). The *shdA* gene is restricted to serotypes of *Salmonella enterica* subspecies I and contributes to efficient and prolonged fecal shedding. *Infect. Immun.* 68, 2720–2727. doi: 10.1128/IAI.68.5.2720-2727.2000
- Kirn, T. J., Jude, B. A., and Taylor, R. K. (2005). A colonization factor links *Vibrio cholerae* environmental survival and human infection. *Nature* 438, 863–866. doi: 10.1038/nature04249
- Kirov, S. M., Castrisios, M., and Shaw, J. G. (2004). *Aeromonas* flagella (polar and lateral) are enterocyte adhesins that contribute to biofilm formation on surfaces. *Infect. Immun.* 72, 1939–1945. doi: 10.1128/IAI.72.4.1939-1945.2004
- Klemm, P., Hjerrild, L., Gjermansen, M., and Schembri, M. A. (2004). Structure-function analysis of the self-recognizing Antigen 43 autotransporter protein from *Escherichia coli*. *Mol. Microbiol.* 51, 283–296. doi: 10.1046/j.1365-2958.2003.03833.x
- Klemm, P., Vejborg, R. M., and Sherlock, O. (2006). Self-associating autotransporters, SAATs: functional and structural similarities. *Int. J. Med. Microbiol.* 296, 187–195. doi: 10.1016/j.ijmm.2005.10.002
- Knowles, T. J., Scott-Tucker, A., Overduin, M., and Henderson, I. R. (2009). Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat. Rev. Microbiol.* 7, 206–214. doi: 10.1038/nrmicro2069
- Kobayashi, K. (2007). *Bacillus subtilis* pellicle formation proceeds through genetically defined morphological changes. *J. Bacteriol.* 189, 4920–4931. doi: 10.1128/JB.00157-07
- Kobayashi, K., and Iwano, M. (2012). BslA(YuaB) forms a hydrophobic layer on the surface of *Bacillus subtilis* biofilms. *Mol. Microbiol.* 85, 51–66. doi: 10.1111/j.1365-2958.2012.08094.x
- Kondo, Y., Ohara, N., Sato, K., Yoshimura, M., Yukitake, H., Naito, M., et al. (2010). Tetratricopeptide repeat protein-associated proteins contribute to the virulence of *Porphyromonas gingivalis*. *Infect. Immun.* 78, 2846–2856. doi: 10.1128/IAI.01448-09
- Kostakioti, M., and Stathopoulos, C. (2004). Functional analysis of the Tsh autotransporter from an avian pathogenic *Escherichia coli* strain. *Infect. Immun.* 72, 5548–5554. doi: 10.1128/IAI.72.10.5548-5554.2004
- Kovacs, A. T., and Kuipers, O. P. (2011). Rok regulates *yuaB* expression during architecturally complex colony development of *Bacillus subtilis* 168. *J. Bacteriol.* 193, 998–1002. doi: 10.1128/JB.01170-10
- Kremer, B. H., Bijlsma, J. J., Kusters, J. G., de Graaff, J., and van Steenberg, T. J. (1999). Cloning of fibA, encoding an immunogenic subunit of the fibril-like surface

