



**HAL**  
open science

# Identification des cibles primaires des ARN non codant de *Staphylococcus aureus* et de leurs réseaux de régulation : mise au point des approches MAPS et Grad-seq

Arnaud Tomasini

► **To cite this version:**

Arnaud Tomasini. Identification des cibles primaires des ARN non codant de *Staphylococcus aureus* et de leurs réseaux de régulation : mise au point des approches MAPS et Grad-seq. Microbiologie et Parasitologie. Université de Strasbourg, 2016. Français. NNT : 2016STRAJ046 . tel-01599266

**HAL Id: tel-01599266**

**<https://theses.hal.science/tel-01599266v1>**

Submitted on 2 Oct 2017

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

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

**ÉCOLE DOCTORALE 414 Vie et Santé**

**UPR 9002 – Architecture et Réactivité de l'ARN**

**THÈSE** présentée par :

**Arnaud Tomasini**

soutenue le : **16 Septembre 2016**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : **Aspects moléculaires et cellulaires de la biologie**

**Identification des cibles primaires  
des ARN non codant de *Staphylococcus aureus*  
et de leurs réseaux de régulation :  
mise au point des approches MAPS et Grad-seq**

**THÈSE dirigée par :**

**Mme Pascale Romby**

Docteur, UPR 9002, Université de Strasbourg

**RAPPORTEURS :**

**M. Fabien Darfeuille**

**M. Harald Putzer**

Docteur, INSERM U 212 - Bordeaux

Docteur, UMR 8261 – IBPC, Paris

**EXAMINATEUR :**

**M. Gilles Prevost**

Professeur, Faculté de Médecine, Strasbourg

« LES BACTERIES REPRESENTENT L'UNE DES PLUS GRANDES REUSSITES DANS L'HISTOIRE DU MONDE. ELLES SONT DE NOS JOURS, ET ONT TOUJOURS ETE, LES ORGANISMES LES PLUS REPANDUS SUR LA TERRE ; ELLES NE POURRAIENT PAS ETRE ANEANTIES PAR UNE CATASTROPHE NUCLEAIRE ET ELLES NOUS SURVIVRONT A NOUS TOUS. CETTE EPOQUE EST LA LEUR, ET NOUS NE VIVONS NULLEMENT A « L'ERE DES MAMMIFERES » COMME LE PROCLAMENT NOS MANUELS, DE FAÇON CHAUVINE. MAIS LE PRIX QU'ELLES PAIENT POUR CETTE REUSSITE EST DE RESTER CONFINÉES INDEFINIMENT AU NIVEAU D'UN MICROMONDE ; ET ELLES NE CONNAITRONT JAMAIS LES JOIES ET LES PEINES DE LA CONSCIENCE. NOUS VIVONS DANS UN UNIVERS OU IL FAUT CHOISIR : COMPLEXITE ET PERSISTANCE NE VONT PAS BIEN ENSEMBLE. »

- STEPHEN JAY GOULD

# Remerciements

En premier lieu je tiens simplement à remercier aux membres du jury, messieurs Darfeuille, Putzer et Prévost, d'avoir accepté d'évaluer mon travail de thèse.

Merci au Pf. Eric Westhof de m'avoir accueilli au sein de l'unité UPR 9002 pendant ces quatre dernières années. Merci à l'Université de Strasbourg et à la Fondation pour la Recherche Médicale qui ont financé mon doctorat.

La thèse est une aventure autant scientifique qu'humaine et je souhaite donc remercier les personnes qui m'auront accompagné pendant ces quatre années. Pascale, merci de m'avoir accepté dans ton équipe, pour tes conseils, ta disponibilité, ton refus perpétuel de me laisser déprécier mon travail, ta compréhension, et ta patience, surtout, parce qu'il a dû t'en falloir. Tu m'as toujours fait confiance et poussé à aller de l'avant et je t'en suis extrêmement reconnaissant.

Isabelle, il s'en sera passé du temps depuis que je frimais dans tes TPs avec la « Gram Team ». Merci pour tout le temps que tu m'as consacré, ce que tu m'as appris, ton optimisme sans faille, tes yeux de la foi, ton extrême patience (oui, c'est un mot qui revient souvent)... et tous ces petits gateaux ! Je remercie également les autres membres de l'équipe : Stefano pour ta disponibilité et ta curiosité intarissable, les conseils et discussions, et tout ce que tu m'as appris à dire en italien (Ai bien dormito ?) qui font que je suis parfaitement bilingue aujourd'hui. Anne-Catherine, merci pour tous les conseils techniques que tu m'as prodigués, j'ai beaucoup appris de toi. Merci à toutes les collègues de galère : Mélodie et son dynamisme qui ne meurt jamais, Alessandra (j'attends toujours ce tiramisu !), les Delphine, mais encore Mejante Marie, Cédric (au final tu auras été le seul garçon à m'épauler en quatre ans !), et tous ceux qui sont passés plus ou moins longtemps chez nous : Maryam, le beau Carlos, et une spéciale pour Zong-Fu : on l'a eu ce RsaC et sa structure secondaire de dingue, on l'a eu, et c'est grâce à toi !

Je remercie également toutes les personnes avec qui j'ai pu collaborer au cours de ces quatre années de travail, notamment nos collègues de Lyon (Karen, François, Tom, Jessica, Claire) et de Wurzburg (Joerg, Sasha, même si maintenant tu es bien plus proche, et merci de m'avoir fait un peu découvrir la ville !).

Remerciements encore aux membres de l'UPR 9002 avec qui j'ai partagé mon quotidien, surtout ceux du 4<sup>ème</sup> (mais pas que) : vous formez une sorte de grosse famille avec laquelle il fait bon travailler. On ne retrouve pas ce genre d'ambiance partout, et ça fait vraiment du bien d'avoir été ici. Merci en particulier aux trublions que sont JC, Guillaume, Philippe Wolf (et mes excuses à vos mamans), Roland, Redmond (my favourite old irish pervert), Semih, Benoît, Marie, Delphine, Romain, Elo, Camille (compagnons d'infortune !), ceux qui sont partis comme Santiago (je me rappellerai

toujours de ton El Spiritu d'Ecuador...) et les nouveaux arrivants comme Damien, Noé, les Emma : c'est vous la relève ! Je remercie également Philippe, Joanna et Laurianne de la plate-forme protéomique qui ont une importance non négligeable dans les résultats que j'ai pu obtenir et avec qui ça a toujours été un plaisir d'expérimenter de nouveaux protocoles pour identifier toutes ces protéines ! Si j'ai oublié de citer quelqu'un, ne le prenez pas personnellement, je m'en excuse mais vous êtes certainement dans mes pensées. Je voudrais aussi remercier toute l'équipe enseignante du Master de BMO, sans qui je n'aurais pu accomplir ce chemin, et particulièrement M. Kammerer, pour m'avoir poussé à donner le meilleur de moi-même en toutes circonstances. Merci aussi à Pascaline avec qui j'ai pris énormément de plaisir à enseigner la biologie aux L1 (j'ai aussi un pensée pour mes étudiants préférés Jeremy, Diane, Gwen, Thomas), ça a été une expérience géniale.

En dehors du laboratoire, maintenant. Je ne serai pas là où j'en suis sans le soutien de mes amis et je profite de cet espace pour les remercier. Les éternels de la fac d'abord, on aura grandi en tant qu'universitaires, et c'est un bonheur de vous voir encore maintenant : Gromilio (même si tu n'assistes pas à ma soutenance, traître), Lionel, Honorable Annie, Maman Vani, Marina, Chassardino, Jordane, Patrick Poncelet, Bol de Vy, Pauline, Françoué... Dédicace spéciale à mon trio de geeks implacables, Papa Raph, Jérôme et Nico Le Clubber, pour le soutien, les soirées GoT, la bière, avoir fait la folie de vivre avec moi (Raph, tu veux pas tenter ?). Merci également aux amis d'enfance, JiBé, Julien et Batou, j'en ai surmonté des épreuves grâce à vous.

Je ne peux pas ne pas faire un paragraphe pour mes amis et collègues de DC Planet, avec qui je partage une sacrée aventure depuis trois ans : JUL, Harley, Nath', Corentin, Zepp', n00dle, Damien, Louis, et tout le reste de l'équipe (hélas trop nombreuse pour que je puisse citer tout le monde), et toutes les personnes que j'ai pu rencontrer ces dernières années avec des discussions et échanges qui, s'ils ne sont pas de l'ordre du scientifique, ont été tout aussi enrichissants : Bruno, Xavier, Katchoo, Sullivan, tous ces auteurs et artistes dont j'ai croisé le chemin : merci à vous.

En dernier lieu je tiens à remercier vivement ma famille. Je me rends compte à quel point vous êtes importants, et ne pourrai jamais assez vous remercier d'avoir été présent quand j'en ai eu besoin. Merci Papa et Maman pour votre soutien, votre confiance et votre amour, merci à mon grand-frère Pierre-Louis (et Pauline, la nouvelle belle-sœur !), et Claire-JULIA, ma jumelle d'exception.

Merci Jeanne, pour tout.

Et parce que je l'avais promis : merci Batman !

# Sommaire

Remerciements.....	2
Sommaire .....	4
Index des figures .....	6
Abréviations.....	8

## INTRODUCTION

Partie 1 - Connaissez-vous le Staphylocoque doré ? .....	11
1. Des amas en grappe de raisin .....	11
2. Des Staphylocoques et des hommes, entre infections et résistance aux antibiotiques .....	15
3. Un génome très plastique .....	20
Partie 2 - Le pouvoir pathogène : .....	27
une histoire de facteurs de virulence.....	27
1. Colonisation et adhésion.....	27
2. Toxines et dissémination.....	29
c. Lutter contre le système immunitaire.....	34
Partie 3 : La régulation de l'expression génétique chez <i>S. aureus</i> .....	39
1. Un grand nombre de protéines... ..	39
2. ... et d'ARN aussi .....	46

## OBJECTIFS

1. Problématique.....	59
2. Démarche expérimentale .....	64

## RESULTATS

Partie 1 – Validation de MgrA comme cible primaire de RsaA.....	67
1. Validation <i>in vivo</i> de la régulation de MgrA par RsaA par diverses approches différentielles de spectrométrie de masse.....	67
2. Validation <i>in vivo</i> de l'importance des deux sites S1 et S2 de RsaA pour la régulation de MgrA .....	70
Partie 2 – Application de l'approche MAPS à RsaA : à la recherche des autres ARNm cibles .....	90
Partie 3 : Recherche et caractérisation .....	130
de la fonction de RsaC.....	130
1. La taille de RsaC varie en fonction des souches à cause de régions répétées.....	130
2. Détermination du démarrage de transcription de RsaC .....	133
3. Détermination de la structure secondaire de RsaC.....	136
4. La stabilité de RsaC est dépendante de la RNase III .....	139
5. Construction du mutant RsaC et protéomique comparative .....	142
6. Application du MAPS à RsaC .....	146
7. Implication de RsaC dans la résistance au stress oxydatif.....	153
8. Un lien possible entre RsaC et la persistance bactérienne.....	158
Partie 4 : Mise au point de la technique Grad-Seq.....	161

## DISCUSSION & PERSPECTIVES

Partie 1 : RsaC, un ARN .....	173
aux propriétés intrigantes .....	173
1. Quelle est la fonction des régions répétées ? .....	173
2. L'ARNm SodA est-il une cible directe de RsaC ? .....	175
3. RsaC est-il vraiment impliqué dans la résistance au stress oxydatif ? .....	178
4. RsaC et persistance.....	181
5. RsaC, RNase III et protéines auxiliaires : quelles interactions ? .....	183
Partie 2 : d'un point de vue technique.....	187
1. L'approche MAPS : un outil puissant, mais perfectionnable .....	187
2. L'approche Grad-seq est-elle réellement utile chez <i>S. aureus</i> ? .....	189

## ANNEXES

Matériel et méthodes .....	195
1. Liste des souches et plasmides utilisés.....	195
2. Liste des oligonucléotides utilisés .....	196
3. Méthodes expérimentales .....	197
Bibliographie .....	201
Publications.....	209
Curriculum Vitae .....	228

## Index des figures

Figure 1 : Découverte et présentation du staphylocoque doré .....	14
Figure 2 : <i>S. aureus</i> , un pathogène redoutable .....	16
Figure 3 : actions des antibiotiques sur la synthèse du peptidoglycane .....	18
Figure 4 : exemples d'éléments génétiques mobiles.....	22
Figure 5 : les éléments génétiques mobiles de <i>S. aureus</i> .....	26
Figure 6 : mécanismes d'action des toxines de <i>S. aureus</i> .....	30
Figure 7 : vue d'ensemble des actions des PSMs .....	32
Figure 8 : <i>S. aureus</i> peut lutter contre les deux types de réponse immunitaire de l'être humain.....	37
Tableau 1: Systèmes à deux composants chez <i>S. aureus</i> .....	44
Figure 9 : Rôle du système agr chez <i>Staphylococcus aureus</i> - .....	45
Figure 10 : Rappel sur les ARN régulateurs chez les bactéries.....	60
Figure 11 : Interactions entre ARNnc et protéines et leurs effets.....	62
Figure 12 : Structure secondaire de RsaA et structure prédite de RsaC selon Geissmann et al., 2009 .....	63
Figure 13 : Représentation schématique de l'approche MAPS.....	65
Figure 14 : RsaA inhibe l'expression de MgrA in vivo .....	69
Figure 15 : Les sites S1 et S2 sont importants pour la régulation de MgrA par RsaA in vivo.....	71
Figure 16 : Représentation schématique d'une partie du réseau de régulation de RsaA	91
Figure 17 : Différences de taille pour RsaC en fonction des souches de <i>S. aureus</i> .....	131
Figure 18 : Le locus de RsaC présente plusieurs régions répétées .....	132
Figure 19 : Expression de RsaC dans différentes souches de <i>S. aureus</i> .....	133
Figure 20 : Stratégie de détermination du TSS de RsaC chez <i>S. aureus</i> HG001 (RN1)....	134
Figure 21 : Expérience de primer extension pour déterminer le TSS de RsaC.....	135
Figure 22 : Séquence finale de RsaC chez <i>S. aureus</i> HG001 .....	136
Figure 23 : Application du SHAPE pour déterminer la structure secondaire de RsaC....	138
Figure 24 : structure secondaire de RsaC .....	139
Figure 25 : impact de la RNase III sur la stabilité de RsaC.....	140
Figure 26 : expression de RsaC chez des souches exprimant différentes versions de la RNase III.....	141
Figure 27 : quantification de RsaC et de RsaC anti-sens par QRT-PCR chez RN1 .....	142
Figure 28 : construction du mutant $\Delta$ rsaC chez <i>S. aureus</i> RN1 .....	143
Tableau 3 : analyse par spectrométrie de masse du protéome de RsaC .....	145
Figure 29 : Expression des variants MS2-RsaC chez <i>S. aureus</i> .....	147
Figure 30 : MS2-RsaC est correctement retenu sur la colonne de chromatographie d'affinité.....	147
Tableau 4 : résultats du MAPS réalisé avec MS2-RsaC.....	151
Figure 31 : L'ARNm <i>sodA</i> est enrichi dans l'éluat provenant de l'extrait brut contenant MS2-RsaC.....	151



Figure 32 : Localisation de sodA sur le génome de <i>S. aureus</i> .....	152
Figure 33 : prédiction d'appariement de séquences entre RsaC et l'ARNm sodA.....	153
Figure 34 : Impact du MV sur la croissance et la survie cellulaire.....	154
Figure 35 : Impact du MV sur l'expression de RsaC, sodA et sodM.....	155
Figure 36 : RsaC joue un rôle dans la résistance au stress oxydatif.....	157
Figure 37 : Expression intracellulaire (cellules MG63) de différents ARNnc de <i>S. aureus</i> .....	158
Figure 38 : Tests de cytotoxicité et de persistance cellulaire dans les cellules MG63 (ostéoblastes).....	159
Figure 39 : Validation de l'approche du gradient de glycérol chez <i>S. aureus</i> .....	163
Figure 40 : Le gradient de glycérol permet de suivre les interactions ARN/ARN et ARN/protéines in vivo.....	164
Figure 41 : Vérifications de la reproductibilité de la collecte des gradients de glycérol par AKTA® et de l'intégrité des ARN après centrifugation.....	165
Figure 42 : Vérifications de la reproductibilité des gradients de glycérol par Northern blot.....	166
Figure 43 : Vérifications de la reproductibilité des gradients de glycérol SDS-PAGE et spectrométrie de masse.....	167
Figure 44 : Visualisation de l'abondance d'ARNnc de <i>S. aureus</i> par quantification et normalisation des données de RNA-seq.....	168
Figure 45 : Visualisation de l'abondance d'ARNnc de <i>S. aureus</i> par quantification et normalisation des données de RNA-seq.....	169
Figure 46 : Résultat de BLAST sur la séquence complète de RsaC.....	173
Figure 47 : Représentation schématique des loci de sodA et psaA chez <i>S. aureus</i> HG001. .....	177
Figure 48 : Modèle schématique du rôle possible de RsaC chez <i>S. aureus</i> .....	181
Figure 49 : Représentation schématique du SPLASH.....	192

# Abréviations

## Bases

A : adénine

C : cytosine

G : guanine

T : thymine

U : uracile

## Acides nucléiques

ADN : acide désoxyribonucléique

ADNc : ADN complémentaire

ARN : acide ribonucléique

ARNm : ARN messager

ARNnc : ARN non codant

sARN : *small non coding RNA*

ARNt : ARN de transfert

ARNr : ARN ribosomique

ARNas : ARN antisens

Rsa : RNA of *S. aureus*

RBS : site de fixation du ribosome

SD : séquence Shine-Dalgarno

TSS : site d'initiation de la transcription

5'UTR : région 5' non traduite

3'UTR : région 3' non traduite

## Métabolites/ions

Fe : fer

Mn : manganèse

Zn : zinc

## Génétique

EGM : éléments génétiques mobiles

SNP : *single nucleotide polymorphism*

SCC : *staphylococcal chromosomal cassette*

CCR : cassette chromosomique de recombinaison

Tn : transposon

IS : séquence d'insertion

P3 : promoteur de l'ARNIII

## Espèces

*S. aureus* : *Staphylococcus aureus*

*S. epidermidis* : *Staphylococcus epidermidis*

*S. fleuretti* : *Staphylococcus fleuretti*

*E. coli* : *Escherichia coli*

*S. enterica* : *Salmonella enterica*

*E. faecalis* : *Enterococcus faecalis*

## Pathogénie

SIDA : syndrome d'immunodéficience humaine acquise

SARM : *S. aureus* résistant à la méticilline

SARV : *S. aureus* résistant à la vancomycine

CMI : concentration minimale inhibitrice

PIA : adhésine intercellulaire polysaccharidique

PSM : *phenol-soluble moduline*

TSST : toxine du syndrome du choc toxique

CWA : protéines ancrées à la paroi cellulaire

MSCRAMM : composants de la surface microbienne qui reconnaissent des molécules de la matrice adhésive

MEC : matrice extracellulaire

IgG : immunoglobuline de type G

β-BFT : protéines formant des tonneaux bêta

CPA : cellule présentatrice de l'antigène

IL : interleukine

CHIPS : protéine inhibitrice de chimiotactisme

MAC : complexe d'attaque de la membrane

ROS : espèces réactive de l'oxygène

AIP : peptide autoinducteur

### **Protéines**

PBP : protéine de liaison à la pénicilline  
MBP : protéine de liaison au maltose  
MS2-CP : protéine de capsid du phage MS2  
ORF : cadre ouvert de lecture  
RNase : ribonucléase  
DNase : désoxyribonucléase  
TCS : système à deux composants  
TF : facteur de transcription  
RNA-P : ARN polymérase  
RT : reverse-transcriptase  
PNPase : polynucléotide phosphorylase

### **Divers**

DiGE : analyse différentielle de gel 2D  
pH : potentiel hydrogène  
pI : point isoélectrique  
kDa : kilo-daltons  
nts : nucléotides  
(q)PCR : réaction de polymérisation en chaîne (quantitative)  
BHI : *Brain-Heart Infusion*  
DIG : digoxygénine  
kb : kilo paire de bases  
SHAPE : *selective 2'-hydroxyl acylation analyzed by primer extension*  
FT : *flow-through*  
EL : éluat  
NBT : nitroblue-tetrazolium  
MV : méthyl-viologen

# INTRODUCTION

# Partie 1 - Connaissez-vous le Staphylocoque doré ?

## Avant-propos

La première partie de cette introduction a été volontairement rédigée dans un langage scientifique vulgarisé, de façon à pouvoir être accessible au plus grand nombre de lecteurs. Pour des raisons personnelles, elle a également constitué un exercice de rédaction dans un style que j'ai eu moins l'occasion de pratiquer au cours de ma thèse.