- structure of *Peptostreptococcus micros*. *J. Bacteriol.* 181, 2485–2491.
- Kroupitski, Y., Golberg, D., Belausov, E., Pinto, R., Swartzberg, D., Granot, D., et al. (2009a). Internalization of *Salmonella enterica* in leaves is induced by light and involves chemotaxis and penetration through open stomata. *Appl. Environ. Microbiol.* 75, 6076–6086. doi: 10.1128/AEM.01084-09
- Kroupitski, Y., Pinto, R., Brandl, M. T., Belausov, E., and Sela, S. (2009b). Interactions of *Salmonella enterica* with lettuce leaves. *J. Appl. Microbiol.* 106, 1876–1885. doi: 10.1111/j.1365-2672.2009.04152.x
- Kroupitski, Y., Pinto, R., Belausov, E., and Sela, S. (2011). Distribution of *Salmonella typhimurium* in romaine lettuce leaves. *Food Microbiol.* 28, 990–997. doi: 10.1016/j.fm.2011.01.007
- Lasa, I. (2006). Towards the identification of the common features of bacterial biofilm development. *Int. Microbiol.* 9, 21–28.
- Lasa, I., and Penadés, J. R. (2006). Bap: a family of surface proteins involved in biofilm formation. *Res. Microbiol.* 157, 99–107. doi: 10.1016/j.resmic.2005.11.003
- Latasa, C., Roux, A., Toledo-Arana, A., Ghigo, J. M., Gamazo, C., Penades, J. R., et al. (2005). BapA, a large secreted protein required for biofilm formation and host colonization of *Salmonella enterica* serovar *Enteritidis*. *Mol. Microbiol.* 58, 1322–1339. doi: 10.1111/j.1365-2958.2005.04907.x
- Latasa, C., Solano, C., Penades, J. R., and Lasa, I. (2006). Biofilm-associated proteins. *C. R. Biol.* 329, 849–857. doi: 10.1016/j.crv.2006.07.008
- Laurenceau, R., Pehau-Arnaudet, G., Bacconnais, S., Gault, J., Malosse, C., Dujeancourt, A., et al. (2013). A Type IV pilus mediates DNA binding during natural transformation in *Streptococcus pneumoniae*. *PLoS Pathog.* 9:e1003473. doi: 10.1371/journal.ppat.1003473
- Lawley, T. D., Klimke, W. A., Gubbins, M. J., and Frost, L. S. (2003). F factor conjugation is a true Type IV secretion system. *FEMS Microbiol. Lett.* 224, 1–15. doi: 10.1016/S0378-1097(03)00430-0
- Le Trong, I., Aprikian, P., Kidd, B. A., Forero-Shelton, M., Tchesnokova, V., Rajagopal, P., et al. (2010). Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like b sheet twisting. *Cell* 141, 645–655. doi: 10.1016/j.cell.2010.03.038
- Leduc, I., White, C. D., Nepluev, I., Throm, R. E., Spinola, S. M., and Elkins, C. (2008). Outer membrane protein DsrA is the major fibronectin-binding determinant of *Haemophilus ducreyi*. *Infect. Immun.* 76, 1608–1616. doi: 10.1128/IAI.00994-07
- Lee, M., Jun, S. Y., Yoon, B. Y., Song, S., Lee, K., and Ha, N. C. (2012). Membrane fusion proteins of Type I secretion system and tripartite efflux pumps share a binding motif for TolC in gram-negative bacteria. *PLoS ONE* 7:e40460. doi: 10.1371/journal.pone.0040460
- Leininger, E., Roberts, M., Kenimer, J. G., Charles, I. G., Fairweather, N., Novotny, P., et al. (1991). Pertactin, an Arg-Gly-Asp-containing *Bordetella pertussis* surface protein that promotes adherence of mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 88, 345–349. doi: 10.1073/pnas.88.2.345
- Lemon, K. P., Higgins, D. E., and Kolter, R. (2007). Flagellar motility is critical for *Listeria monocytogenes* biofilm formation. *J. Bacteriol.* 189, 4418–4424. doi: 10.1128/JB.01967-06
- Leo, J. C., Elovaara, H., Bihan, D., Pugh, N., Kilpinen, S. K., Raynal, N., et al. (2010). First analysis of a bacterial collagen-binding protein with collagen Toolkits: promiscuous binding of YadA to collagens may explain how YadA interferes with host processes. *Infect. Immun.* 78, 3226–3236. doi: 10.1128/IAI.01057-09
- Leo, J. C., Grin, I., and Linke, D. (2012). Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 1088–1101. doi: 10.1098/rstb.2011.0208
- Leyton, D. L., Rossiter, A. E., and Henderson, I. R. (2012). From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis. *Nat. Rev. Microbiol.* 10, 213–225. doi: 10.1038/nrmicro2733
- Ligon, L. S., Hayden, J. D., and Braunstein, M. (2012). The ins and outs of *Mycobacterium tuberculosis* protein export. *Tuberculosis (Edinb.)* 92, 121–132. doi: 10.1016/j.tube.2011.11.005
- Lillehoj, E. P., Kim, B. T., and Kim, K. C. (2002). Identification of *Pseudomonas aeruginosa* flagellin as an adhesin for Muc1 mucin. *Am. J. Physiol. Lung Cell Mol. Physiol.* 282, L751–L756.
- Linke, D., Riess, T., Autenrieth, I. B., Lupas, A., and Kempf, V. A. (2006). Trimeric autotransporter adhesins: variable structure, common function. *Trends Microbiol.* 14, 264–270. doi: 10.1016/j.tim.2006.04.005
- Lipski, S. L., Akimana, C., Timpe, J. M., Wooten, R. M., and Lafontaine, E. R. (2007). The *Moraxella catarrhalis* autotransporter McaP is a conserved surface protein that mediates adherence to human epithelial cells through its N-terminal passenger domain. *Infect. Immun.* 75, 314–324. doi: 10.1128/IAI.01330-06
- Locht, C., Bertin, P., Menozzi, F. D., and Renault, G. (1993). The filamentous haemagglutinin, a multifaceted adhesion produced by virulent *Bordetella* spp. *Mol. Microbiol.* 9, 653–660. doi: 10.1111/j.1365-2958.1993.tb01725.x
- Lyskowski, A., Leo, J. C., and Goldman, A. (2011). Structure and biology of trimeric autotransporter adhesins. *Adv. Exp. Med. Biol.* 715, 143–158. doi: 10.1007/978-94-007-0940-9_9
- Mahajan, A., Currie, C. G., Mackie, S., Tree, J., McAtter, S., McKendrick, I., et al. (2009). An investigation of the expression and adhesin function of H7 flagella in the interaction of *Escherichia coli* O157 : H7 with bovine intestinal epithelium. *Cell Microbiol.* 11, 121–137. doi: 10.1111/j.1462-5822.2008.01244.x
- Mahdavi, J., Sonden, B., Hurtig, M., Olfat, F. O., Forsberg, L., Roche, N., et al. (2002). *Helicobacter pylori* SabA adhesin in persistent infection and chronic inflammation. *Science* 297, 573–578. doi: 10.1126/science.1069076
- Mallick, E. M., Brady, M. J., Luperchio, S. A., Vanguri, V. K., Magoun, L., Liu, H., et al. (2012). Allele- and tir-independent functions of intimin in diverse animal infection models. *Front. Microbiol.* 3:11. doi: 10.3389/fmicb.2012.00011
- Matias, V. R., and Beveridge, T. J. (2005). Cryo-electron microscopy reveals native polymeric cell wall structure in *Bacillus subtilis* 168 and the existence of a periplasmic space. *Mol. Microbiol.* 56, 240–251. doi: 10.1111/j.1365-2958.2005.04535.x
- Matta, S. K., Agarwal, S., and Bhatnagar, R. (2010). Surface localized and extracellular Glyceraldehyde-3-phosphate dehydrogenase of *Bacillus anthracis* is a plasminogen binding protein. *Biochim. Biophys. Acta Proteins Proteomics* 1804, 2111–2120. doi: 10.1016/j.bbapap.2010.08.004
- Mattick, J. S. (2002). Type 4 pili and twitching motility. *Annu. Rev. Microbiol.* 56, 289–314. doi: 10.1146/annurev.micro.56.012302.160938
- Mazar, J., and Cotter, P. A. (2006). Topology and maturation of filamentous haemagglutinin suggest a new model for two-partner secretion. *Mol. Microbiol.* 62, 641–654. doi: 10.1111/j.1365-2958.2006.05392.x
- McBride, M. J., and Zhu, Y. (2013). Gliding motility and Por secretion system genes are widespread among members of the phylum *Bacteroidetes*. *J. Bacteriol.* 195, 270–278. doi: 10.1128/JB.01962-12
- McLaughlin, L. S., Haft, R. J. E., and Forest, K. T. (2012). Structural insights into the Type II secretion nanomachine. *Curr. Opin. Struct. Biol.* 22, 208–216. doi: 10.1016/j.sbi.2012.02.005
- Mendez, M., Huang, I. H., Ohtani, K., Grau, R., Shimizu, T., and Sarker, M. R. (2008). Carbon catabolite repression of type 4 pilus-dependent gliding motility in the anaerobic pathogen *Clostridium perfringens*. *J. Bacteriol.* 190, 48–60. doi: 10.1128/JB.01407-07
- Michel, G. P. E., and Voulhoux, R. (2009). “Chapter 4: Type II secretory system (T2SS) in Gram-negative bacteria: a molecular nanomachine for secretion of Sec and Tat-dependent extracellular proteins,” in *Bacterial Secreted Proteins: Secretory Mechanisms and Role in Pathogenesis*, ed K. Wooldridge (Norwich: Caister Academic Press), 67–92.
- Minamino, T., Imada, K., and Namba, K. (2008). Mechanisms of type III protein export for bacterial flagellar assembly. *Mol. Biosyst.* 4, 1105–1115. doi: 10.1039/b808065h
- Molina, M. A., Ramos, J. L., and Espinosa-Urgel, M. (2006). A two-partner secretion system is involved in seed and root colonization and iron uptake by *Pseudomonas putida* KT2440. *Environ. Microbiol.* 8, 639–647. doi: 10.1111/j.1462-2920.2005.00940.x
- Morgan, E., Campbell, J. D., Rowe, S. C., Bispham, J., Stevens, M. P., Bowen, A. J., et al. (2004). Identification of host-specific colonization factors of *Salmonella enterica* serovar *Typhimurium*. *Mol. Microbiol.* 54, 994–1010. doi: 10.1111/j.1365-2958.2004.04323.x
- Morra, M., and Cassinelli, C. (1996). *Staphylococcus epidermidis* adhesion to films deposited from hydroxyethylmethacrylate plasma. *J. Biomed. Mater. Res.* 31, 149–155. doi: 10.1002/(SICI)1097-4636(199606)31:2<149::AID-JBM1>3.0.CO;2-N
- Mosoni, P., and Gaillard-Martinie, B. (2001). Characterization of a

- spontaneous adhesion-defective mutant of *Ruminococcus albus* strain 20. *Arch. Microbiol.* 176, 52–61. doi: 10.1007/s002030100292
- Nakane, D., Sato, K., Wada, H., McBride, M. J., and Nakayama, K. (2013a). Helical flow of surface protein required for bacterial gliding motility. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11145–11150. doi: 10.1073/pnas.1219753110
- Nakane, D., Sato, K., Wada, H., McBride, M. J., and Nakayama, K. (2013b). Helical flow of surface protein required for bacterial locomotion. *Biophys. J.* 104, 639a–639a. doi: 10.1016/j.bpj.2012.11.3527
- Nakao, R., Ramstedt, M., Wai, S. N., and Uhlin, B. E. (2012). Enhanced biofilm formation by *Escherichia coli* LPS mutants defective in heparin biosynthesis. *PLoS ONE* 7:e51241. doi: 10.1371/journal.pone.0051241
- Nelson, S. S., Glocka, P. P., Agarwal, S., Grimm, D. P., and McBride, M. J. (2007). *Flavobacterium johnsoniae* SprA is a cell surface protein involved in gliding motility. *J. Bacteriol.* 189, 7145–7150. doi: 10.1128/JB.00892-07
- Newman, J. A., Rodrigues, C., and Lewis, R. J. (2013). Molecular basis of the activity of SinR protein, the master regulator of biofilm formation in *Bacillus subtilis*. *J. Biol. Chem.* 288, 10766–10778. doi: 10.1074/jbc.M113.455592
- Niederweis, M. (2003). Mycobacterial porins—new channel proteins in unique outer membranes. *Mol. Microbiol.* 49, 1167–1177. doi: 10.1046/j.1365-2958.2003.03662.x
- Niederweis, M., Danilchanka, O., Huff, J., Hoffmann, C., and Engelhardt, H. (2010). Mycobacterial outer membranes: in search of proteins. *Trends Microbiol.* 18, 109–116. doi: 10.1016/j.tim.2009.12.005
- Nilsson, L. M., Thomas, W. E., Sokurenko, E. V., and Vogel, V. (2006). Elevated shear stress protects *Escherichia coli* cells adhering to surfaces via catch bonds from detachment by soluble inhibitors. *Appl. Environ. Microbiol.* 72, 3005–3010. doi: 10.1128/AEM.72.4.3005-3010.2006
- Nuccio, S. P., and Baumler, A. J. (2007). Evolution of the chaperone/usher assembly pathway: fimbrial classification goes Greek. *Microbiol. Mol. Biol. Rev.* 71, 551–575. doi: 10.1128/MMBR.00014-07
- Nummelin, H., Merckel, M. C., Leo, J. C., Lankinen, H., Skurnik, M., and Goldman, A. (2004). The *Yersinia* adhesin YadA collagen-binding domain structure is a novel left-handed parallel beta-roll. *EMBO J.* 23, 701–711. doi: 10.1038/sj.emboj.7600100
- Nunn, D. (1999). Bacterial Type II protein export and pilus biogenesis: more than just homologies? *Trends Cell Biol.* 9, 402–408. doi: 10.1016/S0962-8924(99)01634-7
- Oberhettinger, P., Schutz, M., Leo, J. C., Heinz, N., Berger, J., Autenrieth, I. B., et al. (2012). Intimin and invasins export their C-terminus to the bacterial cell surface using an inverse mechanism compared to classical autotransport. *PLoS ONE* 7:e47069. doi: 10.1371/journal.pone.0047069
- O'Connell Motherway, M., Zomer, A., Leahy, S. C., Reunanen, J., Bottacini, F., Claesson, M. J., et al. (2011). Functional genome analysis of *Bifidobacterium breve* UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonization factor. *Proc. Natl. Acad. Sci. U.S.A.* 108, 11217–11222. doi: 10.1073/pnas.1105380108
- Odenbreit, S., Till, M., Hofreuter, D., Faller, G., and Haas, R. (1999). Genetic and functional characterization of the *alpAB* gene locus essential for the adhesion of *Helicobacter pylori* to human gastric tissue. *Mol. Microbiol.* 31, 1537–1548. doi: 10.1046/j.1365-2958.1999.01300.x
- Ojha, A., Anand, M., Bhatt, A., Kremer, L., Jacobs, W. R., and Hatfull, G. F. (2005). GroEL1: a dedicated chaperone involved in mycolic acid biosynthesis during biofilm formation in mycobacteria. *Cell* 123, 861–873. doi: 10.1016/j.cell.2005.09.012
- O'Neil, H. S., and Marquis, H. (2006). *Listeria monocytogenes* flagella are used for motility, not as adhesins, to increase host cell invasion. *Infect. Immun.* 74, 6675–6681. doi: 10.1128/IAI.00886-06
- Ostrowski, A., Mehert, A., Prescott, A., Kiley, T. B., and Stanley-Wall, N. R. (2011). YuaB functions synergistically with the exopolysaccharide and TasA amyloid fibers to allow biofilm formation by *Bacillus subtilis*. *J. Bacteriol.* 193, 4821–4831. doi: 10.1128/JB.00223-11
- Ottow, J. C. (1975). Ecology, physiology, and genetics of fimbriae and pili. *Annu. Rev. Microbiol.* 29, 79–108. doi: 10.1146/annurev.mi.29.100175.000455
- Pallen, M. J., Beatson, S. A., and Bailey, C. M. (2005). Bioinformatics, genomics and evolution of non-flagellar Type-III secretion systems: a Darwinian perspective. *FEMS Microbiol. Rev.* 29, 201–229. doi: 10.1016/j.femsre.2005.01.001
- Palumbo, R. N., and Wang, C. (2006). Bacterial invasins: structure, function, and implication for targeted oral gene delivery. *Curr. Drug Deliv.* 3, 47–53. doi: 10.2174/156720106775197475
- Parham, N. J., Pollard, S. J., Chaudhuri, R. R., Beatson, S. A., Desvaux, M., Russell, M. A., et al. (2005). Prevalence of pathogenicity island IICFT073 genes among extraintestinal clinical isolates of *Escherichia coli*. *J. Clin. Microbiol.* 43, 2425–2434. doi: 10.1128/JCM.43.5.2425-2434.2005
- Payot, S., Guedon, E., Desvaux, M., Gelhaye, E., and Petitdemange, E. (1999). Effect of dilution rate, cellobiose and ammonium availabilities on *Clostridium cellulolyticum* sporulation. *Appl. Microbiol. Biotechnol.* 52, 670–674. doi: 10.1007/s002530051577
- Peabody, C. R., Chung, Y. J., Yen, M. R., Vidal-Ingiguardi, D., Pugsley, A. P., and Saier, M. H. Jr. (2003). Type II protein secretion and its relationship to bacterial Type IV pili and archaeal flagella. *Microbiology* 149, 3051–3072. doi: 10.1099/mic.0.26364-0
- Pearson, M. M., Lafontaine, E. R., Wagner, N. J., St Geme, J. W. 3rd, and Hansen, E. J. (2002). A hag mutant of *Moraxella catarrhalis* strain O35E is deficient in hemagglutination, autoagglutination, and immunoglobulin D-binding activities. *Infect. Immun.* 70, 4523–4533. doi: 10.1128/IAI.70.8.4523-4533.2002
- Pegden, R. S., Larson, M. A., Grant, R. J., and Morrison, M. (1998). Adherence of the Gram-positive bacterium *Ruminococcus albus* to cellulose and identification of a novel form of cellulose-binding protein which belongs to the Pil family of proteins. *J. Bacteriol.* 180, 5921–5927.
- Pellic, V. (2008). Type4 pili: e pluribus unum? *Mol. Microbiol.* 68, 827–837. doi: 10.1111/j.1365-2958.2008.06197.x
- Planet, P. J., Kachlany, S. C., Desalle, R., and Figurski, D. H. (2001). Phylogeny of genes for secretion NTPases: identification of the widespread tadA subfamily and development of a diagnostic key for gene classification. *Proc. Natl. Acad. Sci. U.S.A.* 98, 2503–2508. doi: 10.1073/pnas.051436598
- Poole, S. T., McVeigh, A. L., Anantha, R. P., Lee, L. H., Akay, Y. M., Pontzer, E. A., et al. (2007). Donor strand complementation governs intersubunit interaction of fimbriae of the alternate chaperone pathway. *Mol. Microbiol.* 63, 1372–1384. doi: 10.1111/j.1365-2958.2007.05612.x
- Pratt, L. A., and Kolter, R. (1998). Genetic analysis of *Escherichia coli* biofilm formation: roles of flagella, motility, chemotaxis and Type 1 pili. *Mol. Microbiol.* 30, 285–293. doi: 10.1046/j.1365-2958.1998.01061.x
- Pugsley, A. P., and Francetic, O. (1998). Protein secretion in *Escherichia coli* K-12: dead or alive? *Cell Mol. Life Sci.* 54, 347–352. doi: 10.1007/s000180050162
- Puttamreddy, S., Cornick, N. A., and Minion, F. C. (2010). Genome-wide transposon mutagenesis reveals a role for pO157 genes in biofilm development in *Escherichia coli* O157:H7 EDL933. *Infect. Immun.* 78, 2377–2384. doi: 10.1128/IAI.00156-10
- Rabinovich, L., Sigal, N., Borovok, I., Nir-Paz, R., and Herskovits, A. A. (2012). Prophage excision activates *Listeria* competence genes that promote phagosomal escape and virulence. *Cell* 150, 792–802. doi: 10.1016/j.cell.2012.06.036
- Raghunathan, D., Wells, T. J., Morris, F. C., Shaw, R. K., Bobat, S., Peters, S. E., et al. (2011). SadA, a trimeric autotransporter from *Salmonella enterica* serovar Typhimurium, can promote biofilm formation and provides limited protection against infection. *Infect. Immun.* 79, 4342–4352. doi: 10.1128/IAI.05592-11
- Rakotoarivonina, H., Jubelin, G., Hébraud, M., Gaillard-Martinie, B., Forano, E., and Mosoni, P. (2002). Adhesion to cellulose of the Gram-positive bacterium *Ruminococcus albus* involves Type 4 pili. *Microbiology* 148, 1871–1880.
- Rakotoarivonina, H., Larson, M. A., Morrison, M., Girardeau, J. P., Gaillard-Martinie, B., Forano, E., et al. (2005). The *Ruminococcus albus* pilA1-pilA2 locus: expression and putative role of two adjacent pil genes in pilus formation and bacterial adhesion to cellulose. *Microbiology* 151, 1291–1299. doi: 10.1099/mic.0.27735-0
- Ramphal, R., Arora, S. K., and Ritchings, B. W. (1996). Recognition of mucin by the adhesin-flagellar system of *Pseudomonas aeruginosa*. *Am. J. Respir. Crit. Care Med.* 154, S170–S174. doi: 10.1164/ajrccm/154.4.Pt_2.S170
- Renier, S., Chambon, C., Viala, D., Chagnot, C., Hébraud, M., and Desvaux, M. (2013). Exoproteomic analysis of the SecA2-dependent secretion in *Listeria monocytogenes*