Prenez une personne au hasard de votre entourage. Commencez à lui parler du « Staphylocoque doré » et il y a de grandes chances que cette personne fasse de suite une association d'idées avec la « résistance aux antibiotiques » et les « infections en milieu hospitalier ». L'organisme *Staphylococcus aureus* (de son nom scientifique) est en effet reconnu pour être un redoutable pathogène de l'espèce humaine. Pour se faire une petite idée de sa dangerosité, regardons en premier lieu quelques chiffres. Un rapport de février 2015 de l'INSERM estime que le Staphylocoque doré est tenu responsable à hauteur de 16% des 750 000 cas d'infections nosocomiales répertoriés chaque année en France, infections qui entraînent la mort de 4000 personnes. Selon un autre rapport du Centre pour le Contrôle et la Prévention des Maladies (CDC), cette bactérie cause aux Etats-Unis près de 80 000 infections par an et 11 000 décès.

On comprendra dès lors que la bactérie relève d'un intérêt particulier pour la recherche scientifique, notamment pour les enjeux humains présents. Mais il ne faut pas oublier que *Staphylococcus aureus* est avant tout un micro-organisme, et c'est d'abord sous un angle de microbiologiste que je souhaite la présenter : c'est-à-dire, comme un organisme vivant dont l'existence nous a été totalement inconnue jusqu'au XIXème siècle.

## 1. Des amas en grappe de raisin

Le Staphylocoque doré est une bactérie, c'est-à-dire un organisme unicellulaire invisible à l'œil nu, dont l'étude relève de la microbiologie, une discipline qui à l'échelle de l'Histoire des sciences, est encore toute jeune. Les célèbres animalcules observés par

Antoni van Leeuwenhoek ne sont mentionnés la première fois que dans ses observations de 1673, rapportés à la société royale des Pays-Bas. Il faudra attendre la seconde moitié du XIX<sup>ème</sup> siècle pour que la microbiologie puisse véritablement commencer à exister, lorsque le scientifique Louis Pasteur (1822-1895) dénonce la théorie de la génération spontanée. Cette théorie qui prend ses origines dans les croyances religieuses, veut que les organismes observés, par exemple sur des tissus en décomposition, puissent apparaître spontanément. Louis Pasteur est un fervent opposant de cette théorie et réussira à démontrer en 1864 que les micro-organismes présents sur un tissu ou un milieu, proviennent obligatoirement d'une contamination extérieure. C'est à cette période également que se développe la méthode de stérilisation par la chaleur qui sera ensuite appelée pasteurisation. La microbiologie fait un grand pas en avant alors que l'humanité prend conscience d'être entourée de millions de germes invisibles. A l'échelle sociétale, de grandes avancées en résulteront, avec le développement important de l'hygiène publique et la pratique de la stérilisation dans le milieu alimentaire.

Le développement de la microbiologie permet également de démontrer que des micro-organismes sont responsables de maladies, et ce sont les travaux de Koch (1843-1910) – et son fameux postulat<sup>1</sup> – qui valident sa théorie des germes et qui permettent de lier une maladie infectieuse à un organisme donné (Madigan & Martinko, 2007). Et le Staphylocoque doré, alors ? Sa découverte est communément attribuée au chirurgien écossais Alexander Ogston (1844-1929). Celui-ci observa en 1880 au microscope le pus prélevé d'un abcès de patient et pu y contempler ce qu'il décrivit comme « des touffes, enchevêtrements et chaînes d'organismes ronds, qui se démarquent clairement des cellules et autres débris » (Orenstein, 2010). Ces organismes microscopiques et ronds avaient déjà la dénomination de coques (du grec « kokkos » qui signifie graine ou baie) et en 1882, Alexander Ogston désignera ces organismes par le nom de *Staphylococcus* dont la racine grecque vient du mot « staphyl » qui signifie « grappe de raisin » - un nom

---

<sup>1</sup> Un ensemble de règles écrites pour définir le lien entre un agent infectieux et une maladie. Il se constitue ainsi : l'agent infectieux est présent chez les individus malades et absent des individus sains ; il peut être isolé de l'organisme malade et cultivé *in vitro* ; il est responsable de la maladie lorsqu'il est ré-introduit chez un organisme sain ; il doit ensuite être isolé du nouvel hôte rendu malade et identifié comme identique à l'agent infectieux original.

donné en raison de la forme que lui inspiraient ces groupements de cellules. En 1884, le chirurgien allemand Anton J. Rosenbach (1842-1923) (voir **Figure 1a**) isola deux souches de staphylocoques, et les nomma en fonction de la couleur des colonies qu'elles produisent sur milieu solide : les colonies blanches donnent leur nom à *Staphylococcus albus* (du grec « albus », blanc) et celles d'un jaune doré sont appelées *Staphylococcus aureus* (du grec « aureus », l'or) (Cowan & Shaw, 1954). Par la suite, *Staphylococcus albus* sera renommé *Staphylococcus epidermidis*.

Depuis sa découverte, *S. aureus* est donc fortement associé à des enjeux humains. Mais la bactérie a également servi de modèle d'étude pour la compréhension du fonctionnement des bactéries au fur et à mesure des avancées en microbiologie, et on en retrouve de multiples traces, bien qu'elles ne soient pas excessivement nombreuses, dans la bibliographie. Ainsi au cours des années '60 qui ont été marquées par la célèbre caractérisation de l'opéron lactose chez *E. coli* par Jacob et Monod, on peut voir que la communauté scientifique s'est aussi intéressée à cet opéron chez *S. aureus* et au métabolisme du lactose. Ce dernier présente forcément des différences : par exemple l'IPTG ne peut pas être utilisé comme inducteur de l'opéron car il a une action inhibitrice. A cette époque, on pense même utiliser la  $\beta$ -galactosidase de *S. aureus* comme marqueur taxonomique au vu de ses particularités (Mone *et al.* 1968). Des recherches sur le métabolisme de cette bactérie seront effectuées : des expériences réalisées avec des substrats radioactifs permettent de suivre leur cheminement dans la bactérie et les produits qui sont formés, afin de retrouver quelles sont les voies métaboliques présentes chez la bactérie (Strasters & Winkler, 1963). Le tropisme chimio-organotrophe de la bactérie est ainsi caractérisé, tout comme la répression métabolique qu'exerce le glucose sur les enzymes du métabolisme central, ainsi que l'importance du fer pour sa croissance (Theodore & Schade, 1965). Le Staphylocoque doré sera également utilisé comme modèle d'étude pour le métabolisme des lipides et la composition lipidique des membranes cellulaires car elle présente un motif lipidique simple ; l'importance de l'acide téichoïque dans la membrane cellulaire est explicitée (Koch *et al.*, 1984). Quant au pigment qui donne sa couleur caractéristique à *S. aureus* (voir **Figure 1c**), c'est en 1972 qu'on lui donne le nom de staphyloxanthine, et en 1981 qu'il est biochimiquement caractérisé (Marshall & Wilmoth, 1981).

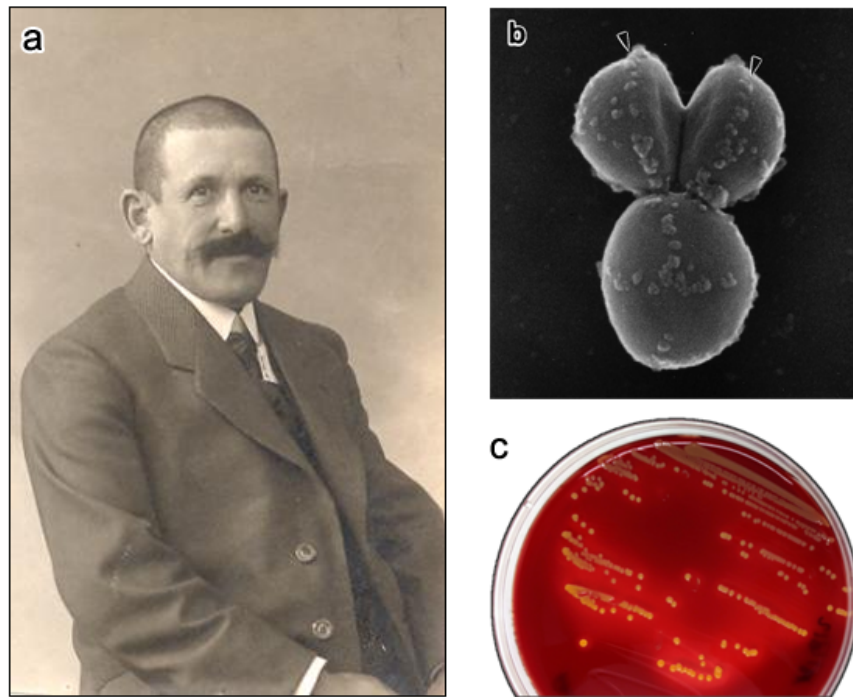


Figure 1 : Découverte et présentation du staphylocoque doré – a : Portrait du chirurgien Anton J. Rosenbach (1842-1923) ; b : cellules de *S. aureus* en division observées en microscopie électronique à balayage (MEB) (Yamada et al., 1996) ; c : colonies de *S. aureus* sur gélose au sang

Les recherches effectuées au cours du dernier siècle nous permettent d'effectuer un portrait plutôt simple de cet organisme : c'est une bactérie à Gram positif à faible pourcentage G-C (dont l'enveloppe cellulaire se constitue d'une épaisse couche de peptidoglycane qui entoure la membrane plasmique) et qui est incapable de produire de spores. Non motile, cette bactérie croît par scissiparité et se regroupe en amas de cellules, dont la forme fait penser à des grappes de raisin. Dotée d'un métabolisme respiratoire aérobie-anaérobie facultatif, elle est capable de pousser dans des milieux à forte concentration en sel et se montre très osmotolérante, même pour une bactérie non-halophile (Graham & Wilkinson, 1992). Cette capacité est d'ailleurs utilisée pour l'isolement sélectif de cette bactérie sur un milieu particulier appelé gélose de Chapman. Si la bactérie est très souvent retrouvée en association avec l'homme et des espèces animales, une certaine partie de la population bactérienne se trouve également dans les sols et milieux extérieurs dans lesquels elle peut survivre grâce à de nombreux mécanismes de résistance aux stress, que ce soient la dessiccation, la disponibilité en nutriments ou celle en métaux (Clements & Foster, 1999).



## 2. Des Staphylocoques et des hommes, entre infections et résistance aux antibiotiques

Retournons maintenant aux liens qui unissent *S. aureus* et les hommes. Le Staphylocoque doré est une bactérie commensale de l'espèce humaine, c'est-à-dire qu'elle se développe sur cet hôte sans que cette relation ne soit néfaste pour ce dernier. A ce titre, on estime que près d'un tiers de la population humaine est porteur sain de *S. aureus*, qu'on retrouve sur la peau et dans les fosses nasales. Le pourcentage total d'être humains colonisés varie selon les études, avec une estimation totale de 25% de porteurs sains (Lindsay, 2014) alors que ce pourcentage est plus important – de 30% à 70% - si on s'intéresse seulement à la présence dans les fosses nasales (Peacock *et al.*, 2001). De fait, les bactéries présentes sur les hôtes colonisés constituent la majeure partie du réservoir infectieux (Lindsay, 2014).

Comment expliquer alors les chiffres que je vous communiquais en tout début d'introduction si *S. aureus* est une espèce commensale de l'homme ? Tout simplement parce que cette bactérie est également un pathogène opportuniste, capable de provoquer de nombreuses maladies chez des hôtes dont le système immunitaire est en proie à des dysfonctionnements (on parle alors de personnes immuno-déprimées), tels que les personnes atteintes du SIDA ou de cancers, ou simplement celles qui sont hospitalisées et sont donc par définition affaiblies au niveau de leur défense immunitaire. Ajoutons à cela les enfants et nourrissons, ainsi que les personnages âgés, dont le système immunitaire n'est soit pas assez développé, soit en perte d'efficacité. L'infection est en général provoquée par une rupture des barrières cutanées ou des muqueuses qui permet à la bactérie de coloniser les tissus adjacents, voire d'intégrer le système de circulation sanguine pour se disséminer au travers de tout l'organisme. Le risque d'infection est également augmenté par la présence de corps étrangers, comme les cathéters utilisés en milieu hospitalier (voir **Figure 2**). En effet, ces derniers sont rapidement recouverts de constituants du sérum comme la fibronectine, ce qui crée un environnement favorable pour l'adhésion puis la colonisation par *S. aureus* (Franklin & Lowy, 1998). La variété des infections provoquées par cette bactérie est liée aux multiples possibilités de sites d'infections sur tout le corps humain, mais également à la

capacité de production d'un grand nombre de facteurs de virulence, qui sera vue plus en détails dans la seconde partie de cette introduction.

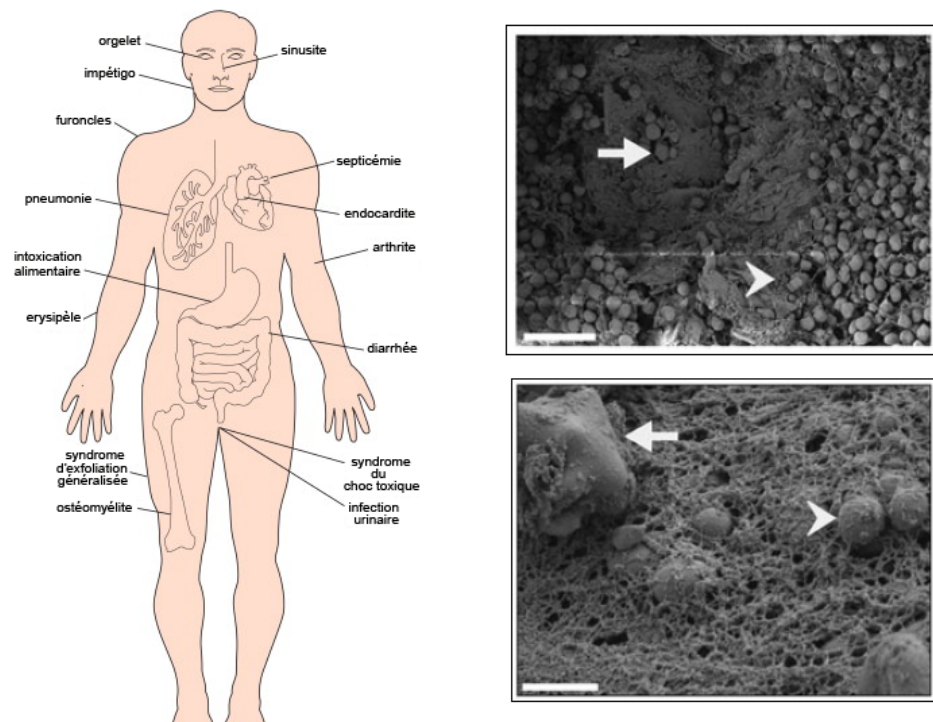
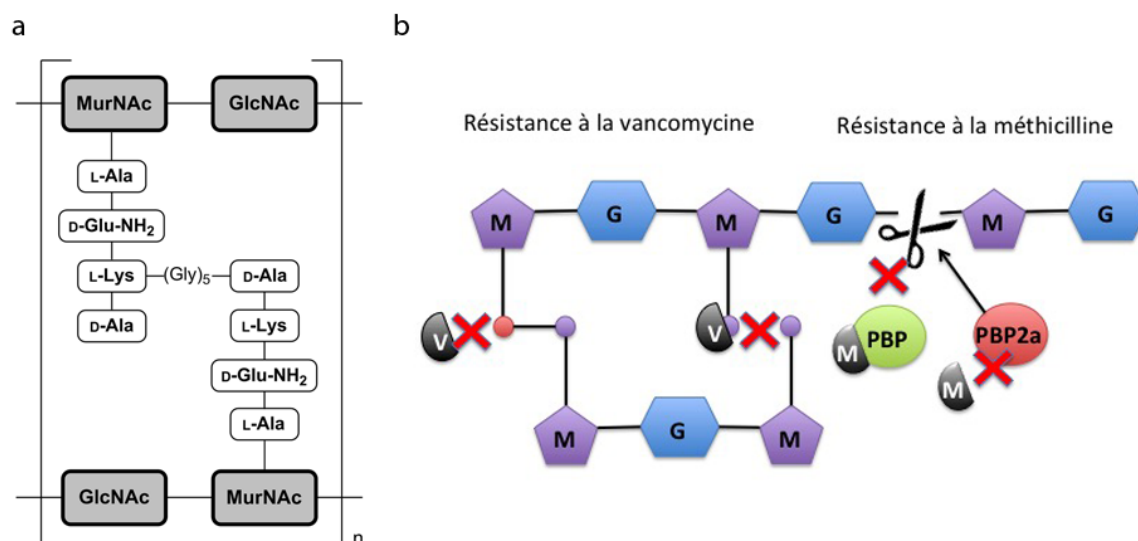


Figure 2 : *S. aureus*, un pathogène redoutable – A gauche : les différents foyers d'infections de *S. aureus* chez l'être humain et maladies associées. A droite : staphylocoques associés en biofilm sur un cathéter, observés au MEB (haut : x 1000 ; bas : x 10 000). Les grandes flèches indiquent des cellules sanguines, les petites flèches en pointe indiquent des staphylocoques au sein du biofilm (Chauhan et al., 2014).

Mais là où le Staphylocoque doré pose également de nombreux problèmes, c'est dans sa capacité à s'adapter et à résister aux antibiotiques. *S. aureus* a d'ailleurs fait partie historiquement de la découverte du premier antibiotique, la pénicilline, en 1928 par Alexander Fleming (1881-1955). Avant cette découverte historique qui a ouvert « l'ère antibiotique », le taux de mortalité suite à une infection par *S. aureus* était de 80%. Après l'introduction des antibiotiques, les infections pouvaient désormais être traitées. Pourtant dès 1942, soit deux ans après l'introduction de la pénicilline dans les usages médicaux, les premières souches résistantes à cet antibiotique ont été observées en milieu hospitalier (Deurenberg & Stobberingh, 2008). La résistance est conférée par une enzyme, la pénicillinase, également appelée  $\beta$ -lactamase, dont l'action est d'hydrolyser le noyau  $\beta$ -lactame de la pénicilline, la rendant ainsi inactive. Face à cette résistance très rapide, une pénicilline résistante aux  $\beta$ -lactamases est introduite en 1959 : la méticilline. C'est la présence d'un groupement ortho-diméthoxyphényl sur la chaîne carbonyle latérale de cet antibiotique qui empêche la liaison avec les pénicillinases et donc

l'hydrolyse de son noyau  $\beta$ -lactame. Mais en 1961 – deux ans seulement, à nouveau – les premières souches de *S. aureus* résistantes à la m $\acute$ eticilline (SARM) sont apparues.

Pour comprendre comment *S. aureus* peut r $\acute$ sister  $\grave$  l'action de la m $\acute$ eticilline, il faut  $\acute$ tudier d'abord son m $\acute$ canisme d'action. Comme je vous l'ai expliqu $\acute$ e auparavant, *S. aureus* est une bact $\acute$ rie  $\grave$  Gram positif, dont l'enveloppe cellulaire est constitu $\acute$ ee d'une  $\acute$ paisse couche de peptidoglycane, un polysaccharide compos $\acute$ e de deux d $\acute$ riv $\acute$ s, le N-ac $\acute$ tylglucosamine (G) et l'acide N-ac $\acute$ tylmuramique (M). Ces compos $\acute$ s forment des feuillet $\acute$ s reli $\acute$ s entre eux par des ponts interpeptidiques, compos $\acute$ s de cinq glycines chez *S. aureus* (la composition du pont interpeptidique varie en fonction des esp $\acute$ ces bact $\acute$ riennes) (Madingan et Martinko, 2007). La couche de peptidoglycane qui entoure la membrane plasmique du Staphylocoque doit  $\acute$ tre rigoureusement imperm $\acute$ eable pour maintenir l'int $\acute$ grit $\acute$ e de la cellule. Mais lors de la division cellulaire, le peptidoglycane doit  $\acute$ tre doubl $\acute$ e, et donc  $\acute$ tre localement et tr $\acute$ s finement cliv $\acute$ e, afin que de nouvelles mol $\acute$ cules de (G) et (M) puissent  $\acute$ tre incorpor $\acute$ ees. Cette action est effectu $\acute$ ee par des prot $\acute$ eines appel $\acute$ ees *penicillin-binding protein* (PBP ayant une activit $\acute$ e de transglycosylase et de transpeptidase). La m $\acute$ eticilline en se liant  $\grave$  cette prot $\acute$ eine bloque l'enzyme, le peptidoglycane ne peut plus  $\acute$ tre cliv $\acute$ e, et la division cellulaire du Staphylocoque est emp $\acute$ ch $\acute$ ee (Deurenberg & Stobberingh, 2008). Les SARM poss $\acute$ dent en revanche une PBP, appel $\acute$ ee PBP2a, sur laquelle la m $\acute$ eticilline ne peut se fixer (**figure 3b**). Cette PBP2a a  $\acute$ te acquise chez *S. aureus* par le g $\acute$ ne *mecA* qui a  $\acute$ te transf $\acute$ er $\acute$ e g $\acute$ n $\acute$ tiquement depuis une esp $\acute$ ce de staphylocoque coagulase-n $\acute$ gative, tr $\acute$ s certainement *S. fleurettii* (Tsubakishita *et al.*, 2010). Lorsque la bact $\acute$ rie poss $\acute$ de cette enzyme, la synth $\acute$ se du peptidoglycane lors de la division cellulaire n' $\acute$ est pas emp $\acute$ ch $\acute$ ee, et la croissance du staphylocoque se poursuit normalement. Si les premiers SARM sont apparus il y a plus de 50 ans, leur diss $\acute$ mination importante au courant des ann $\acute$ es 80-90 fait qu'ils sont pr $\acute$ sents dans 40  $\grave$  60% des h $\acute$ opitaux des pays industrialis $\acute$ s (Lindsay & Holden, 2004).



**Figure 3 : actions des antibiotiques sur la synthèse du peptidoglycane** – **a** : représentation des unités constituant le peptidoglycane. MurNAc (M) : acide N-acétylmuramique. GlcNAc (G) : N-acétylglucosamine. Le pont pentaglycine permet de lier les feuillet de peptidoglycane par les chaînes peptidiques latérales. **b** : représentation schématique du mode d'action de la vancomycine/méthicilline et résistances associées. La vancomycine se lie au résidu D-Ala (violet) et empêche la liaison entre les chaînes peptidiques latérales. Un résidu D-Ala modifié (rouge) apporté par le gène *vanA* contre l'action de la vancomycine. La PBP (verte) clive localement les liaisons entre résidus M et G pour que l'insertion de nouveaux résidus puisse se faire. L'activité de la PBP est inhibée par l'interaction avec la méticilline. La méticilline est incapable d'agir sur une PBP modifiée (PBP2a, rouge) apportée par le gène *mecA*.

De nombreux antibiotiques ont été développés au cours des dernières décennies pour lutter contre le SARM. En 1967, c'est la gentamicine qui est mise sur le marché. Contrairement à la méticilline, la gentamicine fait partie de la famille des aminoglycosides qui ciblent le ribosome de la bactérie, et plus particulièrement l'ARN ribosomique présent dans le site A. Sa présence entraîne des perturbations dans la synthèse des protéines, et l'accumulation de protéines anormales entraîne la mort de la cellule. En 1969, deux années après son introduction, un premier cas de *S. aureus* résistant à la gentamicine est observé (Lacey *et al.*, 1969). Le mécanisme de résistance à la gentamicine résulte de l'action de plusieurs enzymes: la première est une aminoglycoside 6'-N acétyltransférase, la seconde est une aminoglycoside 3'-phosphotransférase. Les deux enzymes ajoutent un groupement à la gentamicine (un acétyle ou un groupement phosphate) pour inactiver l'antibiotique (Dowding, 1976).

Le dernier antibiotique utilisé pour lutter contre le SARM, après sa dissémination importante dans les années '80, est la vancomycine, un antibiotique de la famille des glycopeptides. Introduit sur le marché en 1986, il cible le pont inter-peptidique qui sert à relier les feuillet de peptidoglycane et empêcher la formation de liaisons covalentes entre la D-alanine d'un pont peptidique à la glycine d'un second pont. L'utilisation

massive de cet antibiotique dans les années 80 et 90 va exercer une forte pression de sélection sur le SARM et en 1997 une première souche qui présente un phénotype de résistance intermédiaire (SAIV) est caractérisée (Hiramatsu *et al.*, 1997). Ces souches, que l'on retrouve de plus en plus souvent dans les cas de SARM en milieu hospitalier, se caractérisent par une concentration minimale inhibitrice de croissance (CMI) de 8 µg/ml, soit 4 fois plus importante que celle habituellement retrouvée chez un staphylocoque sensible à la vancomycine. Les premières souches hautement résistantes à la vancomycine (SARV) sont caractérisées quelques années plus tard aux Etats-Unis, en 2002 et 2004. Ces souches possèdent une CMI supérieure à 16 µg/ml et à l'heure actuelle, on a également observé des souches capables de résister à une concentration de 100 µg/ml (Gardete & Tomasz, 2014). A ces concentrations là, le traitement d'un patient avec cet antibiotique n'est plus possible à cause de sa toxicité. Le problème est que la vancomycine était utilisée en dernier recours contre le SARM et il n'existe donc plus de moyen de traitement si l'infection est causée par un SARV. Un constat inquiétant quand on sait que l'industrie pharmaceutique peine grandement à créer de nouveaux antibiotiques, et un signal d'alarme pour pousser cette dernière, en s'associant aux organismes de recherche fondamentale, à trouver d'autres moyens pour lutter efficacement contre les bactéries pathogènes.

Concernant le mécanisme de résistance à la vancomycine, il faut distinguer les cas de SAIV et de SARV. Concernant les souches de SARV, le gène de résistance a été apporté par le transposon Tn1546 qui provient de la bactérie *Enterococcus faecalis*. Ce transposon de 10,8 kb comporte un groupe de sept gènes dont *vanA*, qui code une enzyme qui modifie la cible de la vancomycine : ce sont les groupements D-Ala-D-Ala présents en fin de pont peptidique, qui est modifié en D-Ala-D-Lac (il y a ajout d'un groupement lactate) pour lequel l'antibiotique est bien moins affiné (voir **Figure 3b**) (Périchon & Courvalin, 2009). Pour les SAIV, la résistance est conférée par des mutations sur différents gènes comme *rpoB*, *walkR* ou *vraSR*, ces deux derniers étant directement impliqués dans la synthèse du peptidoglycane (Howden *et al.*, 2014). Des études récentes montrent que le mécanisme d'action des résistances trouvées chez les SAIV, malgré l'hétérogénéité de leurs génotypes, ont un socle commun : c'est l'hypothèse de la fausse cible. En effet il a été constaté que les SAIV présentent une couche de peptidoglycane épaisse et une activité de pontage entre les ponts interpeptidiques

réduite. La baisse de cette activité ainsi que la surproduction de chaînes peptidiques (qui contient les résidus terminaux D-Ala-D-Ala) permettrait d'ajouter ces chaînes en excès dans le peptidoglycane et de leurrer ainsi la vancomycine, qui se liera donc à des peptides situés en dehors des zones de biosynthèse de la paroi cellulaire (Gardete & Tomasz, 2014).

Les nombreuses résistances aux antibiotiques qu'a acquises *S. aureus* depuis l'introduction de la pénicilline sont le fruit de l'évolution du génome de la bactérie suite à l'importante pression de sélection que constituent ces molécules. Mais cette évolution ne s'est pas faite que par un processus de mutation/sélection, et de nombreux éléments génétiques mobiles entrent également en jeu, conférant au génome de *S. aureus* une très grande malléabilité que nous étudions en détail ci-dessous.

### 3. Un génome très plastique

Lorsqu'on parle du génome d'une bactérie, on désigne par là l'ensemble des éléments génétiques qui constituent son patrimoine génétique. Bien entendu, de par la présence des éléments génétiques mobiles (EGM) tels que plasmides, phages et transposons, le génome de deux souches différentes d'une même espèce peut considérablement varier. C'est pour cela qu'on parle de *core* génome pour désigner l'ensemble des éléments génétiques retrouvés chez plus de 95% des souches formant une espèce. Le génome de *S. aureus* a une taille comprise entre 2,820 et 2,903 Mb et des analyses faites *in silico* estiment que son *core* génome comprend 75% du total des gènes. La majorité des gènes inclus dans le *core* génome gouvernent des fonctions essentielles pour le fonctionnement de la cellule – on parle aussi de gènes de ménage – et sont associés également au métabolisme central. On y retrouve aussi quelques gènes spécifiques à l'espèce *S. aureus* comme *spa*, *sodM* ou *icaABDC* qui codent respectivement la protéine A, la superoxyde dismutase permettant la survie dans les neutrophiles ou les adhésines intercellulaire polysaccharidiques (PIA) (Lindsay et Holden, 2004).

Il existe néanmoins une certaine variation génétique dans le *core* génome entre les souches du staphylocoque et un nombre important de mutations ponctuelles (SNP) ont été détectées. Le séquençage de génomes entiers, rendus possibles ces dernières années

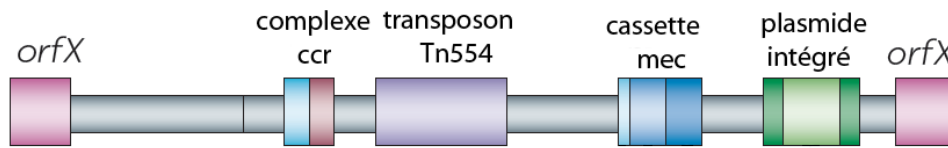
par l'évolution des techniques, a permis d'estimer une fréquence de mutation dans le génome à  $1-3 \cdot 10^{-6}$  substitution par nucléotide et par année, ce qui équivaut à une mutation toutes les 6 à 18 semaines (Lindsay, 2014). Cette « horloge mutationnelle » permet d'expliquer en partie l'évolution des souches de *S. aureus* mais il faut surtout prendre en compte les éléments génétiques mobiles, qui constituent de 15 à 25% du génome. On y trouve parmi ceux-ci les cassettes chromosomiques staphylococcales (SCC), les prophages, les plasmides, les îlots génomiques, les îlots de pathogénicité et les transposons. Ces éléments mobiles (ou mobiles à un temps donné) sont transférés horizontalement entre les souches et participent grandement à la plasticité du génome et à l'acquisition de nouvelles facultés physiologiques.

### 1. Les cassettes chromosomiques staphylococcales

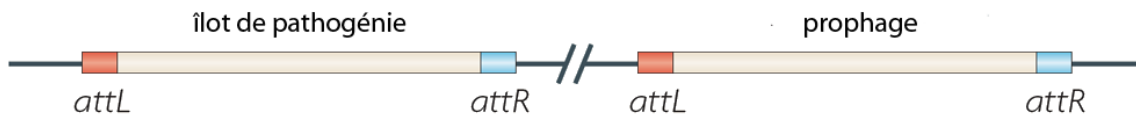
Les SCC sont des éléments génétiques capables de s'intégrer à un seul endroit du génome de *S. aureus*, en aval de l'*orfX* dont le produit de traduction est toujours inconnu. Une SCC comprend un ensemble de gènes qui confèrent la plupart du temps des résistances à des antibiotiques et des métaux lourds comme le mercure (Malachowa & Deleo, 2010), ainsi qu'un duo de gènes qui forment la cassette chromosomique de recombinase (CCR) : il s'agit d'une invertase et d'une résolvasse qui permettent l'intégration et l'excision de la SCC aux sites d'attachement *attB<sub>sc</sub>* situés en 3' de l'*orfX*. L'exemple le plus connu de cassette chromosomique chez *S. aureus* est celle qui confère la résistance à la méticilline, appelée SCC<sub>mec</sub>. Cette dernière comprend le gène *mecA* dont le fonctionnement a été explicité auparavant, mais également d'autres gènes ou éléments mobiles, tels que des prophages ou des transposons. Ainsi, on peut retrouver sur certaines cassettes le plasmide intégré pUB110 qui porte le gène *ant(4)*' permettant la résistance à de nombreux aminoglycosides ; la résistance aux métaux lourds est portée par le plasmide pI258 ; le transposon Tn544 porte le gène *ermA* qui confère la résistance à plusieurs antibiotiques de la famille des macrolides (Deurenberg & Stobberingh, 2008). Ces groupes de gènes forment les complexes *mec* et peuvent donc varier. Ainsi, le type de complexe *mec* tout comme le couple *ccr* permettent d'établir les différents types de cassettes SCC<sub>mec</sub> – dont on a référencé à l'heure actuelle huit grands types (SCC<sub>mec</sub> I-VIII) avec toute une hiérarchie de sous-types. Un exemple de cassette SCC<sub>mec</sub> est schématisé dans la **Figure 4**.

a

Cassette chromosomique staphylococcale



b



**Figure 4 : exemples d'éléments génétiques mobiles** - a : représentation schématique d'une SCC. La cassette *mec* contient le gène de résistance à la méthicilline. Le complexe *ccr* permet l'excision et l'intégration de la SCC au niveau de *orfX*. Les plasmides intégrés et transposons peuvent apporter des gènes supplémentaires (adapté de Chambers & Deleo, 2009). b : représentation schématique d'un îlot de pathogénicité et d'un prophage. Les sites d'intégration permettent une recombinaison entre les deux éléments génétiques (adapté de Novick *et al.*, 2010).

Pour expliquer la grande diversité des cassettes de types *SCCmec*, deux théories s'opposent : celle du clone unique et celles des clones multiples. La première veut que toutes les souches qui possèdent cette cassette aient eu un ancêtre commun qui était sensible à la méthicilline, et qui a acquis la cassette par la suite. Il n'y aurait eu dans ce cas qu'un seul transfert horizontal d'une espèce à l'autre, et c'est à partir d'une seule cassette *SCCmec* que dériveraient toutes celles existantes. La seconde théorie, au contraire stipule que différentes cassettes aient été acquises par plusieurs souches de *S. aureus* au cours de l'évolution. Et c'est cette théorie qui semble la plus probable au vu des preuves scientifiques apportées ces dernières années (Deurenberg et Stobberingh, 2008).

## 2. Les bactériophages

Les bactériophages sont tout simplement des virus qui infectent des bactéries. On distingue généralement deux catégories : les phages lytiques, qui provoquent systématiquement la lyse des cellules infectées, et les phages lysogéniques ou tempérés, qui sont capables de s'intégrer dans le génome de la bactérie – on parle alors de prophage. Certains prophages portent des gènes qui, une fois intégrés dans le génome bactérien, confèrent à la cellule de nouvelles propriétés. Les prophages de *S. aureus* sont classés en cinq grandes familles en fonction des homologues de séquence de leur



intégrase (l'enzyme qui permet, comme son nom l'indique, l'intégration de l'ADN phagique dans le génome bactérien). Une même cellule de *S. aureus* ne peut posséder qu'un seul phage de chaque famille à la fois, lié vraisemblablement à une compétition pour les sites d'insertion, ce qui a des conséquences pour l'évolution du staphylocoque. Parmi les toxines les plus connues portées par des phages, on peut citer la toxine exfoliative A, l'entérotoxine A et la leukocidine de Panton-Valentine, cette dernière étant responsable du syndrome du choc toxique. Les phages participent également au transfert horizontal d'autres gènes et éléments mobiles tels que plasmides, transposons ou îlots de pathogénie, puisqu'une tête de capsidie peut empaqueter jusqu'à 45 kb d'ADN de la cellule hôte à la place du sien, à une fréquence moindre néanmoins (Novick *et al.*, 2010).

### 3. Les îlots de pathogénie et îlots génomiques

Les îlots de pathogénie et îlots génomiques sont des régions du chromosome bactérien flanquées de sites d'intégration qui leur permettent d'être excisé du génome et intégré dans une autre cellule. Les gènes compris dans les îlots de pathogénie sont le plus souvent associés à la pathogénèse : on y trouve des gènes qui codent des super-antigènes, comme la toxine du syndrome du choc toxique (TSST) ou les entérotoxines B et C. Les îlots de pathogénie, dont on connaît au moins 16 représentants, sont classés en quatre grands groupes en fonction de leur intégrase. Comme pour les phages, une même souche n'est capable de posséder qu'un seul îlot par famille (Lindsay & Holden, 2004). Les îlots de pathogénie peuvent être intégrés à six différents sites *att* du génome bactérien (Malachowa & DeLeo, 2010). Ils sont très souvent liés aux prophages, ce qui facilite de façon importante leurs transferts horizontaux, et a également permis de développer une théorie concernant leur origine. Cette théorie suppose qu'il existe un ancêtre commun – très certainement sous la forme d'un prophage – à tous les îlots de pathogénie, qui ont par après acquis leur diversité avec de multiples événements de recombinaison (Yarwood *et al.*, 2002). Les îlots de pathogénie sont transférés d'une cellule à l'autre par le biais d'une encapsidation dans des particules phagiques infectieuses. En absence de phage aidant, l'îlot est alors maintenu dans un état de prophage dormant par un répresseur global, *StI*, qui contrôle l'expression de la plupart des gènes des îlots de pathogénie (Tormo-Mas *et al.*, 2010).