- EGD-e. *J. Proteomics* 80C, 183–195. doi: 10.1016/j.jprot.2012.11.027
- Renier, S., Hebraud, M., and Desvaux, M. (2011). Molecular biology of surface colonization by *Listeria monocytogenes*: an additional facet of an opportunistic Gram-positive foodborne pathogen. *Environ. Microbiol.* 13, 835–850. doi: 10.1111/j.1462-2920.2010.02378.x
- Renier, S., Micheau, P., Talon, R., Hebraud, M., and Desvaux, M. (2012). Subcellular localization of extracytoplasmic proteins in monoderm bacteria: rational secretomics-based strategy for genomic and proteomic analyses. *PLoS ONE* 7:e24982. doi: 10.1371/journal.pone.0042982
- Riess, T., Andersson, S. G., Lupas, A., Schaller, M., Schafer, A., Kyme, P., et al. (2004). *Bartonella* adhesin A mediates a proangiogenic host cell response. *J. Exp. Med.* 200, 1267–1278. doi: 10.1084/jem.20040500
- Rigel, N. W., and Braunstein, M. (2008). A new twist on an old pathway—accessory Sec systems. *Mol. Microbiol.* 69, 291–302. doi: 10.1111/j.1365-2958.2008.06294.x
- Rincon, M. T., Ding, S. Y., McCrae, S. I., Martin, J. C., Aurilia, V., Lamed, R., et al. (2003). Novel organization and divergent dockerin specificities in the cellulosome system of *Ruminococcus flavefaciens*. *J. Bacteriol.* 185, 703–713. doi: 10.1128/JB.185.3.703-713.2003
- Roder, H. L., Hansen, L. H., Sorensen, S. J., and Burmolle, M. (2013). The impact of the conjugative IncP-1 plasmid pKJK5 on multispecies biofilm formation is dependent on the plasmid host. *FEMS Microbiol. Lett.* 344, 186–192. doi: 10.1111/1574-6968.12175
- Rojas, C. M., Ham, J. H., Deng, W. L., Doyle, J. J., and Collmer, A. (2002). HecA, a member of a class of adhesins produced by diverse pathogenic bacteria, contributes to the attachment, aggregation, epidermal cell killing, and virulence phenotypes of *Erwinia chrysanthemi* EC16 on *Nicotiana glauca* seedlings. *Proc. Natl. Acad. Sci. U.S.A.* 99, 13142–13147. doi: 10.1073/pnas.202358699
- Romero, D., Aguilar, C., Losick, R., and Kolter, R. (2010). Amyloid fibers provide structural integrity to *Bacillus subtilis* biofilms. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2230–2234. doi: 10.1073/pnas.0910560107
- Romero, D., Vlamakis, H., Losick, R., and Kolter, R. (2011). An accessory protein required for anchoring and assembly of amyloid fibres in *B. subtilis* biofilms. *Mol. Microbiol.* 80, 1155–1168. doi: 10.1111/j.1365-2958.2011.07653.x
- Rose, J. E., Meyer, D. H., and Fives-Taylor, P. M. (2003). Aae, an autotransporter involved in adhesion of *Actinobacillus actinomycetemcomitans* to epithelial cells. *Infect. Immun.* 71, 2384–2393. doi: 10.1128/IAI.71.5.2384-2393.2003
- Rossiter, A. E., Browning, D. E., Leyton, D. L., Johnson, M. D., Godfrey, R. E., Wardius, C. A., et al. (2011a). Transcription of the plasmid-encoded toxin gene from enteroaggregative *Escherichia coli* is regulated by a novel co-activation mechanism involving CRP and Fis. *Mol. Microbiol.* 81, 179–191. doi: 10.1111/j.1365-2958.2011.07685.x
- Rossiter, A. E., Leyton, D. L., Tveen-Jensen, K., Browning, D. E., Sevastyanovich, Y., Knowles, T. J., et al. (2011b). The essential b-barrel assembly machinery complex components BamD and BamA are required for autotransporter biogenesis. *J. Bacteriol.* 193, 4250–4253. doi: 10.1128/JB.00192-11
- Roussel-Jazede, V., Jongerius, I., Bos, M. P., Tommassen, J., and Van Ulsen, P. (2010). NalP-mediated proteolytic release of lactoferrin-binding protein B from the meningococcal cell surface. *Infect. Immun.* 78, 3083–3089. doi: 10.1128/IAI.01193-09
- Roux, A., Beloin, C., and Ghigo, J. M. (2005). Combined inactivation and expression strategy to study gene function under physiological conditions: application to identification of new *Escherichia coli* adhesins. *J. Bacteriol.* 187, 1001–1013. doi: 10.1128/JB.187.3.1001-1013.2005
- Roy, K., Hilliard, G. M., Hamilton, D. J., Luo, J., Ostrmann, M. M., and Fleckenstein, J. M. (2009). Enterotoxigenic *Escherichia coli* EtpA mediates adhesion between flagella and host cells. *Nature* 457, 594–598. doi: 10.1038/nature07568
- Ruiz-Perez, F., Henderson, I. R., and Nataro, J. P. (2010). Interaction of FkpA, a peptidyl-prolyl cis/trans isomerase with EspP autotransporter protein. *Gut Microbes* 1, 339–344. doi: 10.4161/gmic.1.5.13436
- Ryder, C., Byrd, M., and Wozniak, D. J. (2007). Role of polysaccharides in *Pseudomonas aeruginosa* biofilm development. *Curr. Opin. Microbiol.* 10, 644–648. doi: 10.1016/j.mib.2007.09.010
- Salacha, R., Kovacic, F., Brochier-Armanet, C., Wilhelm, S., Tommassen, J., Filloux, A., et al. (2010). The *Pseudomonas aeruginosa* patatin-like protein PlpD is the archetype of a novel Type V secretion system. *Environ. Microbiol.* 12, 1498–1512.
- Salmond, G. P., and Reeves, P. J. (1993). Membrane traffic wardens and protein secretion in Gram-negative bacteria. *Trends Biochem. Sci.* 18, 7–12. doi: 10.1016/0968-0004(93)90080-7
- Sandkvist, M. (2001). Biology of Type II secretion. *Mol. Microbiol.* 40, 271–283. doi: 10.1046/j.1365-2958.2001.02403.x
- Sarvas, M., Harwood, C. R., Bron, S., and Van Dijk, J. M. (2004). Post-translocational folding of secretory proteins in Gram-positive bacteria. *Biochim. Biophys. Acta Mol. Cell. Res.* 1694, 311–327. doi: 10.1016/j.bbamcr.2004.04.009
- Sato, K., Naito, M., Yukitake, H., Hirakawa, H., Shoji, M., McBride, M. J., et al. (2010). A protein secretion system linked to bacteroidete gliding motility and pathogenesis. *Proc. Natl. Acad. Sci. U.S.A.* 107, 276–281. doi: 10.1073/pnas.0912010107
- Sato, K., Yukitake, H., Narita, Y., Shoji, M., Naito, M., and Nakayama, K. (2013). Identification of *Porphyromonas gingivalis* proteins secreted by the Por secretion system. *FEMS Microbiol. Lett.* 338, 68–76. doi: 10.1111/1574-6968.12028
- Sauer, F. G., Remaut, H., Hultgren, S. J., and Waksman, G. (2004). Fiber assembly by the chaperone-usher pathway. *Biochim. Biophys. Acta Mol. Cell. Res.* 1694, 259–267. doi: 10.1016/j.bbamcr.2004.02.010
- Saunders, N. J., Jeffries, A. C., Peden, J. F., Hood, D. W., Tettelin, H., Rappuoli, R., et al. (2000). Repeat-associated phase variable genes in the complete genome sequence of *Neisseria meningitidis* strain MC58. *Mol. Microbiol.* 37, 207–215. doi: 10.1046/j.1365-2958.2000.02000.x
- Sava, I. G., Heikens, E., Kropec, A., Theilacker, C., Willems, R., and Huebner, J. (2010). Enterococcal surface protein contributes to persistence in the host but is not a target of opsonic and protective antibodies in *Enterococcus faecium* infection. *J. Med. Microbiol.* 59, 1001–1004. doi: 10.1099/jmm.0.020578-0
- Scarselli, M., Serruto, D., Montanari, P., Capecci, B., Adu-Bobie, J., Veggi, D., et al. (2006). *Neisseria meningitidis* NnhA is a multifunctional trimeric autotransporter adhesin. *Mol. Microbiol.* 61, 631–644. doi: 10.1111/j.1365-2958.2006.05261.x
- Schmitt, C., Turner, D., Boesl, M., Abele, M., Frosch, M., and Kurzai, O. (2007). A functional two-partner secretion system contributes to adhesion of *Neisseria meningitidis* to epithelial cells. *J. Bacteriol.* 189, 7968–7976. doi: 10.1128/JB.00851-07
- Schneewind, O., and Missiakas, D. M. (2012). Protein secretion and surface display in Gram-positive bacteria. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 1123–1139. doi: 10.1098/rstb.2011.0210
- Serruto, D., Adu-Bobie, J., Scarselli, M., Veggi, D., Pizzi, M., Rappuoli, R., et al. (2003). *Neisseria meningitidis* App, a new adhesin with autocatalytic serine protease activity. *Mol. Microbiol.* 48, 323–334. doi: 10.1046/j.1365-2958.2003.03420.x
- Serruto, D., Spadafina, T., Ciocchi, L., Lewis, L. A., Ram, S., Tontini, M., et al. (2010). *Neisseria meningitidis* GNA2132, a heparin-binding protein that induces protective immunity in humans. *Proc. Natl. Acad. Sci. U.S.A.* 107, 3770–3775. doi: 10.1073/pnas.0915162107
- Sexton, J. A., and Vogel, J. P. (2002). Type IVb secretion by intracellular pathogens. *Traffic* 3, 178–185. doi: 10.1034/j.1600-0854.2002.030303.x
- Shatalkin, A. I. (2004). Highest level of division in classification of organisms. 3. Monodermata and Didermata. *Zh. Obshch. Biol.* 65, 195–210.
- Shaw, R. K., Berger, C. N., Feys, B., Knutton, S., Pallen, M. J., and Frankel, G. (2008). Enterohemorrhagic *Escherichia coli* exploits EspA filaments for attachment to salad leaves. *Appl. Environ. Microbiol.* 74, 2908–2914. doi: 10.1128/AEM.02704-07
- Shaw, R. K., Berger, C. N., Pallen, M. J., Sjoling, A., and Frankel, G. (2011). Flagella mediate attachment of enterotoxigenic *Escherichia coli* to fresh salad leaves. *Environ. Microbiol. Rep.* 3, 112–117. doi: 10.1111/j.1758-2229.2010.00195.x
- Sherlock, O., Schembri, M. A., Reisner, A., and Klemm, P. (2004). Novel roles for the AIDA adhesin from diarrheagenic *Escherichia coli*: cell aggregation and biofilm formation. *J. Bacteriol.* 186, 8058–8065. doi: 10.1128/JB.186.23.8058-8065.2004
- Sherlock, O., Vejborg, R. M., and Klemm, P. (2005). The TibA adhesin/invasin from enterotoxigenic *Escherichia coli* is self recognizing and induces bacterial aggregation and biofilm formation. *Infect. Immun.* 73, 1954–1963. doi: 10.1128/IAI.73.4.1954-1963.2005
- Shimoji, Y., Ogawa, Y., Osaki, M., Kabeya, H., Maruyama, S., Mikami, T., et al. (2003).