Concernant les îlots génomiques, ils diffèrent surtout par les gènes qu'ils contiennent. On en distingue trois chez *S. aureus*, qui portent l'appellation  $\nu\text{Sa}\alpha$ ,  $\nu\text{Sa}\beta$  et  $\nu\text{Sa}\gamma$ . Certains de leurs gènes sont conservés entre les souches qui les portent, mais la variabilité existe quand même, chacun portant un certain nombre de gènes de lipoprotéines, lipases, voire même de leucocidines (Lindsay & Holden, 2004). L'îlot  $\nu\text{Sa}\gamma$  quant à lui porte le gène d'une « *phenol soluble moduline* » (PSM), un petit peptide sécrété qui joue un rôle important dans la virulence et la formation de biofilms (Malachowa & DeLeo, 2010). La capacité de ces îlots à être transférés horizontalement n'a pas été testée et on pense que leur variabilité est plutôt due à des phénomènes d'intégrations, de recombinaisons, et de pertes de gènes.

#### 4. Les plasmides

Les plasmides sont des molécules d'ADN circulaires auto-répliquatives qui peuvent être échangés entre souches, voire entre différentes espèces, si les conditions le permettent (incompatibilité plasmidique). Chez *S. aureus*, les plasmides retrouvés sont regroupés en trois grandes classes en fonction de leur taille et des gènes qui y sont portés :

- les petits plasmides multi-copies à fonction cryptique ou qui ne portent qu'un seul gène de résistance ;
- les plasmides de taille moyenne (15-30 kpb) à faible nombre de copies (4-6/cellule) qui portent généralement plusieurs marqueurs de résistance ;
- les plasmides conjugatifs de multi-résistance (Malachowa & DeLeo, 2010)

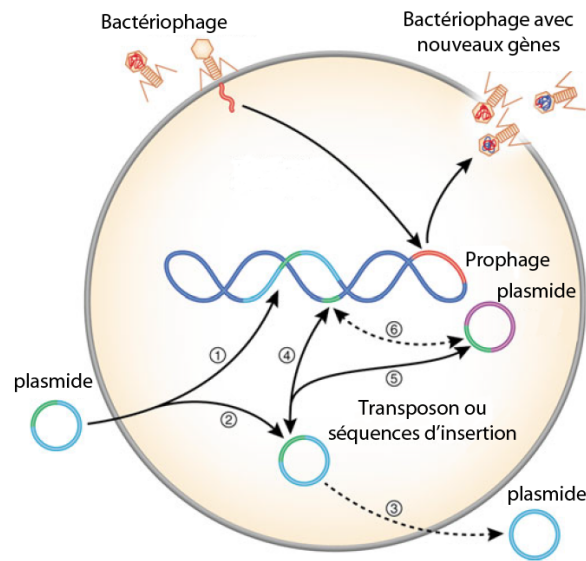
On y retrouve donc des gènes de résistance aux antibiotiques, à des antiseptiques mais aussi à des métaux lourds. Certains gènes de facteurs de virulence, comme celui de la toxine exfoliative B ou de l'entérotoxine D, sont également retrouvés sur des plasmides (Lindsay & Holden, 2004). La majorité des plasmides est acquise par transduction plutôt que par conjugaison, puisque les plasmides conjugatifs ne sont pas courants chez cette espèce bactérienne, et que la transformation ne peut se faire car *S. aureus* n'est pas naturellement compétente. Il existe en revanche des plasmides qui peuvent être transférés horizontalement depuis *E. faecalis* vers *S. aureus* si cette dernière sécrète un petit peptide qui est une phéromone qui active le processus de conjugaison entre les deux bactéries. On soupçonne notamment que l'acquisition du gène *vanA*, porté par

ailleurs sur le transposon Tn1456, puisse se faire par ce genre de plasmide (Showsh *et al.*, 2001). Ce plasmide ne peut pas se répliquer dans le staphylocoque mais sa survie dans la cellule est assez longue pour que le transposon s'intègre ensuite dans le génome (Lindsay & Holden, 2004).

### 5. Les transposons et séquences d'insertion

Ces deux éléments génétiques sont très présents dans le génome de *S. aureus*. Ils peuvent être présents en simple ou multiple copies et être placés sur le génome ou en association avec d'autres EGM. Les séquences d'insertion (IS) sont généralement associées en paires sur le génome, et peuvent s'insérer à de nombreux endroits grâce au gène de la transposase qu'elles codent. De ce fait, elles sont responsables de changements d'expression de gène, soit en s'insérant en plein milieu d'une phase codante, soit dans la région de régulation située à proximité (Malachowa & DeLeo, 2010). A contrario, si l'IS contient un promoteur, elle peut aussi réguler positivement l'expression d'un gène adjacent.

Les transposons sont, comme les IS, des séquences d'ADN capables de s'insérer au hasard dans le génome. Les transposons de *S. aureus* portent généralement des gènes de résistance aux antibiotiques. Les plus petits se retrouvent en plusieurs copies dans le génome et peuvent être associés à d'autres EGM comme des plasmides ou des SCC. Les plus grands (taille jusqu'à 18 kpb) ne sont présents qu'en une seule copie et contiennent des gènes de résistance à la tétracycline, aux aminoglycosides ou encore à la vancomycine (Malachowa & Deleo, 2010).



**Figure 5 : les éléments génétiques mobiles de *S. aureus*** - 1 : incorporation de plasmide ou d'éléments plasmidiques dans le chromosome. 2 : Le plasmide peut se maintenir sous forme libre. 3 : le plasmide peut également ressortir de la cellule. 4 : Transfert d'un transposon ou d'une séquence d'insertion depuis le plasmide vers l'ADN chromosomique. 5 : Transfert d'un transposon ou d'une IS entre plasmides. 6 : Transfert d'un transposon ou d'une IS depuis l'ADN chromosomique vers un plasmide. Adapté de Malachowa & Deleo, 2010.

**Depuis sa découverte il y a plus d'un siècle, le staphylocoque doré s'est affirmé comme l'un des pathogènes les plus problématiques pour l'espèce humaine. La grande plasticité de son génome lui a permis d'évoluer et de s'adapter constamment aux activités humaines, et la dissémination de souches multi-résistantes aux antibiotiques pose plus que jamais un grave problème de santé publique. Cette première partie visait à décrire les moyens de défense et d'adaptation que *S. aureus* possède. La prochaine partie va maintenant s'attarder sur ces capacités d'attaque : les facteurs de virulence, qui font de la bactérie un pathogène redouté.**

## Partie 2 - Le pouvoir pathogène : une histoire de facteurs de virulence

J'énonçais en première partie la capacité de *S. aureus* à provoquer diverses maladies lorsqu'elle réussit à percer les barrières cutanées ou des muqueuses. En effet, la bactérie produit un très grand nombre de facteurs de virulence qui interviennent temporellement à différentes étapes du processus infectieux, que l'on peut diviser ainsi : premièrement l'adhésion à une surface (par exemple, épithéliale) ; dans un second temps la destruction des tissus environnants, la multiplication et la dissémination ; et enfin la protection contre le système immunitaire.

Les gènes des facteurs de virulence produits se trouvent sur le *core* génome mais également sur des EGM, et certains ne se limitent pas qu'à une seule fonction et interviennent donc dans plusieurs processus de l'infection.

### 1. Colonisation et adhésion

De façon récurrente dans la littérature, les protéines du staphylocoque doré sont divisées en deux catégories : les protéines attachées à l'enveloppe cellulaire (CWA pour *cell wall anchored proteins*) et celles qui sont excrétées dans le milieu extracellulaire. Lorsque *S. aureus* profite de la rupture d'une barrière pour se disséminer dans son hôte, elle exprime un certain nombre de protéines CWA qui lui permettront d'adhérer au tissu colonisé, afin de pouvoir commencer à se diviser. On distingue chez cette espèce près de 24 protéines CWA, ce qui est bien plus que d'autres espèces de staphylocoques comme *S. epidermidis* (Foster *et al.*, 2014).

La grande majorité de ces protéines portent l'acronyme de MSCRAMM – qu'on peut traduire assez brutalement par « les composants de la surface microbienne qui reconnaissent des molécules de la matrice adhésive ». Cet acronyme a été donné en premiers lieu pour désigner les protéines de surface de *S. aureus* qui permettent l'adhésion à des composants de la matrice extracellulaire (MEC) – c'est à dire, le fibrinogène, la fibronectine, des polysaccharides, parfois de l'ADN ou encore du

collagène - mais sa définition est aujourd'hui plus large. Cette famille de protéine se caractérise par des domaines IgG (immunoglobulines G) répétés en tandem ; mais si ces protéines sont associées à l'adhésion du staphylocoque au tissu de l'hôte, les études menées ont montré que les rôles peuvent être multiples pour certaines d'entre elles. De même, les gènes de certaines des protéines, notamment la protéine de liaison au fibrinogène, sont présents en plusieurs copies sur le génome (au nombre de cinq, pour la protéine susnommée). Cette redondance génétique peut expliquer premièrement, qu'un mutant d'un gène de protéine CWA n'aura pas un phénotype de mutant complet (c'est-à-dire que la bactérie n'est plus du tout capable d'adhérer aux tissus de l'hôte), et deuxièmement que ces protéines sont difficiles à cibler pour une stratégie thérapeutique (de par leur grand nombre, il est notamment impossible d'envisager d'utiliser des anticorps monoclonaux).

Une autre caractéristique de ces protéines de surface est qu'elles possèdent toutes un signal de sécrétion dépendant du système *sec* dans leur partie N-terminale, et aussi un motif d'adressage - comprenant un motif LPXTG (Leu-Pro-X-Thr-Gly ; le X peut être n'importe quel acide aminé) -, un domaine hydrophobique (pour pouvoir s'ancrer dans le peptidoglycane) ainsi qu'une suite de résidus chargés positivement.

Les représentants majeurs des MSCRAMM sont les protéines ClfA et ClfB (pour *clumping factor* A et B), qui sont capables de se lier au fibrinogène, mais ClfB est également capable de promouvoir l'adhésion au cellules épithéliales desquamées et serait impliquée dans la colonisation des fosses nasales (Mulcahy *et al.*, 2012). Les protéines FnBPA et FnBPB (pour *fibronectin binding protein*) sont impliquées dans l'adhésion à la MEC, tout comme l'adhésine au collagène (Cna) qui, comme son nom l'indique, permet à la bactérie de se lier au collagène (Foster *et al.*, 2014). Notons également que plusieurs composants de l'enveloppe cellulaire ont eux-mêmes une fonction de facteur de virulence. C'est ainsi le cas du peptidoglycane, qui peut être reconnu par le système immunitaire et l'activer, tout comme les acides téichoïques, qui ont une fonction de super-antigène pour provoquer une réaction immunitaire non contrôlée (le mécanisme sera détaillé plus loin).

La production des MSCRAMM est cruciale pour que la bactérie adhère à un support biologique lors de l'infection et pour la formation de biofilm. Le biofilm est une structure multi-cellulaire dans laquelle les bactéries sont enveloppées d'une matrice extra-cellulaire composée de divers polysaccharides, de protéines et parfois d'ADN. Tout en permettant la croissance dans un milieu protégé des attaques du système immunitaire, cette structure promeut également une résistance (par imperméabilité) aux antibiotiques. En même temps que la croissance initiale se produit, les bactéries vont synthétiser une grande gamme d'enzymes et de toxines qui vont lui permettre d'attaquer les tissus environnants, pour récupérer des éléments nutritifs, mais aussi pour se frayer un chemin vers le système circulatoire, afin d'assurer la dissémination de l'organisme.

## 2. Toxines et dissémination

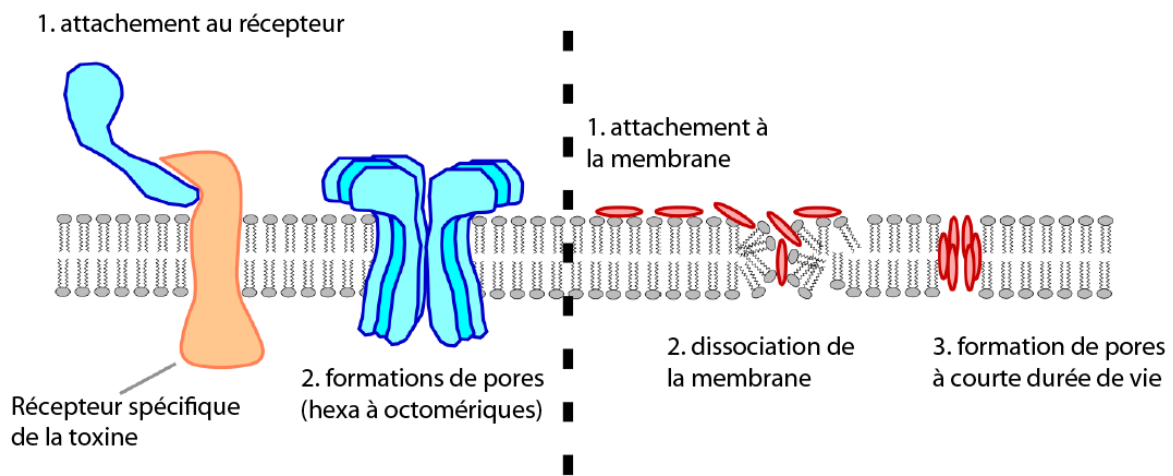
Les toxines de *S. aureus* sont également nombreuses et l'objectif de cette partie n'est pas de les énumérer toutes (voir Otto, 2014 pour une revue récente sur le sujet). Néanmoins nous pouvons les regrouper en différentes catégories en fonction de leur activité.

### a. Cytotoxines et PSM

Ces toxines qui sont excrétées hors du milieu cellulaire ciblent la membrane des cellules environnantes (certaines peuvent reconnaître des récepteurs de façon spécifique) et s'assemblent en multimères pour y former des pores, ce qui conduit à l'extrusion du milieu intracellulaire de la cellule touchée, et en définitive à la lyse cellulaire. Les toxines qui lysent les hématies sont appelées hémolysines, alors que celles qui lysent les leucocytes sont les leucotoxines (Otto, 2014).

Parmi les toxines dont l'action est liée à l'interaction avec un récepteur, l'alpha-toxine (aussi connue sous le nom d'hémolysine alpha) est certainement l'une des plus connues et étudiées. C'est également la première qui a été identifiée comme faisant partie de la famille des  *$\beta$ -barrel forming toxins* ( $\beta$ -BFT) de par sa structure protéique, composée principalement de feuilletts bêta. A de très fortes concentrations, l'alpha-toxine n'a plus besoin de récepteurs pour induire la formation de pores, sous forme heptamérique, mais

à de plus faibles concentrations, elle interagit avec le récepteur ADAM10. Cette interaction avec ce récepteur, qui a une activité de métalloprotéinase, mène à la disruption des adhésions environnantes. Plus particulièrement, le clivage de la cadhérine E dans au niveau des cellules épithéliales mène à la rupture de la barrière épithéliale, ce qui permet la dissémination de la bactérie dans le tissu. L'alpha-toxine est également capable de provoquer la lyse des monocytes, des lymphocytes B et T, mais pas des neutrophiles (Otto, 2014).



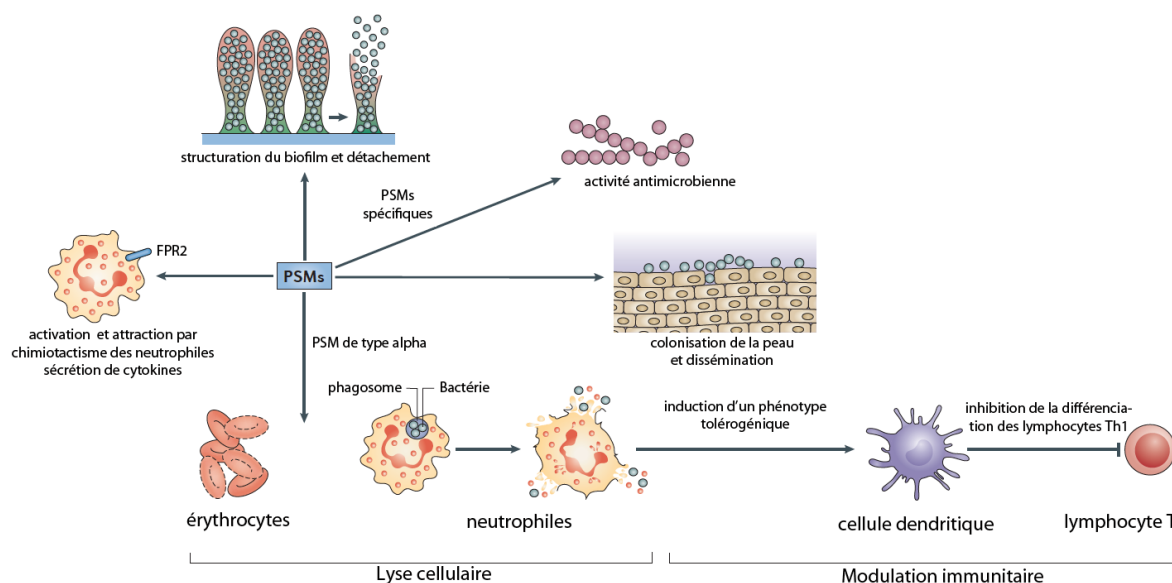
**Figure 6 : mécanismes d'action des toxines de *S. aureus*** – A gauche : mécanismes nécessitant la reconnaissance spécifique d'un récepteur qui concernent les alpha-toxines, les leucocidines bi-partites (LukABn LukDE, PVL) et les hémolysines gamma (pour ces dernières, les récepteurs spécifiques n'ont pas été déterminés). A droite : mécanisme indépendant de récepteurs qui concerne la plupart des PSM $\alpha$  ; la faible durée de vie des pores a été confirmée pour les  $\delta$ -toxines. Adapté de Otto, 2014.

De nombreuses autres toxines font partie de la famille des  $\beta$ -BFT, avec la particularité d'être formées de deux sous-unités. C'est le cas des hémolysines gamma (codées par les gènes *hlgA*, *hlgB* et *hlgC*) ; ces dernières s'assemblent en octamères pour former le pore qui viendra lyser la cellule cible. On peut distinguer deux types d'hémolysine gamma, qui diffèrent en fonction des associations des sous-unités : HglAB et HlgCB. Ces deux protéines diffèrent également de par les récepteurs qu'elles reconnaissent : ce sont les récepteurs à chimiokines CXCR1, CXCR2 et CCR2 pour HglAB, alors que HlgCB reconnaît les récepteurs du complément C5aR et C5L2 (tous ces récepteurs étant impliqués dans la défense immunitaire) (Thammavongsa *et al.*, 2015). Ces deux toxines seraient également impliquées dans l'acquisition de fer par *S. aureus* en ciblant les érythrocytes qui, une fois lysés, laissent échapper les composés ferreux (hémoglobine chargée) dans le milieu extracellulaire. On peut également mentionner les leucocidines LukAB (aussi appelée LukGH) et LukDE (voir **Figure 6**). Cette dernière se lie aux



récepteurs à chimiokines CCR5, CXCR5 et CXCR2 et induit la lyse des macrophages, des cellules dendritiques (les cellules présentatrices de l'antigène ou CPA) et des lymphocytes T. On peut enfin citer la leucocidine de Panton-Valentine, qui a la particularité d'être portée par un prophage, est plus souvent retrouvée chez les staphylocoques associés à la communauté (c'est-à-dire les souches qui ne se transmettent pas en priorité en milieu hospitalier), et capable de provoquer des pneumonies nécrosantes. Cette toxine composée des sous-unités LukS et LukF cible le récepteur du complément C5aR des neutrophiles, monocytes et macrophages (voir **figure 8**) mais ce faisant, elle n'induit pas seulement la lyse de ces cellules, mais est également capable d'activer les neutrophiles, les conduisant à relarguer leurs granules pro-inflammatoires.