- Adhesive surface proteins of *Erysipelothrix rhusiopathiae* bind to polystyrene, fibronectin, and type I and IV collagens. *J. Bacteriol.* 185, 2739–2748. doi: 10.1128/JB.185.9.2739-2748.2003
- Shircliff, M. E., Mader, J. T., and Camper, A. K. (2002). Molecular interactions in biofilms. *Chem. Biol.* 9, 859–871. doi: 10.1016/S1074-5521(02)00198-9
- Shoham, Y., Lamed, R., and Bayer, E. A. (1999). The cellulosome concept as an efficient microbial strategy for the degradation of insoluble polysaccharides. *Trends Microbiol.* 7, 275–281. doi: 10.1016/S0966-842X(99)01533-4
- Shoji, M., Sato, K., Yukitake, H., Kondo, Y., Narita, Y., Kadowaki, T., et al. (2011). Por secretion system-dependent secretion and glycosylation of *Porphyromonas gingivalis* hemin-binding protein 35. *PLoS ONE* 6:e21372. doi: 10.1371/journal.pone.0021372
- Shoseyov, O., Shani, Z., and Levy, I. (2006). Carbohydrate binding modules: biochemical properties and novel applications. *Microbiol. Mol. Biol. Rev.* 70, 283–295. doi: 10.1128/MMBR.00028-05
- Shrivastava, A., Johnston, J. J., Van Baaren, J. M., and McBride, M. J. (2013). *Flavobacterium johnsoniae* GldK, GldL, GldM, and SprA are required for secretion of the cell surface gliding motility Adhesins SprB and RemA. *J. Bacteriol.* 195, 3201–32012. doi: 10.1128/JB.00333-13
- Shrivastava, A., Rhodes, R. G., Pochiraju, S., Nakane, D., and McBride, M. J. (2012). *Flavobacterium johnsoniae* RemA is a mobile cell surface lectin involved in gliding. *J. Bacteriol.* 194, 3678–3688. doi: 10.1128/JB.00588-12
- Silverman, J. M., Brunet, Y. R., Cascales, E., and Mougous, J. D. (2012). Structure and regulation of the Type VI secretion system. *Ann. Rev. Microbiol.* 66, 453–472. doi: 10.1146/annurev-micro-121809-151619
- Skerker, J. M., and Shapiro, L. (2000). Identification and cell cycle control of a novel pilus system in *Caulobacter crescentus*. *EMBO J.* 19, 3223–3234. doi: 10.1093/emboj/19.13.3223
- Skorko-Glonek, J., and Sobiecka-Szatula, A. (2008). The extracytoplasmic protein quality control in bacterium *Escherichia coli*: the role of proteases and the folding factors. *Postepy Biochem.* 54, 317–326.
- Slakeski, N., Seers, C. A., Ng, K., Moore, C., Cleal, S. M., Veith, P. D., et al. (2011). C-terminal domain residues important for secretion and attachment of RgpB in *Porphyromonas gingivalis*. *J. Bacteriol.* 193, 132–142. doi: 10.1128/JB.00773-10
- Song, H., Sandie, R., Wang, Y., Andrade-Navarro, M. A., and Niederweis, M. (2008). Identification of outer membrane proteins of *Mycobacterium tuberculosis*. *Tuberculosis (Edinb.)* 88, 526–544. doi: 10.1016/j.tube.2008.02.004
- Soto, G. E., and Hultgren, S. J. (1999). Bacterial adhesins: common themes and variations in architecture and assembly. *J. Bacteriol.* 181, 1059–1071.
- St Geme, J. W. 3rd., and Cutter, D. (2000). The *Haemophilus influenzae* Hia adhesin is an autotransporter protein that remains uncleaved at the C terminus and fully cell associated. *J. Bacteriol.* 182, 6005–6013. doi: 10.1128/JB.182.21.6005-6013.2000
- St Geme, J. W. 3rd., Falkow, S., and Barenkamp, S. J. (1993). High-molecular-weight proteins of nontypable *Haemophilus influenzae* mediate attachment to human epithelial cells. *Proc. Natl. Acad. Sci. U.S.A.* 90, 2875–2879. doi: 10.1073/pnas.90.7.2875
- St Geme, J. W. 3rd., and Yeo, H. J. (2009). A prototype two-partner secretion pathway: the *Haemophilus influenzae* HMW1 and HMW2 adhesin systems. *Trends Microbiol.* 17, 355–360. doi: 10.1016/j.tim.2009.06.002
- Stauder, M., Huq, A., Pezzati, E., Grim, C. J., Ramoino, P., Pane, L., et al. (2012). Role of GbpA protein, an important virulence-related colonization factor, for *Vibrio cholerae*'s survival in the aquatic environment. *Environ. Microbiol. Rep.* 4, 439–445. doi: 10.1111/j.1758-2229.2012.00356.x
- Stoop, E. J., Bitter, W., and Van Der Sar, A. M. (2012). Tubercle bacilli rely on a type VII army for pathogenicity. *Trends Microbiol.* 20, 477–484. doi: 10.1016/j.tim.2012.07.001
- Stover, A. G., and Driks, A. (1999). Control of synthesis and secretion of the *Bacillus subtilis* protein Yqxm. *J. Bacteriol.* 181, 7065–7069.
- Strom, M. S., and Lory, S. (1993). Structure-function and biogenesis of the Type 4 pili. *Annu. Rev. Microbiol.* 47, 565–596. doi: 10.1146/annurev.mi.47.100193.003025
- Suez, J., Porwollik, S., Dagan, A., Marzel, A., Schorr, Y. I., Desai, P. T., et al. (2013). Virulence gene profiling and pathogenicity characterization of non-typhoidal *Salmonella* accounted for invasive disease in humans. *PLoS ONE* 8:e58449. doi: 10.1371/journal.pone.0058449
- Sutcliffe, I. C. (2010). A phylum level perspective on bacterial cell envelope architecture. *Trends Microbiol.* 18, 464–470. doi: 10.1016/j.tim.2010.06.005
- Swierczynski, A., and Ton-That, H. (2006). Type 3 pilus of *Corynebacteria*: pilus length is determined by the level of its major pilin subunit. *J. Bacteriol.* 188, 6318–6325. doi: 10.1128/JB.00606-06
- Tampakaki, A. P., Fadoulglou, V. E., Gazi, A. D., Panopoulos, N. J., and Kokkinidis, M. (2004). Conserved features of Type III secretion. *Cell Microbiol.* 6, 805–816. doi: 10.1111/j.1462-5822.2004.00432.x
- Tampakaki, A. P., Skandalis, N., Gazi, A. D., Bastaki, M. N., Sarris, P. F., Charova, S. N., et al. (2010). Playing the “Harp”: evolution of our understanding of hrp/hrc genes. *Annu. Rev. Phytopathol.* 48, 347–370. doi: 10.1146/annurev-phyto-073009-114407
- Tang, J., Chen, J., Li, H., Zeng, P., and Li, J. (2013). Characterization of adhesin genes, staphylococcal nuclease, hemolysis, and biofilm formation among *Staphylococcus aureus* strains isolated from different sources. *Foodborne Pathog. Dis.* 10, 757–763. doi: 10.1089/fpd.2012.1474
- Tang, X., Xiao, Y., and Zhou, J. M. (2006). Regulation of the Type III secretion system in phytopathogenic bacteria. *Mol. Plant Microbe Interact.* 19, 1159–1166. doi: 10.1094/MPMI-19-1159
- Telford, J. L., Barocchi, M. A., Margarit, I., Rappuoli, R., and Grandi, G. (2006). Pili in Gram-positive pathogens. *Nat. Rev. Microbiol.* 4, 509–519. doi: 10.1038/nrmicro1443
- Tendolkar, P. M., Baghdayan, A. S., Gilmore, M. S., and Shankar, N. (2004). Enterococcal surface protein, Esp, enhances biofilm formation by *Enterococcus faecalis*. *Infect. Immun.* 72, 6032–6039. doi: 10.1128/IAI.72.10.6032-6039.2004
- Terrasse, R., Tacnet-Delorme, P., Moriscot, C., Perard, J., Schoehn, G., Vernet, T., et al. (2012). Human and pneumococcal cell surface glyceraldehyde-3-phosphate dehydrogenase (GAPDH) proteins are both ligands of human C1q protein. *J. Biol. Chem.* 287, 42620–42633. doi: 10.1074/jbc.M112.423731
- Thanassi, D. G., Bliska, J. B., and Christie, P. J. (2012). Surface organelles assembled by secretion systems of Gram-negative bacteria: diversity in structure and function. *FEMS Microbiol. Rev.* 36, 1046–1082. doi: 10.1111/j.1574-6976.2012.00342.x
- Timpe, J. M., Holm, M. M., Vanlerberg, S. L., Basrur, V., and Lafontaine, E. R. (2003). Identification of a *Moraxella catarrhalis* outer membrane protein exhibiting both adhesin and lipolytic activities. *Infect. Immun.* 71, 4341–4350. doi: 10.1128/IAI.71.8.4341-4350.2003
- Tjalsma, H. (2007). Feature-based reappraisal of the *Bacillus subtilis* exoproteome. *Proteomics* 7, 73–81. doi: 10.1002/pmic.200600520
- Tjalsma, H., Bolhuis, A., Jongbloed, J. D., Bron, S., and Van Dijk, J. M. (2000). Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. *Microbiol. Mol. Biol. Rev.* 64, 515–547. doi: 10.1128/MMBR.64.3.515-547.2000
- Todhanakasem, T., and Young, G. M. (2008). Loss of flagellum-based motility by *Listeria monocytogenes* results in formation of hyperbiofilms. *J. Bacteriol.* 190, 6030–6034. doi: 10.1128/JB.00155-08
- Tomich, M., Planet, P. J., and Figurski, D. H. (2007). The *tad* locus: postcards from the widespread colonization island. *Nat. Rev. Microbiol.* 5, 363–375. doi: 10.1038/nrmicro1636
- Ton-That, H., Marraffini, L. A., and Schneewind, O. (2004). Protein sorting to the cell wall envelope of Gram-positive bacteria. *Biochim. Biophys. Acta Mol. Cell. Res.* 1694, 269–278. doi: 10.1016/j.bbamcr.2004.04.014
- Ton-That, H., and Schneewind, O. (2003). Assembly of pili on the surface of *Corynebacterium diphtheriae*. *Mol. Microbiol.* 50, 1429–1438. doi: 10.1046/j.1365-2958.2003.03782.x
- Ton-That, H., and Schneewind, O. (2004). Assembly of pili in Gram-positive bacteria. *Trends Microbiol.* 12, 228–234. doi: 10.1016/j.tim.2004.03.004
- Tormo, M. A., Knecht, E., Gotz, F., Lasa, I., and Penades, J. R. (2005). Bap-dependent biofilm formation by pathogenic species of *Staphylococcus*: evidence of horizontal gene transfer? *Microbiology* 151, 2465–2475. doi: 10.1099/mic.0.27865-0
- Tormo, M. A., Ubeda, C., Marti, M., Maiques, E., Cucarella, C., Valle, J., et al. (2007). Phase-variable expression of the biofilm-associated

- protein (Bap) in *Staphylococcus aureus*. *Microbiology* 153, 1702–1710. doi: 10.1099/mic.0.2006/005744-0
- Totsika, M., Wells, T. J., Beloin, C., Valle, J., Allsopp, L. P., King, N. P., et al. (2012). Molecular characterization of the EhaG and UpaG trimeric autotransporter proteins from pathogenic *Escherichia coli*. *Appl. Environ. Microbiol.* 78, 2179–2189. doi: 10.1128/AEM.06680-11
- Tran, V. B., Fleiszig, S. M., Evans, D. J., and Radke, C. J. (2011). Dynamics of flagellum- and pilus-mediated association of *Pseudomonas aeruginosa* with contact lens surfaces. *Appl. Environ. Microbiol.* 77, 3644–3652. doi: 10.1128/AEM.02656-10
- Tresse, O., Lebre, V., Benezech, T., and Faille, C. (2006). Comparative evaluation of adhesion, surface properties, and surface protein composition of *Listeria monocytogenes* strains after cultivation at constant pH of 5 and 7. *J. Appl. Microbiol.* 101, 53–62. doi: 10.1111/j.1365-2672.2006.02968.x
- Troge, A., Scheppach, W., Schroeder, B. O., Rund, S. A., Heuner, K., Wehkamp, J., et al. (2012). More than a marine propeller—the flagellum of the probiotic *Escherichia coli* strain Nissle 1917 is the major adhesin mediating binding to human mucus. *Int. J. Med. Microbiol.* 302, 304–314. doi: 10.1016/j.ijmm.2012.09.004
- Tsai, J. C., Yen, M. R., Castillo, R., Leyton, D. L., Henderson, I. R., and Saier, M. H. Jr. (2010). The bacterial intimins and invasins: a large and novel family of secreted proteins. *PLoS ONE* 5:e14403. doi: 10.1371/journal.pone.0014403
- Tunio, S. A., Oldfield, N. J., Ala'aldin, D. A., Wooldridge, K. G., and Turner, D. P. (2010). The role of glyceraldehyde 3-phosphate dehydrogenase (GapA-1) in *Neisseria meningitidis* adherence to human cells. *BMC Microbiol.* 10:280. doi: 10.1186/1471-2180-10-280
- Uchiyama, T., Kawano, H., and Kusuha, Y. (2006). The major outer membrane protein rOmpB of spotted fever group rickettsiae functions in the rickettsial adherence to and invasion of Vero cells. *Microbes Infect.* 8, 801–809. doi: 10.1016/j.micinf.2005.10.003
- Valle, J., Latasa, C., Gil, C., Toledo-Arana, A., Solano, C., Penades, J. R., et al. (2012). Bap, a biofilm matrix protein of *Staphylococcus aureus* prevents cellular internalization through binding to GP96 host receptor. *PLoS Pathog.* 8:e1002843. doi: 10.1371/journal.ppat.1002843
- Valle, J., Mabbett, A. N., Ulett, G. C., Toledo-Arana, A., Wecker, K., Totsika, M., et al. (2008). UpaG, a new member of the trimeric autotransporter family of adhesins in uropathogenic *Escherichia coli*. *J. Bacteriol.* 190, 4147–4161. doi: 10.1128/JB.00122-08
- Van Der Woude, A. D., Mahendran, K. R., Ummels, R., Piersma, S. R., Pham, T. V., Jimenez, C. R., et al. (2013). Differential detergent extraction of *Mycobacterium marinum* cell envelope proteins identifies an extensively modified threonine-rich outer membrane protein with channel activity. *J. Bacteriol.* 195, 2050–2059. doi: 10.1128/JB.02236-12
- Van Der Woude, M. W., and Henderson, I. R. (2008). Regulation and function of Ag43 (flu). *Annu. Rev. Microbiol.* 62, 153–169. doi: 10.1146/annurev.micro.62.081307.162938
- Van Dijk, J. M., Bolhuis, A., Tjalsma, H., Jongbloed, J. D., De Jong, A., and Bron, S. (2001). “Protein transport pathways in *Bacillus subtilis*: a genome-based road map,” in *Bacillus Subtilis and Its Closest Relatives: from Genes to Cells*, eds A. L. Sonenshein, J. A. Hoch, and R. Losick. (Washington, DC: ASM press), 337–355.
- Van Ulsen, P., Adler, B., Fassler, P., Gilbert, M., Van Schilfgaarde, M., Van Der Ley, P., et al. (2006). A novel phase-variable autotransporter serine protease, AusI, of *Neisseria meningitidis*. *Microbes Infect.* 8, 2088–2097. doi: 10.1016/j.micinf.2006.03.007
- Van Ulsen, P., Van Alphen, L., Ten Hove, J., Fransens, F., Van Der Ley, P., and Tommassen, J. (2003). A Neisserial autotransporter NalP modulating the processing of other autotransporters. *Mol. Microbiol.* 50, 1017–1030. doi: 10.1046/j.1365-2958.2003.03773.x
- Varga, J. J., Nguyen, V., O'Brien, D. K., Rodgers, K., Walker, R. A., and Melville, S. B. (2006). Type 4 pili-dependent gliding motility in the Gram-positive pathogen *Clostridium perfringens* and other *Clostridia*. *Mol. Microbiol.* 62, 680–694.
- Varga, J. J., Therit, B., and Melville, S. B. (2008). Type 4 pili and the CcpA protein are needed for maximal biofilm formation by the Gram-positive anaerobic pathogen *Clostridium perfringens*. *Infect. Immun.* 76, 4944–4951. doi: 10.1128/IAI.00692-08
- Vatanyooaisarn, S., Nazli, A., Dodd, C. E., Rees, C. E., and Waites, W. M. (2000). Effect of flagella on initial attachment of *Listeria monocytogenes* to stainless steel. *Appl. Environ. Microbiol.* 66, 860–863. doi: 10.1128/AEM.66.2.860-863.2000
- Vengadesan, K., and Narayana, S. V. (2011). Structural biology of Gram-positive bacterial adhesins. *Protein Sci* 20, 759–772. doi: 10.1002/pro.613
- Verstraeten, N., Braeken, K., Debkumari, B., Fauvart, M., Franssaer, J., Vermant, J., et al. (2008). Living on a surface: swarming and biofilm formation. *Trends Microbiol.* 16, 496–506. doi: 10.1016/j.tim.2008.07.004
- Vlamakis, H., Aguilar, C., Losick, R., and Kolter, R. (2008). Control of cell fate by the formation of an architecturally complex bacterial community. *Genes Dev.* 22, 945–953. doi: 10.1101/gad.1645008
- Vlamakis, H., Chai, Y., Beauregard, P., Losick, R., and Kolter, R. (2013). Sticking together: building a biofilm the *Bacillus subtilis* way. *Nat. Rev. Microbiol.* 11, 157–168. doi: 10.1038/nrmicro2960
- Voigt, B., Hieu, C. X., Hempel, K., Becher, D., Schluter, R., Teeling, H., et al. (2012). Cell surface proteome of the marine planctomycete *Rhodopirellula baltica*. *Proteomics* 12, 1781–1791. doi: 10.1002/pmic.201100512
- Voth, D. E., Broederdorf, L. J., and Graham, J. G. (2012). Bacterial Type IV secretion systems: versatile virulence machines. *Future Microbiol.* 7, 241–257. doi: 10.2217/fmb.11.150
- Voulhoux, R., Ball, G., Ize, B., Vasil, M. L., Lazdunski, A., Wu, L. F., et al. (2001). Involvement of the twin-arginine translocation system in protein secretion via the Type II pathway. *EMBO J.* 20, 6735–6741. doi: 10.1093/emboj/20.23.6735
- Vu, B., Chen, M., Crawford, R. J., and Ivanova, E. P. (2009). Bacterial extracellular polysaccharides involved in biofilm formation. *Molecules* 14, 2535–2554. doi: 10.3390/molecules14072535
- Wagner, C., Polke, M., Gerlach, R. G., Linke, D., Stierhof, Y. D., Schwarz, H., et al. (2011). Functional dissection of SiiE, a giant non-fimbrial adhesin of *Salmonella enterica*. *Cell Microbiol.* 13, 1286–1301. doi: 10.1111/j.1462-5822.2011.01621.x
- Waksman, G., and Fronzes, R. (2010). Molecular architecture of bacterial Type IV secretion systems. *Trends Biochem. Sci.* 35, 691–698. doi: 10.1016/j.tibs.2010.06.002
- Waksman, G., and Hultgren, S. J. (2009). Structural biology of the chaperone-usher pathway of pilus biogenesis. *Nat. Rev. Microbiol.* 7, 765–774. doi: 10.1038/nrmicro2220
- Wallden, K., Rivera-Calzada, A., and Waksman, G. (2010). Type IV secretion systems: versatility and diversity in function. *Cell Microbiol.* 12, 1203–1212. doi: 10.1111/j.1462-5822.2010.01499.x
- Wehrli, W., Brinkmann, V., Jungblut, P. R., Meyer, T. F., and Szczepek, A. J. (2004). From the inside out—processing of the *Chlamydial* autotransporter PmpD and its role in bacterial adhesion and activation of human host cells. *Mol. Microbiol.* 51, 319–334. doi: 10.1046/j.1365-2958.2003.03838.x
- Wells, T. J., McNeilly, T. N., Totsika, M., Mahajan, A., Gally, D. L., and Schembri, M. A. (2009). The *Escherichia coli* O157:H7 EhaB autotransporter protein binds to laminin and collagen I and induces a serum IgA response in O157:H7 challenged cattle. *Environ. Microbiol.* 11, 1803–1814. doi: 10.1111/j.1462-2920.2009.01905.x
- Wells, T. J., Srocklock, O., Rivas, L., Mahajan, A., Beatson, S. A., Torpdahl, M., et al. (2008). EhaA is a novel autotransporter protein of enterohemorrhagic *Escherichia coli* O157:H7 that contributes to adhesion and biofilm formation. *Environ. Microbiol.* 10, 589–604. doi: 10.1111/j.1462-2920.2007.01479.x
- Wong, E., Vaaje-Kolstad, G., Ghosh, A., Hurtado-Guerrero, R., Konarev, P. V., Ibrahim, A. F., et al. (2012). The *Vibrio cholerae* colonization factor GbpA possesses a modular structure that governs binding to different host surfaces. *PLoS Pathog.* 8:e1002373. doi: 10.1371/journal.ppat.1002373
- Wong, K. K., McClelland, M., Stillwell, L. C., Sisk, E. C., Thurston, S. J., and Saffer, J. D. (1998). Identification and sequence analysis of a 27-kilobase chromosomal fragment containing a *Salmonella* pathogenicity island located at 92 minutes on the chromosome map of *Salmonella enterica* serovar typhimurium LT2. *Infect. Immun.* 66, 3365–3371.
- Wurpel, D. J., Beatson, S. A., Totsika, M., Petty, N. K., and Schembri, M. A. (2013). Chaperone-usher fimbriae of *Escherichia coli*. *PLoS ONE* 8:e52835. doi: 10.1371/journal.pone.0052835
- Xu, Q., Barak, Y., Kenig, R., Shoham, Y., Bayer, E. A., and Lamed, R. (2004). A novel *Acetivibrio cellulolyticus*