Les « *Phenol-Soluble Moduline* » (PSM) forment une toute autre catégorie de toxines. Identifiées en 1999 et nommées par leur propriété d'extraction au phénol-chloroforme, ces protéines sont majoritairement codées par des îlots de pathogénie retrouvés chez la plupart des souches de *S. aureus* (Psechel & Otto, 2013). Il existe aussi une PSM, le peptide PSM-mec que l'on retrouve sur la cassette chromosomique staphylococcale SCC-mec. En résumé, on peut distinguer trois *loci* sur lesquels se trouvent les gènes de ces petites protéines: l'opéron *psm $\alpha$*  qui encode quatre peptides PSM $\alpha$ 1- PSM $\alpha$ 4, l'opéron *psm $\beta$*  qui encode les deux peptides PSM $\beta$ 1 et PSM $\beta$ 2, et enfin la séquence codante retrouvée sur l'ARNIII (cet ARN sera décrit dans la troisième partie de cette introduction) qui encode la PSM $\delta$ , autrefois appelée  $\delta$ -toxine. De manière intrigante, les PSM sont regroupées en fonction de leur taille : les  $\alpha$ -peptides (PSM $\alpha$ 1-4 et  $\delta$ -toxine) ont une taille de 20 à 25 acides aminés alors que les  $\beta$ -peptides en ont une d'environ 44 acides aminés (Psechel & Otto, 2013).



**Figure 7 : vue d'ensemble des actions des PSMs** – La production de PSMs de l'ordre du  $\mu\text{M}$  permet la structuration des biofilms de *S. aureus* et le détachement des cellules de cette structure. Certaines PSMs ont également une activité antimicrobienne spécifique pour d'autres espèces. Les PSM $\alpha$  possèdent une activité cytolytique indépendante de la présence de récepteurs et peuvent donc provoquer la lyse de nombreux types cellulaires. La lyse des neutrophiles peut se faire après la phagocytose, ce qui en fait un très bon moyen de défense contre ce type de cellules. Les PSMs peuvent également moduler la réponse immunitaire en induisant un phénotype tolérogénique chez les cellules dendritiques et en empêchant la différenciation des lymphocytes T « helper » en lymphocytes T cytotoxiques. La production de PSMs dans l'ordre du nM active le récepteur FPR2 qui mène à l'activation, l'attraction des neutrophiles et à l'excrétion de cytokines. Adapté de Peschel & Otto, 2013.

Contrairement aux toxines précédemment citées, les PSM n'ont pas besoin de récepteur particulier pour exercer leur action cytolytique et peuvent donc cibler virtuellement toutes les cellules eucaryotes sans distinction (voir **Figures 6 et 7**). Ainsi, les PSM $\alpha$  provoquent la lyse des érythrocytes et des leucocytes avec une activité la plus forte pour le PSM $\alpha$ 3, alors que l'activité de la  $\delta$ -toxine est plus modérée. Les PSM $\beta$ , elles, n'ont pas d'activité cytolytique. Si cette activité est récepteur-indépendante, les PSM $\alpha$  reconnaissent malgré tout le récepteur FPR2 qui leur permet d'induire une réponse inflammatoire. Une étude récente a démontré que la production de ces PSM contribue à la lyse des neutrophiles une fois la bactérie phagocytée, ce qui constitue alors un mécanisme de défense contre le système immunitaire (Otto, 2014). Enfin, il a été démontré que la  $\delta$ -toxine est capable d'induire la dégranulation des mastocytes, ce qui conduit chez le patient à une dermatite. Ainsi, malgré leur proximité fonctionnelle, les PSM ont acquis des fonctions divergentes liées à la pathogenèse au cours de l'évolution.

### **b. Entérotoxines et autres enzymes**

Dans cette sous-partie je noterai les toxines dont la fonction première n'est pas d'induire la lyse de cellules cibles.

Les entérotoxines sont des toxines secrétées d'un poids moléculaire compris entre 20 et 30 kDa qui interfèrent avec les fonctions des cellules intestinales et sont la cause de vomissements et de diarrhées. Ce sont des molécules super-antigènes, c'est-à-dire qu'elles vont provoquer une réaction immunitaire qui échappe au contrôle des cellules de l'hôte infecté. Ces toxines vont activer la prolifération clonale des lymphocytes T sans que celles-ci n'aient été présentées à un antigène par une CPA. La prolifération s'accompagne d'une excrétion massive de cytokines non-spécifiques de l'agent infectieux, ce qui empêche le système immunitaire d'avoir une réponse calibrée. En fonction de la localisation de l'infection, cette excrétion de cytokines peut conduire au choc septique. Parmi ces entérotoxines, on retrouve l'entérotoxine B (*staphylococcal enterotoxine B, SAB*) – cette dernière étant considérée comme une arme biologique – et l'entérotoxine C (SAC), qui est capable de provoquer endocardites et infections des reins, notamment lorsqu'elle est présente chez une souche de *S. aureus* associée à la communauté (Otto, 2014). Les mécanismes d'action précis des entérotoxines ne sont pas encore bien connus.

Une autre molécule de type super-antigène, et certainement la plus connue de cette famille, est la toxine du syndrome du choc toxique (TSST), qui est responsable du syndrome du choc toxique (TSS) en induisant l'excrétion de cytokines pro-inflammatoires telles que les interleukines IL-1 et IL2, le TNF- $\alpha$  et d'autres cytokines. Malgré sa dangerosité, la TSST et les entérotoxines ne sont présentes que chez un faible pourcentage des souches de *S. aureus* existantes (Otto, 2014).

Le staphylocoque doré produit au cours de l'infection un grand nombre d'enzymes aux propriétés diverses, qui permettent de dégrader la matrice extracellulaire environnante ainsi que d'interférer avec le métabolisme des cellules adjacentes ou leurs voies de signalisation. La sécrétion de protéases, lipases et hyaluronidase permettent la destruction du tissu extra-cellulaire de façon générale, mais certaines protéines ont une

action plus ciblée. L'auréolysine, une protéase, dégrade de nombreuses protéines dont l'insuline B et permet également la maturation d'autres protéases de *S. aureus* dont la sérine endopeptidase SspA. La fonction d'autres protéases à sérine n'est pas exactement comprise, exceptée pour les toxines exfoliatives A et B, qui coupent les cadhérines présentes sur les couches superficielles de la peau, ce qui mène chez l'hôte au « *scalded skin syndrome* », une maladie qui entraîne d'importantes lésions cutanées. Enfin, on peut également citer la collagénase, une enzyme qui comme son nom l'indique dégrade le collagène, composante importante du milieu extra-cellulaire chez l'être humain.

### c. Lutter contre le système immunitaire

Dans le cadre d'une infection, la balance de la relation entre le pathogène et l'hôte infecté dépend des facteurs de virulence développés par le premier, et des moyens de défense du système immunitaire du second. J'ai dit auparavant que *S. aureus* est un pathogène opportuniste, qui provoque la maladie lorsque le système immunitaire de son hôte est défaillant. Ce n'est pas pour autant que ledit système est inactif et le staphylocoque doré doit donc faire face aux mécanismes de défense de l'hôte, qu'ils soient moléculaires ou cellulaires. C'est pourquoi je décrirai dans cette sous-partie les facteurs de virulence que *S. aureus* produit pour lutter ou détourner le système immunitaire de l'hôte.

L'un des premiers mécanismes de défense une fois que les barrières naturelles de l'hôte ont été entravées est la migration des neutrophiles depuis les vaisseaux sanguins, vers le foyer de l'infection, et *S. aureus* est capable d'entraver ce processus. Pour ce faire, la bactérie produit une série de molécules qui bloquent le chimiotactisme de ces cellules et leur migration : il s'agit des *superantigen like proteins* (SSL), des protéines inhibitrices de chimiotactisme (CHIPS), du *formyl peptide receptor-like 1 inhibitor* (FLIPr) mais également des PSM que nous avons vues précédemment. Les SSLs sont une famille de molécules structurellement proches des super-antigènes de staphylocoque et dont les gènes sont regroupés sur le génome sur un îlot génomique et un *cluster* (groupe) de gènes impliqués dans l'évasion immunitaire. Les SSL sont capables de reconnaître une large gamme de glycoprotéines présentes à la surface des leucocytes ; elles empêchent ainsi la reconnaissance par les neutrophiles de la sélectine, qui est nécessaire dans le

processus de migration des cellules vers le foyer infectieux ; la liaison à certains récepteurs empêche également la reconnaissance de certains déterminants antigéniques présents à la surface du staphylocoque, par exemple le peptidoglycane ou certaines lipoprotéines ; alors que d'autres se lient à des récepteurs de lymphocytes, empêchant la production de chimiokines qui attirent les neutrophiles (Thammavongsva *et al.*, 2015). La migration des leucocytes est également empêchée par la protéine d'adhérence extracellulaire (Eap). On peut également citer la staphopaine (ScpA), une cystéine protéase sécrétée dont le rôle est de cliver certaines chimiokines impliquées dans la migration des neutrophiles.

Les protéines CHIPS reconnaissent spécifiquement les récepteurs FPR et C5aR ((Thammavongsva *et al.*, 2015), ce qui empêche d'un côté la reconnaissance des peptides formylés de la bactérie (qui constituent un motif spécifiquement associé à la présence d'un pathogène ou PAMP) et d'un autre côté empêche l'activation des leucocytes par le C5a, un effecteur du système du complément, qui facilite la reconnaissance des pathogènes à la phagocytose en se liant dessus (un processus que l'on appelle opsonisation). La protéine FLIPr est également capable de se lier au récepteur FPR mais reconnaît aussi le récepteur FPR2 qui lie spécifiquement les PSM (Prat *et al.*, 2006). Enfin, il a été démontré que cette protéine agit comme un inhibiteur compétitif des récepteurs FcγR qui reconnaissent les immunoglobulines G (IgG), ce qui empêche le processus de phagocytose par les neutrophiles. Les protéines CHIPS et FLIPr sont codées par des îlots de pathogénie et sont retrouvées à une fréquence élevée parmi les souches de *S. aureus* (Otto, 2014).

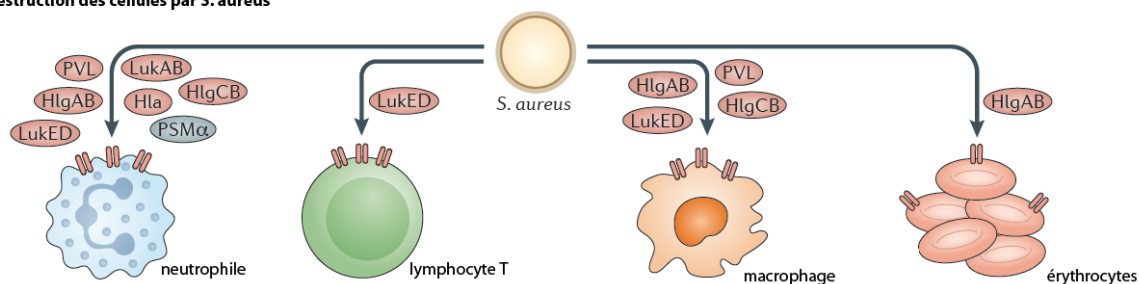
*S. aureus* possède d'autres moyens que les CHIPS pour contrer les molécules du complément, et sa première ligne de défense est tout simplement l'épaisse couche de peptidoglycane qui entoure la cellule. En effet, certaines molécules du complément comme la C5b ont pour rôle de se lier à la membrane cellulaire de la bactérie pour y former des complexes d'attaque de la membrane (MAC). La formation de ce complexe est empêchée car l'épaisse couche de peptidoglycane rend la membrane inaccessible pour la C5b. De plus, plusieurs souches cliniques de *S. aureus* produisent également des polysaccharides qui vont former une couche de protection supplémentaire, la capsule. Deux types de sucres capsulaires existent, le type 5 et le type 8, et la formation de

capsule est souche spécifique (Thammavongsva *et al.*, 2015). La capsule empêche la phagocytose en présence des molécules du complément. Mais *S. aureus* produit d'autres protéines qui empêchent le dépôt de ces molécules à la surface de la bactérie : on trouve parmi celles-ci l'auréolysine, une métalloprotéinase qui, associée aux facteurs du complément H et I, coupe et permet la dégradation des molécules C3b, empêchant son accumulation à la surface bactérienne. Une autre protéine est l'inhibiteur staphylococcal du complément (SCIN) qui se lie à la C3 convertase et inhibe son action, ce qui empêche la production des molécules C3a, C3b et C5a, impliquées dans l'opsonisation et le recrutement des neutrophiles. Des actions similaires sont effectuées par les protéines de liaison au fibrinogène extracellulaire (Efb) et au complément extracellulaire (Ecb) et on peut également mentionner la protéine de liaison aux immunoglobulines (Sbi) et la collagène adhésine (Cna) qui inhibent la voie classique du complément, alors que la staphylokinase dégrade un ensemble de molécules produites par le système immunitaire, telles que les défensines, les IgG, ou la C3b.

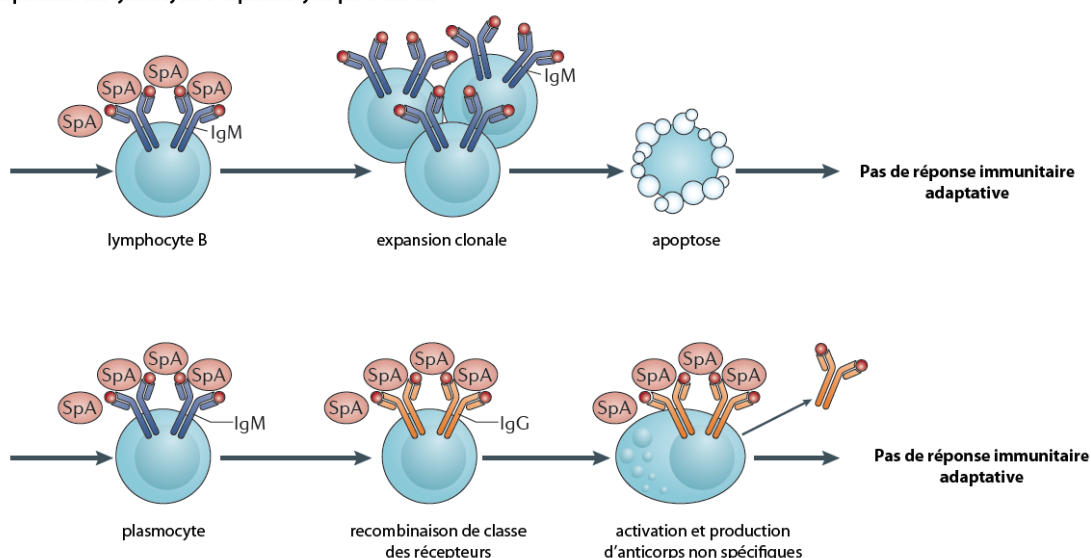
Malgré les moyens de défense produits pour empêcher l'arrivée des neutrophiles, la reconnaissance par le complément ou la phagocytose, il se peut que *S. aureus* se retrouve néanmoins internalisé par une cellule de l'hôte, telle qu'un macrophage, un neutrophile mais aussi des cellules spécialisées comme les ostéoblastes (Dasthgeyb & Otto, 2015). Une fois la bactérie internalisée par ce genre de cellule, elle se retrouve en contact avec de nombreux composés qui lui sont toxiques : des peptides antimicrobiens, des espèces réactives de l'oxygène (ROS) telles que le peroxyde d'hydrogène ( $H_2O_2$ ), les radicaux hydroxyles, mais également le monoxyde d'azote (NO). Mais le staphylocoque doré a également développé des moyens de lutter contre ces molécules. Elle est en effet capable de modifier les sucres du peptidoglycane, l'acide téichoïque ou encore le phosphatidyl-glycérol, tous présents dans la paroi bactérienne, pour empêcher la reconnaissance et liaison des peptides antimicrobiens ou la dégradation de la paroi par le lysozyme. Le pigment qui donne sa couleur au staphylocoque doré, la staphyloxanthine, offre également une résistance contre les ROS. Deux enzymes, les superoxyde dismutases SodA et SodM permettent d'éliminer les groupements superoxyde produits par les neutrophiles, alors que la catalase (KatG) et l'alkylhydroperoxyde réductase (AhpC) protègent la bactérie contre l' $H_2O_2$ . Enfin, la flavohémoglobine (Hmp) permet de détoxifier le NO et la L-lactate déshydrogénase, une

enzyme habituellement impliquée dans le métabolisme central, permet de maintenir l'homéostasie oxydo-réductrice au sein des neutrophiles via la production de lactate.

**a. Destruction des cellules par *S. aureus***



**b. Manipulation des lymphocytes B et plasmocytes par *S. aureus***



**Figure 8 : *S. aureus* peut lutter contre les deux types de réponse immunitaire de l'être humain** – a. La production de nombreuses toxines permet de cibler virtuellement tous les types cellulaires impliqués dans la réponse immunitaire innée. b. La protéine A (Spa) permet d'empêcher le développement d'une réponse immunitaire adaptative, en provoquant la multiplication et l'apoptose massive des lymphocytes B, ou en induisant une activation de plasmocytes et la production d'anticorps non spécifiques de *S. aureus*. Adapté de Thammavongsa *et al.*, 2015.

En dernier lieu, il me faut évidemment parler de la protéine A (spa) produite par *S. aureus* qui a une importance capitale dans l'échappement à la réponse immunitaire adaptative. La protéine A est une protéine de surface avec une très forte affinité pour les immunoglobulines des vertébrés, comprenant celles exprimées chez l'être humain : IgA, IgD, IgG1-IgG4, IgM et IgE. La protéine est d'abord exprimée à la surface de la bactérie avant d'être détachée par l'action de la peptidoglycane hydrolase LytM. C'est une protéine qui est exprimée par tous les isolats de *S. aureus* et dont les domaines de liaison aux immunoglobulines sont conservés. Les capacités immunosuppressives de Spa sont attribuées à deux activités distinctes : la liaison aux domaines Fcγ et Fab des anticorps. La liaison au premier domaine des IgG empêche toute phagocytose de la cellule puisque les anticorps s'attachent « à l'envers » à la surface de la protéine, et ne peuvent plus être

reconnu ensuite par les macrophages, alors que la liaison au second domaine provoque un effet super-antigène sur les lymphocytes B : les cellules prolifèrent de façon non contrôlée et subissent une apoptose de masse (Thammavongsva *et al.*, 2015). La liaison aux récepteurs IgM des lymphocytes B entraîne également un changement de classe de ces récepteurs en IgG, suivie ensuite d'une sécrétion d'anticorps qui ne sont pas spécifiques de *S. aureus*, empêchant ainsi le développement d'une réponse immunitaire adaptée (voir **figure 8b**).