- anchoring scaffoldin that bears divergent cohesins. *J. Bacteriol.* 186, 5782–5789. doi: 10.1128/JB.186.17.5782-5789.2004
- Xu, Q., Gao, W., Ding, S. Y., Kenig, R., Shoham, Y., Bayer, E. A., et al. (2003). The cellulosome system of *Acetivibrio cellulolyticus* includes a novel type of adaptor protein and a cell surface anchoring protein. *J. Bacteriol.* 185, 4548–4557. doi: 10.1128/JB.185.15.4548-4557.2003
- Yakovenko, O., Sharma, S., Forero, M., Tchesnokova, V., Arikian, P., Kidd, B., et al. (2008). FimH forms catch bonds that are enhanced by mechanical force due to allosteric regulation. *J. Biol. Chem.* 283, 11596–11605. doi: 10.1074/jbc.M707815200
- Yang, H. S., and Hooper, J. K. (1998). Molecular cloning of a light-inducible gene (*lipA*) encoding a novel pilin from *Arthrobacter photogonimos*. *FEMS Microbiol. Lett.* 163, 85–89. doi: 10.1111/j.1574-6968.1998.tb13030.x
- Yang, H. S., and Hooper, J. K. (1999). Regulation of *lipA* gene expression by cell surface proteins in *Arthrobacter photogonimos*. *Curr. Microbiol.* 38, 92–95. doi: 10.1007/s002849900409
- Zambrano, M. M., and Kolter, R. (2005). Mycobacterial biofilms: a greasy way to hold it together. *Cell* 123, 762–764. doi: 10.1016/j.cell.2005.11.011
- Zavialov, A., Zav'yalova, G., Korpela, T., and Zav'yalov, V. (2007). FGL chaperone-assembled fimbrial polyadhesins: anti-immune armament of Gram-negative bacterial pathogens. *FEMS Microbiol. Rev.* 31, 478–514. doi: 10.1111/j.1574-6976.2007.00075.x
- Zav'yalov, V., Zavialov, A., Zav'yalova, G., and Korpela, T. (2010). Adhesive organelles of Gram-negative pathogens assembled with the classical chaperone/usher machinery: structure and function from a clinical standpoint. *FEMS Microbiol. Rev.* 34, 317–378. doi: 10.1111/j.1574-6976.2009.00201.x
- Zechner, E. L., Lang, S., and Schildbach, J. F. (2012). Assembly and mechanisms of bacterial Type IV secretion machines. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 1073–1087. doi: 10.1098/rstb.2011.0207
- Zhang, W., Huhe, Pan, Y., Toyofuku, M., Nomura, N., Nakajima, T., and Uchiyama, H. (2013). Dechlorination of chloral hydrate is influenced by the biofilm adhesin protein LapA in *Pseudomonas putida* LF54. *Appl. Environ. Microbiol.* 79, 4166–4169. doi: 10.1128/AEM.00804-13
- Zhao, K., Tseng, B. S., Beckerman, B., Jin, F., Gibiansky, M. L., Harrison, J. J., et al. (2013). Psl trails guide exploration and microcolony formation in *Pseudomonas aeruginosa* biofilms. *Nature* 497, 388–391. doi: 10.1038/nature12155
- Zhou, J. M., and Chai, J. (2008). Plant pathogenic bacterial Type III effectors subdue host responses. *Curr. Opin. Microbiol.* 11, 179–185. doi: 10.1016/j.mib.2008.02.004
- that could be construed as a potential conflict of interest.

Received: 18 July 2013; accepted: 22 September 2013; published online: 14 October 2013.

Citation: Chagnot C, Zorgani MA, Astruc T and Desvaux M (2013) Proteinaceous determinants of surface colonization in bacteria: bacterial adhesion and biofilm formation from a protein secretion perspective. *Front. Microbiol.* 4:303. doi: 10.3389/fmicb.2013.00303

This article was submitted to *Microbial Physiology and Metabolism*, a section of the journal *Frontiers in Microbiology*. Copyright © 2013 Chagnot, Zorgani, Astruc and Desvaux. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Mohamed Amine ZORGANI



Caractérisation des ARN régulateurs chez *Streptococcus agalactiae*

Résumé

Streptococcus agalactiae, appelé aussi Group B *Streptococcus* (GBS), est une bactérie commensale du tractus digestif et génital de diverses espèces animales dont l'espèce humaine. Elle représente la première cause d'infections néonatales et est aussi un pathogène émergent chez l'adulte immunodéprimé. L'objectif de ma thèse est la caractérisation fonctionnelle et mécanistique des ARNrég. J'ai étudié plus particulièrement l'ARNrég CetR (pour «cell-envelope-targeting RNA»). Il module la résistance au peptide antimicrobiens (PAM) et la virulence à travers la régulation post-transcriptionnelle de l'ARNm *dltD* codant une protéine de biosynthèse de l'acide D-alanyl-lipotéichoïque. La délétion de *cetR* induit des changements dans la morphologie cellulaire, une diminution de la formation du biofilm et de la résistance aux PAM. Une zone d'interaction, CetR-*dltD*, de 27 nucléotides a été prédite *in silico*. Des mutations compensatoires chez GBS montrent que CetR interagit directement avec l'ARNm *dltD* et que la perturbation de la zone d'appariement est suffisante pour observer les phénotypes associés à CetR. La quantification des niveaux d'ARNm et de la protéine DltD nous a permis de montrer que CetR active la traduction de *dltD* et que la perturbation du duplex CetR-*dltD* induit une diminution spectaculaire de la protéine DltD. De plus, en utilisant un modèle murin d'infection et en quantifiant la survie des bactéries dans les macrophages, nous avons montré que CetR et DltD sont cruciaux pour la virulence de GBS. Enfin, une approche protéomique globale nous a permis de montrer que CetR joue un rôle important dans l'expression des protéines dites «*moonlighting*» et de certains facteurs de virulence potentiels. Cet ARNrég peut jouer un rôle important dans la capacité de *S. agalactiae* à s'établir dans son biotope et à exprimer ses facteurs de virulence. Enfin, les résultats de ces recherches sont des prérequis au développement de stratégies permettant de réduire le risque des infections néonatales dues à *S. agalactiae*.

Mots clés : ARN régulateurs, *Streptococcus agalactiae*, opéron *dlt*, virulence, protéome global

Characterization of regulatory RNAs in *Streptococcus agalactiae*

Abstract

The opportunistic pathogen group B *Streptococcus* (GBS) is the leading cause of neonatal infections. The aim of this work is the characterization of a 680 nt-long regulatory RNA, CetR (cell-envelope-targeting RNA). It modulates antimicrobial peptides (AMPs) resistance and virulence through posttranscriptional regulation of *dltD* mRNA which encodes a D-alanyl-lipoteichoic acid biosynthesis protein. Deletion of *cetR* leads to cell morphology changes, reduced biofilm formation and AMPs resistance. A 27 nt-long CetR-*dltD* interacting region is predicted *in silico*. Compensatory base pair exchanges in GBS demonstrate that CetR interacts directly with *dltD* mRNA and that disruption of this RNA pairing is sufficient to observe the CetR-associated phenotypes. By quantifying both mRNA and protein, we demonstrate that CetR enhances *dltD* translation and disruption of the CetR/*dltD* mRNA interaction results in a dramatic decrease in DltD protein. Moreover, using an infection murine model and quantifying bacterial survival in macrophages, we observe that both CetR and DltD are crucial for GBS virulence. Finally, we highlight CetR pleiotropic role in the expression of several moonlighting proteins and potential virulence factors. This regulatory RNA may play an important role in the ability of GBS to settle in its biotope and express its virulence factors.

Keywords: Regulatory RNA, *Streptococcus agalactiae*, *dlt* operon, virulence, global proteome