**Les différentes phases de l'infection provoquées par *S. aureus* peuvent être facilement expliquées par sa capacité à synthétiser un nombre très important de facteurs de virulence, dont l'action permet la colonisation des tissus de l'hôte, leur destruction puis la dissémination de la bactérie, ainsi que la subversion et la résistance au différents moyens mis en place par le système immunitaire pour lutter contre le pathogène. Mais l'expression de ces facteurs de virulence, pour être efficace, doit être très finement régulée dans l'espace en fonction de l'endroit où se trouve la bactérie, et dans le temps pour être en accord avec les différentes phases de croissance. C'est à ce niveau qu'interviennent de multiples stratégies de régulation de l'expression des gènes, que nous allons voir en détails dans la troisième et dernière partie de cette introduction.**



## Partie 3 : La régulation de l'expression génétique chez *S. aureus*

A l'échelle d'une bactérie la croissance cellulaire est une aventure de tous les dangers. Les changements de l'environnement se ressentent bien plus pour un organisme unicellulaire et même des micro-variations de température, de pH ou de disponibilités en nutriments peuvent se révéler fatales si l'organisme n'est pas capable de s'adapter rapidement. Dans le cadre d'un organisme tel que *S. aureus*, le passage d'un mode de vie commensal à la surface épithéliale de l'homme, à celui de pathogène, dans les différents tissus qui sont colonisables, sont synonymes de changements environnementaux encore plus drastiques. Et c'est sans compter sur les multiples attaques du système immunitaire qui prennent des formes très différentes (phagocytose, cytotoxicité, anticorps...). L'adaptation rapide et contrôlée à ces changements est plus que nécessaire pour que l'infection par la bactérie soit efficace. Si nous avons vu dans la précédente partie bon nombre de facteurs de virulence qui sont cautionnaires de cette efficacité, il faut maintenant regarder comment *S. aureus* contrôle l'expression de ces gènes pour qu'elle se fasse au bon endroit et au bon moment.

### 1. Un grand nombre de protéines...

*S. aureus* possède un nombre important de protéines qui sont capables d'assurer la régulation des gènes impliqués dans les processus infectieux. Par exemple l'expression de la protéine A est effectuée par au moins 21 protéines en réponse à des stimuli environnementaux tels que la densité cellulaire, la disponibilité en nutriments, fer ou oxygène (Bischoff & Romby, 2016). Les facteurs de régulation de nature protéique peuvent se ranger dans plusieurs catégories, dont nous allons étudier quelques représentants.

#### a. Les facteurs sigma alternatifs

Un facteur sigma est une protéine auxiliaire au complexe de l'ARN polymérase qui est chargé de guider ce dernier vers les gènes qui doivent être transcrits. Pour ce faire, ce facteur se lie à des séquences d'ADN spécifiques (le promoteur) ce qui permet d'amener

l'ARN polymérase au bon endroit sur le génome pour initier la transcription. Ce mécanisme permet la transcription de gènes et d'opérons dont les fonctions ne sont pas liées, du moment qu'ils partagent une même séquence promotrice. Les gènes qui sont exprimés de façon constitutive pour le bon fonctionnement de la cellule, appelés aussi gènes de ménage, possèdent des promoteurs reconnus par le facteur sigma 70, qui est le facteur sigma conventionnel et également appelé  $\sigma^A$  chez *S. aureus*. Mais des facteurs sigmas alternatifs existent également, et vont reconnaître des séquences promotrices bien spécifiques dont l'expression est nécessaire dans des conditions particulières. Chez *S. aureus* ces facteurs sigma alternatifs sont au nombre de trois : ce sont les facteurs  $\sigma^B$ ,  $\sigma^H$  et  $\sigma^S$ .

Le facteur  $\sigma^B$  est le facteur alternatif le plus étudié chez *S. aureus*. C'est le facteur qui répond aux phénomènes de stress de la bactérie, qu'ils soient liés à la phase de croissance (la disponibilité en nutriments et l'accumulation de sous-produits toxiques du métabolisme ont lieu en phase stationnaire) ou à des changements physico-chimiques du milieu extracellulaire tels que la température, le pH ou la présence de composés toxiques comme les antibiotiques. De même, il a été démontré que son expression est plus importante lorsque la bactérie est internalisée par des macrophages et lors d'expériences d'infections dans des modèles murins. De nombreuses études de transcriptomique et de protéomique ont montré que le régulon (c'est à dire, l'ensemble des gènes dont l'expression est régulée) de ce facteur sigma comprend plus de 229 gènes et opérons, dont les fonctions impliquent la production de facteurs de virulence, la régulation de la transcription, l'adhésion cellulaire ou l'expression de toxines. De fait, ce facteur a très certainement un rôle important à jouer dans la pathogénicité de la bactérie en modulant l'expression d'un grand nombre de facteurs de virulence de manière fine en réponse à des changements environnementaux (Bischoff & Romby, 2016).

Le facteur  $\sigma^H$  est le second facteur alternatif à avoir été identifié chez *S. aureus*. Cependant, les conditions d'expression de ce facteur ne sont pas encore tout à fait comprises, et le régulon comprend une quarantaine de gènes (Fagerlund et al., 2014). Parmi ceux-là, on trouve les gènes *comE* et *comG* qui sont impliqués dans la compétence de *S. aureus* et sa capacité à être transformée avec de l'ADN extracellulaire, bien que cette capacité chez la bactérie soit très faible. Enfin, le facteur  $\sigma^S$  n'a été identifié que très

récemment (Shaw *et al.*, 2008) et si les études menées en laboratoire ont démontré qu'il est capable de se lier à l'ARN polymérase et de réguler positivement la transcription de son propre gène, son expression reste très faible en conditions de culture standard et son rôle n'est pas intégralement compris. Des études plus récentes ont néanmoins démontré que l'expression du gène *sigS* est fortement activée en présence de composés chimiques qui induisent des dommages à l'ADN et causent une instabilité de la paroi cellulaire (Miller *et al.* 2012). Et si le régulon de ce facteur n'a pas encore été établi, son rôle supposé se trouverait dans la réponse au stress et la virulence (Bischoff & Romby, 2016).

### **b. Les facteurs de transcription**

Ces protéines, tout comme les facteurs  $\sigma$ , sont capables de reconnaître des séquences d'ADN de manière spécifique et de s'y lier. Ce faisant, la fixation sur l'ADN va entraîner des changements de conformation de l'ADN qui vont rendre les séquences promotrices plus accessibles d'une part, et la protéine va d'autre part recruter les sous-unités alpha de l'ARN polymérase – et donc le complexe tout entier pour ensuite initier la transcription. Le nombre exact de facteurs de transcription peut varier d'une souche de *S. aureus* à l'autre (ce qui est dû à la plasticité du génome), mais le *core* génome comprend 112 gènes de ces facteurs, qui peuvent être subdivisés en 36 familles, parmi lesquelles la famille des protéines Sar, les protéines de liaison aux métabolites, et les systèmes à deux-composants que nous verrons dans une troisième sous-partie.

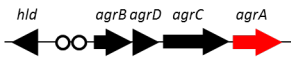

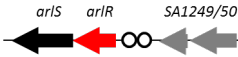
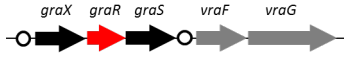
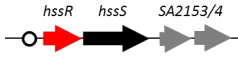
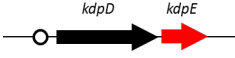
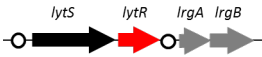
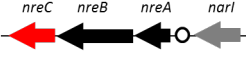
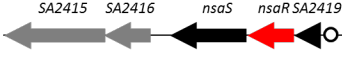
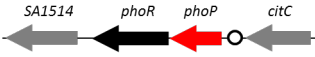
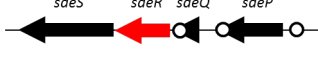

La famille des protéines Sar est représentée par le facteur de transcription SarA – dont une dizaine d'homologues existent chez *S. aureus* : par exemple les facteurs SarR, SarS, ou MgrA). La protéine SarA active la synthèse de la protéine de liaison à la fibronectine et d'un certain nombre d'exotoxines alors qu'elle réprime la synthèse de la protéine A et de plusieurs protéases en se liant aux séquences promotrices de ces gènes au niveau de motifs riches en A-T (Bischoff & Romby, 2016). Des études récentes ont montré de façon intéressante que cette protéine pourrait également se lier à des ARN messagers (ARNm) et influencer leur dégradation – et notamment l'ARNm codant la protéine A (Monk *et al.*, 2012). Cette dualité de fonction montre qu'une protéine peut ne pas se cantonner qu'à un seul rôle. A côté de ce facteur, les protéines SarR et SarS ont un rôle antagoniste à

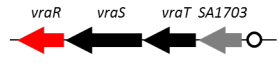

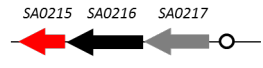
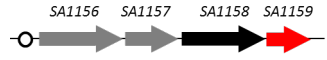
SarA. SarR bloque la transcription de *sarA* en se fixant à sa région promotrice et agit comme compétiteur sur les sites de fixation reconnus par SarA. SarS agit de façon similaire, ce facteur active la synthèse de protéine A et réprime la production d'exotoxines (Bischoff & Romby, 2016).

La protéine MgrA (*multiple gene regulator*) est un homologue de la famille Sar dont le régulon comprend au moins 355 gènes et opérons et qui joue un rôle important dans l'expression de facteurs de virulence. Les études menées sur ce facteur ont notamment démontré qu'il régule positivement l'expression des sucres de la capsule, de l'alpha-toxine et de leucocidines alors qu'il réprime la synthèse de protéines de surface ainsi que la formation de biofilms. Il exerce également un contrôle sur l'expression de pompes à efflux, ce qui a un impact sur la résistance à plusieurs antibiotiques. Son rôle sera discuté plus en profondeur dans les parties Résultats et Discussion.

### c. Les systèmes à deux-composants et l'importance du système agr

Les systèmes à deux-composants (ou TCS pour *two-component system*) forment une classe particulière des facteurs de transcription qui sont aussi la forme la plus répandue de régulateurs aux stimuli environnementaux chez les bactéries. Comme leur nom l'indique, il s'agit de systèmes qui fonctionnent avec deux unités : une protéine de membrane qui répond à un stimulus du milieu extracellulaire et qui possède une activité histidine kinase, et une protéine cytoplasmique qui est le régulateur de réponse. Lorsque la protéine membranaire perçoit le changement environnemental qui lui est spécifique, elle s'autophosphoryle par son activité histidine kinase et le groupement phosphate est ensuite transmis au régulateur de réponse qui, une fois activé, va reconnaître les séquences promotrices des gènes dont la transcription doit être initiée. Il existe chez *S. aureus* 16 TCS qui ont été identifiés chez les souches dont le génome est correctement annoté, et la plupart de ces systèmes sont compris dans des circuits de régulation plus complexes qui font intervenir d'autres protéines comme celles citées précédemment (Bischoff & Romby, 2016). Ces systèmes à deux composants permettent de relayer un ou plusieurs signaux extérieurs à la reprogrammation génétique pour ajuster le métabolisme et l'expression des facteurs de virulence (**Tableau 1**).

Nom	ID du gène	Organisation génétique	Fonction	Famille régulatrice	Régulé par
AgrCA	SA1843 SA1844		Quorum sensing, régulation de la synthèse des exoprotéines au niveau transcriptionnel et post-transcriptionnel	LytTR	AgrA, ArlR, CcpA, CcpE, CodY, RpiRC, Rsr, SarA, SigB, SrrA
AirSR (YhcSR)	SA1667 SA1666		Régulation du métabolisme de la paroi cellulaire dépendante du potentiel Redox, modulation de la chaîne respiratoire des nitrates en conditions anaérobies, contrôle transcriptionnel des opérons <i>lac</i> et <i>opuC</i> ; gène essentiel	LuxR	NsaR
ArlRS	SA1247 SA1248		Régulation de la synthèse des facteurs de virulence et de la formation de biofilm; non-essentiel	OmpR	Agr, SarA
GraRS	SA0614 SA0615		Fait partie d'un réseau de transduction de signal en 5 parties, qui contrôle la détection et la résistance aux peptides antimicrobiens cationiques chez <i>S. aureus</i>	OmpR	
HssRS	SA2151 SA2152		Un TCS sensible aux groupements hème et impliqués dans leur détoxification; non-essentiel	OmpR	MgrA
KdpDE (KdpD-PstA)	SA1882 SA1883		Régulation dépendante de la croissance de la synthèse de la capsule, régulation dépendant du c-di-AMP du transport des ions, régulation dépendante des ions K <sup>+</sup> des facteurs de virulence ; non essentiel	OmpR	Agr, LuxS
LytSR	SA0250 SA0251		Régulation dépendante du potentiel de membrane du système holine/antiholine et de la formation de biofilm, adaptation aux peptides antimicrobiens cationiques de l'hôte ; non essentiel	LytR	MgrA
NreBC	SA2179 SA2180		Régulation dépendante de l'oxygène de gènes impliqués dans la réduction et le transport des nitrates et des nitrites ; non essentiel	OmpR	Rex, AirR
NsaRS (BceRS) (BraRS)	SA2417 SA2418		Ce TCS reconnaît les dommages de la paroi et de la membrane cellulaire, et affecte la formation de biofilm, la virulence et la synthèse de la paroi. Fait possiblement partie d'un réseau de transduction de signal à 5 composants ; non essentiel	OmpR	Agr
PhoPR	SA1515 SA1516		Fonction inconnue ; non essentiel	OmpR	
SaeRS	SA0660 SA0661		Régulation de la synthèse d'exoprotéines, impliquées essentiellement dans l'échappement au système immunitaire ; non essentiel	OmpR	Agr, Fur, Rot, SaeR, SigB, WalR
SrrAB (SrhSR)	SA1322 SA1323		Répression de la synthèse d'exoprotéines et activation de la formation des polysaccharides intercellulaires d'adhésion en basse concentration d'O <sub>2</sub> ; non essentiel	OmpR	Rex, SrrA

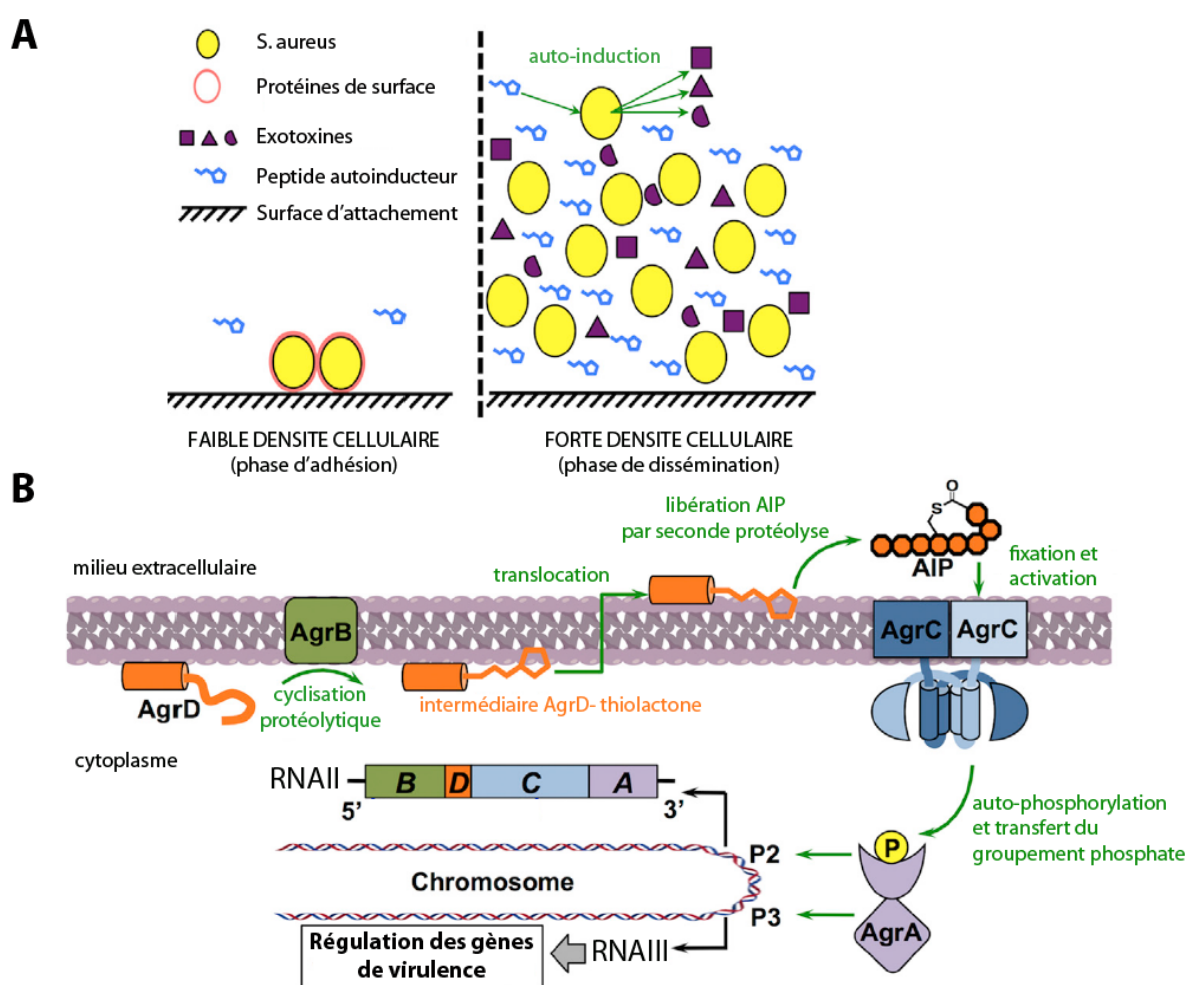
VraSR	SA1700 SA1701		Régulation du „stimulon de la paroi cellulaire“ en réponse à la présence d’anti-microbiens ciblant la paroi ; non essentiel	LuxR	VraR
WalKR (VicRS) (YycGF)	SA0017 SA0018		Régulation du métabolisme de la paroi et de la virulence ; rôle supposé dans la colonisation des fosses nasales ; non essentiel	OmpR	Rex, WalR
-	SA0215 SA0216		Fonction inconnue, rôle suggéré dans la réponse de <i>S. aureus</i> à l’anaérobiose ; non essentiel	AraC	NsaR
-	SA1158 SA1159		Fonction inconnue ; non essentiel	LuxR	

**Tableau 1: Systèmes à deux composants chez *S. aureus*** – Les ID des gènes sont basés sur la séquence du génome de *S. aureus* N315. Les gènes des régulateurs de réponse sont indiqués par des flèches rouges, les gènes codant les autres composants du système sont indiqués par des flèches noires. Les gènes environnants dont la fonction n’est pas encore connue sont indiqués par des flèches grises. Les promoteurs sont indiqués par des cercles. Adapté de Bischoff & Romby, 2016.

Le système à deux composants le plus connu et le plus étudié chez *S. aureus* est le système *agr* (*accessory gene regulator*) et est un très bon exemple de TCS intégré dans un réseau de régulation complexe. Il fait en effet intervenir un mécanisme de mesure de la densité cellulaire (*quorum sensing*) par la production d’un peptide auto inducteur (AIP) qui est excrété dans le milieu extracellulaire et qui, lorsque la densité cellulaire atteint une concentration seuil, va induire chez toutes les cellules l’expression simultanée d’un ensemble de gènes. Cette expression massive du même ensemble génique fait passer la bactérie lors de l’infection d’une phase d’adhésion et de production de biofilm à une phase d’excrétion d’exotoxines et de dissémination.

Le locus *agr* se compose de deux unités transcriptionnelles divergentes dont les transcrits sont appelés ARNII et ARNIII et dont la transcription est initiée par les promoteurs P2 et P3 respectivement (Bischoff & Romby, 2016). L’ARNII est un transcrit polycistronique qui contient l’opéron *agrBDCA*. Les protéines AgrB et AgrD sont les composantes du système de *quorum sensing*. AgrD est le précurseur de l’AIP qui est mûri par la peptidase transmembranaire AgrB. Le composé intermédiaire qui résulte de cette maturation est ensuite exporté au travers de la membrane cellulaire tout en subissant une seconde maturation pour être l’AIP effectif une fois arrivé dans le milieu extracellulaire. Ce peptide constitue le signal qui sera reconnu par le système à deux composants formé par AgrC et AgrA. La protéine AgrC est le récepteur membranaire à activité histidine kinase qui reconnaît l’AIP. S’ensuit alors une cascade de phosphorylation qui conduit à la phosphorylation du régulateur de réponse AgrA qui, une fois activé, régule positivement la transcription de l’ARNII en se fixant au promoteur

P2 et également celle de l'ARNIII en se fixant au promoteur P3 (Wang & Muir, 2016). De fait, la fixation d'AgrA sur le promoteur P2 alimente une boucle de régulation positive pour la production des quatre protéines du système *agr* et donc du peptide auto-inducteur, élément clé du système de *quorum sensing*. Mais la protéine AgrA va également aller activer l'expression des gènes des PSM $\alpha$  et des PSM $\beta$  en se fixant directement dans les régions promotrices, tout en inhibant l'expression de gènes impliqués dans le métabolisme des acides aminés et des sucres par un mécanisme encore inconnu (Queck *et al.*, 2008).



**Figure 9 : Rôle du système *agr* chez *Staphylococcus aureus*** - A. Représentation des deux phases physiologiques de *S. aureus* qui sont modulées par le système *agr*. La production du peptide auto-inducteur (AIP) permet, une fois une concentration seuil atteinte, d'activer l'expression des gènes d'exotoxines et d'inhiber celle des protéines d'adhésion. B. Représentation schématique du système *agr*. Le fonctionnement est explicité en détails dans le texte suivant. Adapté de Wang & Muir, 2016.

L'ARNIII est l'effecteur intracellulaire du système *agr* et un régulateur central pour la transition de la bactérie de sa phase d'adhésion à celle de dissémination. C'est un ARN

multifonctionnel qui possède une ORF (un cadre ouvert de lecture) dans sa séquence qui encode la  $\delta$ -toxine, la seule PSM dont l'expression n'est pas régulée directement par AgrA (voie revue Tomasini et al., 2014). En plus de cette propriété messagère, l'ARNIII est également un ARN régulateur qui influence l'expression d'ARNm cibles par appariements de base pour soit activer la traduction du messager (par exemple, en rendant accessible un site de liaison du ribosome qui était séquestré dans une structure secondaire), soit l'inhiber en s'appariant grâce à ses motifs riches en cytosines à la séquence Shine-Dalgarno (SD) de l'ARNm cible et provoquer sa dégradation par une RNase spécifique des régions ARN à double brin (dans ce cas, il s'agit de la RNase III). L'ARNIII s'apparie directement avec l'ARNm de la protéine A (*Spa*) et inhibe sa traduction. À l'inverse, son appariement avec l'ARNm *hla* (qui code l'alpha-toxine) active sa traduction. De plus, l'ARNIII s'apparie également avec l'ARNm de *rot*, qui code un facteur de transcription qui active l'expression de *spa* et inhibe la production d'exotoxines. Cet appariement, qui conduit à une répression de l'expression de *rot*, permet donc à l'ARNIII d'assurer la production de toxines et l'inhibition de la synthèse des protéines d'adhésion à deux niveaux : au niveau transcriptionnel en interférant avec l'action de *rot*, et au niveau post-transcriptionnel en se liant directement aux ARNm des protéines concernées (Wang & Muir, 2016). L'ARNIII, via le système *agr*, permet donc de lier la croissance cellulaire à la formation de biofilms et la production de toxines. Il est un élément central d'un réseau de régulation complexe dans lequel les facteurs sigma, les facteurs de transcription et autres molécules d'ARN sont imbriqués. Ce réseau ne cesse de se complexifier au fur et à mesure que les études chez *S. aureus* progressent.

L'exemple de l'ARNIII en tant qu'ARN régulateur est très loin d'être un cas isolé chez *S. aureus* puisque depuis plusieurs décennies, des ARN de ce genre ont été découverts par dizaines chez la bactérie.

## 2. ... et d'ARN aussi

Les ARN régulateurs, sont souvent non codants (ARNnc). Ce sont des molécules d'ARN qui, contrairement aux ARNm, ne vont pas coder de protéine. La plus grande classe des ARNnc agissent comme régulateur de la traduction d'ARNm cibles, en séquestrant ou en rendant accessible la séquence SD au ribosome, et/ou en modulant la stabilité du



messager. L'ARNIII est le premier ARNnc à avoir été découvert chez *S. aureus* (Novick *et al.*, 1993). Depuis, l'utilisation combinée d'approches *in silico* et expérimentales telles que les puces à ADN, le séquençage à haut débit, a conduit à l'identification chez la bactérie d'au moins 250 ARNnc, pour la plupart desquels la fonction n'a pas encore été élucidée. Par ailleurs, on s'aperçoit que le transcriptome de la bactérie est bien plus complexe que nous pouvions l'imaginer. Les ARNm ne servent pas qu'à être traduits mais peuvent aussi contenir des éléments régulateurs qui agissent en cis (thermosenseur, riboswitch) ou constituer le réservoir d'ARNnc. Ils peuvent même agir comme ARN régulateur pour réguler le taux d'un ARNnc cible. Il devient parfois difficile de déterminer qui peut être le régulateur ou la cible. Par ailleurs, il apparaît qu'un grand nombre d'opérons divergeants ou convergeants se chevauchent et selon leur taux respectifs entraineraient la régulation de l'un ou de l'autre opéron (Sesto *et al.*, 2013 ; Lasa *et al.*, 2012). Même des régions complètement inattendues servent d'ARN régulateurs comme les espaceurs des opérons d'ARNt (Lalaouna *et al.*, 2015). On doit donc s'attendre à voir émerger de plus en plus d'ARN aux fonctions et mécanismes d'action multiples. Dans ce qui va suivre, je me focaliserai principalement sur les interactions ARN-ARN et la classe prédominante des ARNnc qui ciblent les ARNm.

Revenons brièvement aux ARNnc de *S. aureus*. Leurs gènes sont dispersés sur l'ensemble du *core* génome de *S. aureus* mais également sur des éléments génétiques mobiles tels que les plasmides et îlots de pathogénie (Caldelari *et al.*, 2011). Ils sont impliqués dans tous les processus cellulaires et lient le métabolisme central à la virulence, en passant par les réponses aux stress. Dans le laboratoire, un très grand nombre d'ARN possiblement régulateurs ont été identifiés dont onze ARNnc, nommés de RsaA à RsaK, qui font l'objet d'une étude plus approfondie (Geissmann *et al.*, 2009). L'équipe s'attache à élucider la fonction de ces ARNs tout en les remplaçant dans les réseaux de régulation de la bactérie. Pour mettre l'accent sur l'importance des ARNnc de *S. aureus*, leur diversité autant dans la forme qu'ils prennent que dans leurs mécanismes d'action, ainsi que la multiplicité des processus dans lesquels ils sont impliqués, j'ai participé en premier auteur à la rédaction d'une revue scientifique (Tomasini *et al.*, 2014) qui fait l'état des lieux sur ce sujet et qui est rattachée à la suite de cette introduction, dont elle vient apporter la conclusion.



## Review

The importance of regulatory RNAs in *Staphylococcus aureus*

Arnaud Tomasini<sup>a</sup>, Patrice François<sup>b</sup>, Benjamin P. Howden<sup>c,d,e,f</sup>, Pierre Fechter<sup>a</sup>, Pascale Romby<sup>a</sup>, Isabelle Caldelari<sup>a,□</sup>

<sup>a</sup> Architecture et Réactivité de l'ARN, Université de Strasbourg, CNRS, IBMC, Strasbourg, France

<sup>b</sup> Genomic Research Laboratory, Department of Medical Specialties, Geneva University Hospitals, University of Geneva, Genève, Switzerland

<sup>c</sup> Austin Centre for Infection Research (ACIR), Infectious Diseases Department, Austin Health, Heidelberg, Victoria, Australia

<sup>d</sup> Microbiology Department, Austin Health, Heidelberg, Victoria, Australia

<sup>e</sup> Department of Microbiology and Immunology, University of Melbourne, Parkville, Victoria, Australia

<sup>f</sup> Department of Microbiology, Monash University, Clayton, Victoria, Australia

## a r t i c l e i n f o

## Article history:

Available online xxxx

## Keywords:

Staphylococcus  
Regulatory RNA  
Virulence

## a b s t r a c t

RNA molecules with regulatory functions in pathogenic bacteria have benefited from a renewed interest these two last decades. In *Staphylococcus aureus*, recent genome-wide approaches have led to the discovery that almost 10–20% of genes code for RNAs with critical regulatory roles in adaptive processes. These RNAs include trans-acting RNAs, which mostly act through binding to target mRNAs, and cis-acting RNAs, which include regulatory regions of mRNAs responding to various metabolic signals. Besides recent analysis of *S. aureus* transcriptome has revealed an unprecedented existence of pervasive transcription generating a high number of weakly expressed antisense RNAs along the genome as well as numerous mRNAs with overlapped regions. Here, we will illustrate the diversity of trans-acting RNAs and illustrate how they are integrated into complex regulatory circuits, which link metabolism, stress response and virulence.

© 2013 Elsevier B.V. All rights reserved.

## Contents

1. Introduction	00
2. <i>S. aureus</i> sRNA at the crossroads of bacterial metabolism, stress response, and virulence	00
2.1. sRNA and small-colony variant phenotype	00
2.2. sRNA and the TCA cycle	00
2.3. SigmaB-dependent sRNAs	00
2.4. Expression of sRNAs under antibiotic exposure	00
3. sRNAs acting directly on virulence gene expression	00
3.1. Two intertwined regulatory networks quorum sensing and virulence factor control	00
3.2. <i>S. aureus</i> RNAIII regulates the temporal expression of virulence determinants	00
3.3. SprD mediates a crosstalk between pathogenicity island and the core genome to regulate virulence	00
3.4. A pathogenicity-island sRNA attenuates MRSA virulence	00
3.5. Another long sRNA contributes to pathogenesis	00
3.6. Expression of sRNAs in <i>S. aureus</i> during infection	00
4. RNA-binding proteins	00
4.1. Hfq protein, a mysterious protein in <i>S. aureus</i>	00
4.2. Ribonucleases as possible co-factors of sRNAs	00
4.3. A transcriptional regulator affects RNA stability	00
5. Concluding remarks	00
Acknowledgements	00
References	00

□ Corresponding author.

E-mail address: [I.caldelari@ibmc-cnrs.unistra.fr](mailto:I.caldelari@ibmc-cnrs.unistra.fr) (I. Caldelari).

## 1. Introduction

*Staphylococcus aureus* is a versatile microorganism that is able to adapt to various ecological niches. It is also present in a large proportion of the population (almost 30% of individuals) as a commensal of skin and anterior nares, and conversely, causes a large range of hospital-acquired and community infections (Chambers and Deleo, 2009; Tristan et al., 2007). This large variety of lifestyles is due to the fact that the bacteria produces a large array of virulence factors responsible for adhesion, invasion, and dissemination in host tissues, the acquisition of nutrients, and for combating the host immune system (Novick, 2003). In addition, two component systems, transcriptional regulatory proteins, and RNAs form intricate interactions and regulatory circuits that sense and respond to a large variety of signals to coordinate gene expression in a dynamic manner (Cheung et al., 2004; Romilly et al., 2012a).

During the last decades, tremendous works have shown that besides transcriptional regulatory proteins, small and usually non-coding RNAs (sRNAs) act as important players in adaptive processes in *Escherichia coli* and many other bacteria (Caldelari et al., 2013; Papenfort and Vogel, 2010). In *S. aureus*, the first reported regulatory RNA, the so-called RNAIII, was shown to be one of the main intracellular effectors of the agr quorum-sensing system (Novick and Geisinger, 2008; Novick et al., 1993). Later on, research on the discovery and functional characterization of sRNAs in *S. aureus* has exploded leading to an impressive repertoire of various regulatory RNAs (reviewed in Felden et al., 2011; Romilly et al., 2012a). Approaches for identifications were based on various computing softwares (Geissmann et al., 2009; Marchais et al., 2009; Pichon and Felden, 2005), Affymetrix microarrays (Anderson et al., 2006; Roberts et al., 2006) and tiling arrays (Lasa et al., 2011), conventional cloning/sequencing of small sized cDNAs (Abu-Qatouseh et al., 2010), and high throughput sequencing (Beaume et al., 2010; Bohn et al., 2010; Howden et al., 2013; Lasa et al., 2011; Lioliou et al., 2012). All these studies (see Table 1 for details) revealed that the sRNA genes are located randomly in the core genome and are also present in the horizontally acquired mobile elements. Transcription of many of these sRNAs is tightly controlled and responds to growth phase and to stress conditions suggesting that they are engaged in regulatory functions (Beaume et al., 2010; Bohn et al., 2010; Geissmann et al., 2009; Howden et al., 2013; Lasa et al., 2011; Roberts et al., 2006).

As in many bacteria, sRNAs in *S. aureus* are heterogeneous in size, structure and function. They can be very long (>500 nts) and some of them encode peptides. A large number of regulatory RNAs act on independently expressed targets (mRNA or proteins). Moreover, genome-wide analysis unexpectedly revealed the expression of a large number of short and long antisense RNAs (asRNAs), as well as mRNAs that present overlaps with the entire length or with some parts of their 5' or 3'-untranslated regions (Lasa et al., 2011). The biological consequences of such antisense trans2013135-crip-tion requires further experimental investigation, but some examples have recently indicated regulatory roles (Howden et al., 2013; Lasa et al., 2011; Lioliou et al., 2012; Sayed et al., 2011). Another prominent class of regulatory RNA elements, the so-called riboswitches, are transcribed as part of their target mRNAs (Breaker, 2011; Dambach and Winkler, 2009). They are mostly located in the 5' UTRs of some mRNAs and exhibit a structured receptor domain specifically recognized by a small metabolite. The binding induces a conformational change of the downstream mRNA that provokes premature transcription arrest, translation repression/activation, or cleavage. In general, riboswitches regulate the associated downstream genes, which are involved in the uptake or metabolism of the sensed metabolite (Breaker, 2011; Serganov, 2010). In *S. aureus*, riboswitches control almost 2% of the genes

associated with metabolism and the most common respond to the intracellular concentrations of S-adenosylmethionine, thiamine pyrophosphate, flavin mononucleotide, lysine, glycine, guanine, 7-aminomethyl-7-deazaguanine, and glucosamine-6-phosphate (reviewed in Caldeleri et al., 2011; Felden et al., 2011). Riboswitches are considered as promising targets for the design of novel antibacterial compounds that can overcome the current mechanisms of resistance, because they regulate essential genes and their structures reveal specific pockets for binding small compounds (Breaker, 2009; Mulhbachter et al., 2010b).

Conservation of most sRNAs is restricted to the genus *Staphylococcus*, except for the cis-acting regulatory sequences and house-keeping RNAs. Although many of them are encoded within the *S. aureus* species, their expression appears to be strain-specific (Bohn et al., 2010; Geissmann et al., 2009; Pichon and Felden, 2005). Therefore the sRNA expression profiles might contribute to the diversity of phenotypes. With the new development of genome-wide approaches, the functional analysis of *S. aureus* sRNAs is expected to undergo a paradigm shift because it is now possible to monitor the expression profiles of, in principle, all RNAs and to gain knowledge on the impact of sRNAs on metabolism, adaptability to environment or host, resistance pathways or pathogenicity mechanisms. In this review, we will focus on trans-acting sRNAs in *S. aureus* for which the function has been solved. We will illustrate by specific examples that numerous sRNAs and associated proteins are found embedded in complex regulatory circuits, which links metabolism, stress response and virulence.

## 2. *S. aureus* sRNA at the crossroads of bacterial metabolism, stress response, and virulence

*S. aureus* is exposed to a number of environmental stress conditions during the infection process. It has been well documented that the transcription and synthesis of the virulence determinants are influenced by environmental and nutritional stresses, and by the energy metabolism (Somerville and Proctor, 2009). Interestingly, the synthesis of trans-encoded sRNA is induced upon stresses such as temperature change, acid or alkali, nutrient limitation, oxidative stress or antibiotic treatment that can be encountered during the host infection (Anderson et al., 2006; Beaume et al., 2010; Bohn et al., 2010; Geissmann et al., 2009). Other sRNAs are presumably part of the growth phase specific transcriptional responses controlling metabolism and possibly virulence.

### 2.1. sRNA and small-colony variant phenotype

Abu-Qatouseh et al. (2010) constructed cDNA libraries from samples collected at different growth stages for two isogenic derivatives of a clinical strain displaying the normal and the variant small-colony variant (SCV) phenotypes, respectively. SCVs are slow-growing and persistent bacteria with an increased ability to form biofilm and high resistance to aminoglycoside antibiotics. Although many alterations in metabolism can be associated with slow-growth, a limited effect on deficient electron transport has been described for the *S. aureus* SCV phenotype (Proctor et al., 2006). Additional features concern decreased pigmentation, hemolytic and coagulase activity, and respiration that can be linked to electron transport deficiency. A total of 142 sRNAs were identified and expression of 18 candidates was validated by Northern blot. These sRNAs originated from intergenic regions (IGRs) or were antisense or partially overlapping with ORFs. The expression of the newly identified sRNAs was growth phase regulated and most interestingly, for a subset of them comprising RNAIII, was phenotype-specific. Certain sRNAs were exclusively expressed in the

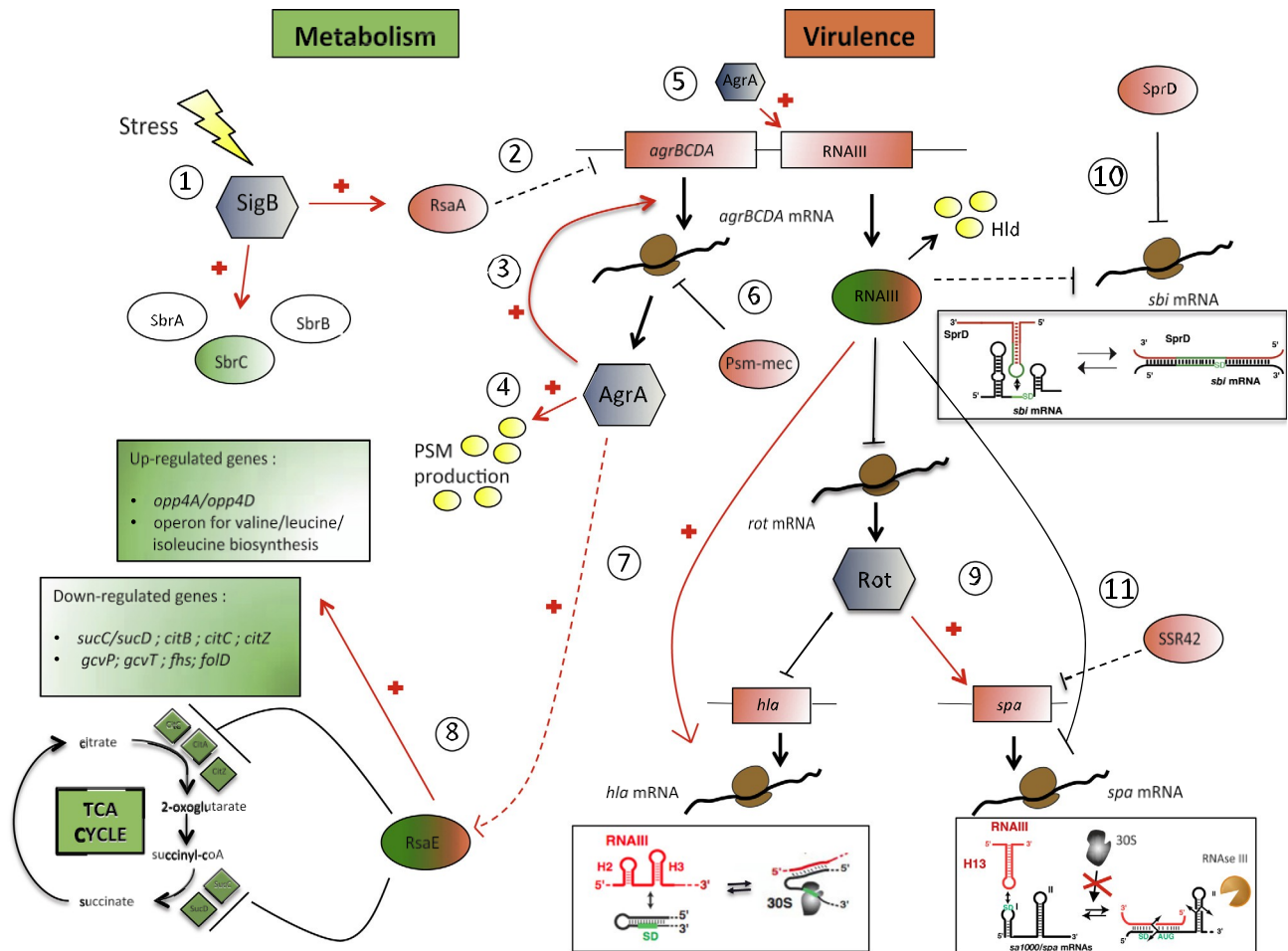


Fig. 1. RNAs and transcriptional factors link metabolism and virulence in *S. aureus*. The left panel of the figure shows molecular effectors involved mainly in *S. aureus* general metabolism (green boxes) and the right panel those involved in virulence (red boxes). Transcriptional factors are represented by grey hexagons and all regulatory RNAs by red ovals (except for RNAIII and RsaE which are also colored in green to emphasize their dual roles in central metabolism and virulence while SbrA/C are uncolored because no functions have been assigned yet). Red arrows and red crosses shows a positive effect on expression, in contrast black connections with a terminal bar denote a negative effect, dashed lines represent putative interactions: (1) Several stress factors (pH, T°C, nutrients availability) can trigger the expression of the stress-response sigma factor Sigma B, which activates the transcription of RsaA among other genes (SbrA, B, C). (2) The *agr* operon is probably indirect. The *AgrA* transcriptional factor autoregulates its own transcription (3), induces PSM production (4) but also enhances RNAIII expression (5). (6) *AgrA* itself is subjected to a negative control by the *Psm-mec* RNA in methicillin resistant *S. aureus* strains. (7) It has a direct positive effect on PSMs (phenol soluble modulins) production (PSMs are small molecule with hemolytic effect) and putatively on the non-coding RsaE transcription. (8) RsaE controls negatively several genes coding for enzymes involved in the TCA cycle but also a positive effect on genes coding for membrane proteins (*opp4A/D*) or enzymes contributing to valine/leucine/isoleucine synthesis pathways. (9) RNAIII is a bi-functional RNA, which contains its own ORF encoding hemolysin-o, and possessing regulatory properties: RNAIII acts at the post-transcriptional level on *hla* and *spa* mRNA by base-pairings. The effect on target mRNAs translation is either positive (*hla*) or negative (*spa*, *rot*). *Rot* inhibits transcription of *hla* and *spa* genes. (10) *SprD* negative regulation of *sbi* mRNA translation occurs via the ncRNA *SprD*, but hypothetically also via RNAIII. (11) The long ncRNA *SSR42* plays potentially an indirect role on surface proteins like protein A. All the references are given in the text. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

wild-type (WT) strain like RNAIII or were up-regulated in the SCV like RsaA and the pH responsive riboswitch element *yybP-ykoY*. Six deregulated candidate sRNAs encoded from pathogenicity islands or inserted phage sequences were also identified. An SCV-specific sRNA *Sau-66*, not expressed in the WT strain, may have a role in controlling inorganic ion transport through antisense control of SA0671. Whether the deregulation of sRNAs is a consequence or is responsible for the development of the SCV phenotype in *S. aureus* has to be clarified.

## 2.2. sRNA and the TCA cycle

The tricarboxylic acid (TCA) cycle is one of the pathways leading to the production of biosynthetic intermediates required for the synthesis of macromolecules. When environmental conditions change and nutrients become growth limiting, *S. aureus* increases TCA cycle activity to catabolize nonpreferred carbon sources, i.e., acetate. Numerous studies have shown that the TCA cycle has an

impact on the production of virulence factors through the inhibition of biosynthetic enzyme activity but also by regulating the transcription of target genes (Somerville and Proctor, 2009). Interestingly, two studies have shown that RsaE, a small non-coding RNA, is involved in the regulation of several metabolic pathways (Bohn et al., 2010; Geissmann et al., 2009). RsaE is conserved in *Staphylococcus*, *Micrococcus* and in *Bacillus* (Geissmann et al., 2009), all derive from a common Gram-positive ancestor. Using combined transcriptome/proteome analysis, it was demonstrated that RsaE regulates enzymes involved in amino acid and peptide transport, cofactor synthesis, lipid, purine and carbohydrate metabolisms, and the TCA cycle (Fig. 1). Consistently, overproduction of RsaE induced a growth defect that is partially alleviated by acetate (Bohn et al., 2010). For three targets, *oppB* (amino acid and peptide transporter), *sucD* (succinyl-CoA synthetase of the TCA cycle) and SA0873 (unknown function) mRNAs, direct repression was proposed to occur by the formation of base-pairings between a conserved and unpaired C-rich motif of RsaE with the SD sequence

of these mRNAs to inhibit translation (Geissmann et al., 2009). More recently, new RsaE target candidates have been identified using bioinformatics combined with high throughput sequencing approaches (Lioliou et al., 2012, 2013). In many clinical strains, RsaE is expressed at the exponential phase and repressed at the stationary phase of growth (Bohn et al., 2010; Song et al., 2012). Hence, RsaE would facilitate the transition of energy metabolism, purine biosynthesis and amino acid transport in response to nutrient availability, and would contribute to the cell reprogramming during the growth phase. However, in RN6390, a  $\sigma^B$ -deficient strain, the expression of RsaE was enhanced at the stationary phase of growth in a manner dependent on the agr system (Geissmann et al., 2009). Due to the interplay between metabolism and virulence factor production (Somerville and Proctor, 2009), RsaE is expected to play an indirect role in virulence (Fig. 1).

### 2.3. SigmaB-dependent sRNAs

The stress-dependent activation of Sigma B has been the subject of intensive works in order to understand the influence of environmental conditions on virulence factor biosynthesis. Sigma B was found activated under general stress conditions, upon growth phase transitions and during morphological changes of *S. aureus* (reviewed in Van Schaik and Abee, 2005). Sigma B is also involved in the regulatory networks controlling the expression of virulence determinants, modulates antibiotic resistance and biofilm formation. Hence the Sigma B regulon includes many different cellular processes. Sigma B regulates the expression of genes by binding to a consensus promoter sequence upstream of the target genes. However, it was also found that many of these Sigma-B-dependent genes are not preceded by a consensus promoter sequence, implicating a role for other downstream regulators. This is why several teams searched for sRNAs that might be part of the Sigma B regulon (Geissmann et al., 2009; Nielsen et al., 2011) (Fig. 1). Among them, *rsaA* gene is under the control of a typical Sigma B-dependent promoter (Geissmann et al., 2009). Recent work suggests that RsaA regulates the translation of a master regulatory protein to regulate capsule and biofilm formation (Romilly et al., unpublished data). Prediction of Sigma B-promoters within intergenic regions of *S. aureus* revealed three other sRNAs, which were induced by KOH treatment. Two of them might be bifunctional RNAs because they potentially encode small and highly basic peptides, which are conserved among staphylococci (Beaume et al., 2010; Nielsen et al., 2011). The third encodes a genuine asRNA, which interacts in vitro with the 3'-end of the downstream gene on the opposite strand (Nielsen et al., 2011). Although antisense regulation has not yet been demonstrated in vivo, these data suggested that SbrC would regulate the synthesis of an ABC transporter specific for the uptake of manganese that is important for effective host colonization by *S. aureus* (Nielsen et al., 2011). Global approaches such as chromatin immunoprecipitation followed by HTS would be useful to reveal the whole set of sRNAs that are part of the Sigma B regulon.

### 2.4. Expression of sRNAs under antibiotic exposure

Antibiotic treatment can be regarded as one of the most selective pressures on *S. aureus*, including within the host environment. As a major public health problem, some *S. aureus* clones have become resistant to most antimicrobial agents including the commonly used penicillin-related antibiotics. Multi-drug resistant strains of *S. aureus* have been now reported with increasing frequency worldwide. Strains that are resistant to methicillin (methicillin-resistant *S. aureus*; MRSA) are often also resistant to lincosamides, macrolides, aminoglycosides, fluoroquinolones, or combinations of these antibiotics. More recently, intermediate

resistance to vancomycin – a glycopeptide antibiotic that is one of the last line agents active against MRSA – has been reported globally. Community-associated MRSA (CA-MRSA) strains seem also to be particularly virulent (Chambers and Deleo, 2009). Recently, the transcriptional response of sequence type 239 (ST239) multidrug resistant *S. aureus* to antibiotic exposure was investigated using Illumina HTS, with a particular focus to decipher the potential role of sRNAs in antimicrobial response (Howden et al., 2013). In this study, four antibiotics were used (vancomycin, linezolid, ceftibiprole and tigecycline). A large number of putative sRNAs were identified in this strain (409 in total), and 42% of them (171) were antisense to mRNAs. Seventy-nine of the sRNAs were encoded from the *S. aureus* accessory genome, indicating that 330 core genome encoded putative sRNAs are conserved in all *S. aureus* included in the analysis (19 strains). Upon drug exposure, the most significant changes were differential regulation of asRNAs, predominantly targeted against genes encoding proteins involved in translation and ribosomal function, such as subunits of RNA polymerase, translation elongation factors and ribosomal proteins. Linezolid, an oxazolidinone antibiotic that acts by binding to 23S rRNA, produced the most profound transcriptional responses. These results suggested a contribution, through control of protein synthesis and of asRNAs to the bacterial response to antimicrobials. In addition, one putative sRNAs (sRNA260) and one experimentally validated sRNA RNA131 (also called RsaOI) were upregulated. RsaOI is a small non-coding RNA conserved in all staphylococcal strains and its synthesis accumulates at the stationary phase of growth (Marchais et al., 2009). RsaOI was proposed to adopt a particular structure involving two long distant interactions (Marchais et al., 2010). Intriguingly, *rsaOI* gene is located in the vicinity of the *vra* operon that is upregulated in some vancomycin intermediate *S. aureus* strains (Howden et al., 2013). Ongoing work is underway to definitely link sRNA expression to drug resistance.

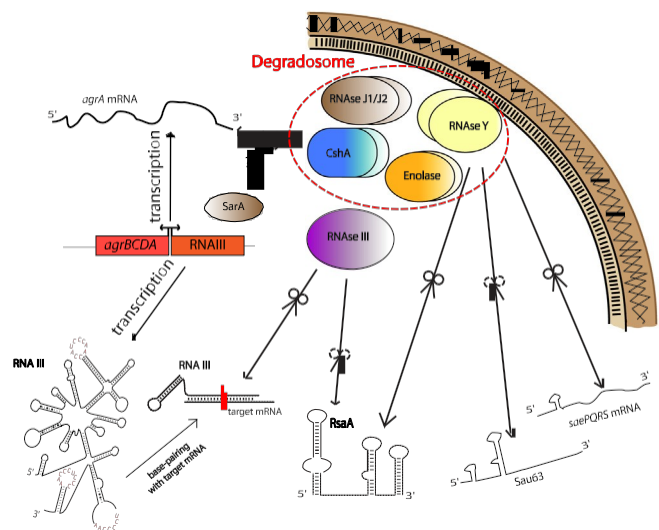


Fig. 2. Interplay between RNAs and RNA-binding proteins in *Staphylococcus aureus*. The major complex of RNA degradation, the so-called degradosome (circled in red), comprises the membrane-anchored RNase Y, the two non-essential RNase J1/J2, the RNA helicase CshA and the central metabolism-related enolase. PNPase and phosphofruktokinase are also part of the complex but not shown for figure simplicity (see Roux et al., 2011 for further details). RNase Y is involved in the degradation of several mRNAs (*saePQRS* mRNAs shown for example) and ncRNAs RsaA and sau63 (Marincola et al., 2012). The double-stranded RNA specific RNase III is not associated with the degradosome but is nonetheless one of the most important enzymes involved in RNA degradation. The *agrBCDA* mRNA degradation is mediated by CshA, the RNA helicase component of the degradosome (Oun et al., 2013). It has also been shown that the transcriptional factor SarA has a positive effect on *agrA* mRNA stability but whether it is indirect by inhibiting degradation by CshA is yet to be elucidated (Morrison et al., 2012a).

### 3. sRNAs acting directly on virulence gene expression

Regulatory RNAs dedicated to the regulation of virulence genes are thought to be acquired more recently in the evolution of the bacteria, and could be considered as key adaptors to the host (Queck et al., 2008). In this paragraph, we will highlight recent studies to illustrate the direct role of sRNAs in the virulence of *S. aureus*, their integration in pre-existing regulatory circuits and the diversity of their expression according to the strain and/or infection, that may contribute to the adaptation of the bacteria to the variability of the host responses.

#### 3.1. Two intertwined regulatory networks quorum sensing and virulence factor control

Quorum sensing is widely used by pathogens to regulate virulence gene expression (Bassler and Losick, 2006). In *S. aureus*, quorum sensing is mainly accomplished by the *agr* system. This system is composed of two divergent transcripts, RNAII encoding a quorum sensing cassette and a two-component system, and RNAIII encoding hemolysin delta. The quorum sensing cassette produces the autoinducer peptide, AIP, which upon a threshold concentration, activates the membrane kinase AgrC and the response regulator AgrA through a phosphorylation mechanism. AgrA induces the expression of RNAII, and RNAIII. While most studies were dedicated to decipher the function of RNAIII, the main intracellular effector of gene virulence (Novick et al., 1993; Novick, 2003), a recent study on a methicillin-resistant *S. aureus* strain revealed that the *agr* system is composed of two independent regulatory circuits, with two principal effectors, AgrA and RNAIII (Queck et al., 2008). While RNAIII coordinates the expression of virulence factors, AgrA up-regulates Phenol-soluble-modulin (PSM) cytolytic genes by direct binding to promoters and down-regulates genes involved in carbohydrate and amino acid metabolism by an unknown mechanism (Queck et al., 2008). Thus the AgrA-dependant regulatory circuit may be reminiscent of the ancient function of the quorum sensing system, in a non-pathogenic life style of *S. aureus*, where *agr* enables cells to adapt to low nutrition conditions in high cell density population. Otherwise not only the *agr* operon is controlled by transcriptional factors like SarA, MgrA, but *agrA* mRNA itself is negatively regulated by different factors (see Figs. 1 and 2 and text below). The *agr* system is confirmed as a pivotal element in *S. aureus* virulence allowing the bacteria to integrate a multitude of signals.

#### 3.2. *S. aureus* RNAIII regulates the temporal expression of virulence determinants

The bifunctional RNAIII encodes a PSM, delta hemolysin. If both delta hemolysin and RNAIII play roles in virulence, they do not act synergistically. Interestingly, *hld* expression is delayed by 1 h after the transcription of RNAIII (Novick, 2003). This is not in line with the accessibility of the ribosome binding site (RBS) sequence of *hld* since it was shown that ribosomes bind to RNAIII in vitro (Béni-to et al., 2000). Thus, it is possible that a trans-acting factor could be involved either to repress *hld* translation during the first hour, or to activate its translation. As a regulator, RNAIII promotes the switch between the expression of surface proteins and the synthesis of excreted toxins in a time-dependent manner (Novick, 2003). The regulation mechanisms by RNAIII have been extensively documented in the last years (reviewed in Caldelari et al., 2011; Romby and Charpentier, 2010; Romilly et al., 2012a). It controls gene expression at the post-transcriptional level via an antisense mechanism. The 5'UTR of *hld* binds to the leader region of *hla* mRNA to facilitate ribosome recruitment while the large 3'UTR is acting

primarily as a repressor domain. The 3'UTR contains three redundant hairpin structures with a conserved C-rich sequence located in the apical loops. This motif is often used to bind the G-rich sequences of the ribosome binding sites of mRNAs. The duplexes formed between the RNAIII and the targeted mRNAs prevent ribosome binding and recruit the double-strand specific RNase III, which initiates the rapid degradation of both the mRNAs and the RNAIII (Boisset et al., 2007; Chevalier et al., 2010; Huntzinger et al., 2005). These mRNAs encode virulence factors expressed at the surface of the cell (protein A, coagulase, SA1000), and the transcriptional repressor of toxins, Rot. Through the inhibition of Rot, RNAIII indirectly activates the transcription of exotoxins and represses the transcription of adhesins such as protein A (Boisset et al., 2007; Geisinger et al., 2006). Thus, RNAIII acts directly at the post-transcriptional level by targeting mRNAs, and indirectly at the transcriptional level by repressing the translation of Rot, through the formation of two coherent feed-forward regulatory loops as exemplified in Fig. 1. Such networks ensure a tight regulation of gene expression regulation (Shimoni et al., 2007, 2009). They may also explain the delay between repression of adhesin synthesis, which occurs directly through RNA-RNA pairing, and the induction of exoprotein synthesis, which requires first the inhibition of rot translation (Novick, 2003). Interestingly, RNAIII also represses several hydrolases and amidases involved in the metabolism of peptidoglycan and hence it may contribute to the cell wall integrity at high cell density (Boisset et al., 2007; Chunhua et al., 2012; Dunman et al., 2001) (Lioliou et al., personal communication). Regulation of peptidoglycan metabolism, and not virulence only, might be the common function of RNAIII in all staphylococcal species including *Staphylococcus epidermidis*.

RNAIII is thus a multi-functional RNA and this is probably not an exception in *S. aureus* (Kaito et al., 2011; Fig. 1). Although it is clear that RNAIII regulates gene expression at the post-transcriptional level, the whole set of direct targets has not yet been fully elucidated. Strong evidences suggested that RNase III acts as a co-factor of RNAIII to induce rapid degradation of repressed target mRNAs. However, no protein such as Hfq has been found to be involved in the formation of RNAIII-mRNA complexes, or in RNAIII stability. Do the two functions of RNAIII, synthesis and regulation, interfere with each other? Recent work suggested that part of the coding sequence of *hld* could bind to map mRNA encoding the major histocompatibility complex class II analogous protein to activate translation (Liu et al., 2011). In addition, part of the ribosome binding site of *hld* is also probably engaged in base pairings with *hla* mRNA. Therefore activation of translation of target mRNAs and *hld* translation are most probably two exclusive functions of RNAIII. Finally, although all sequenced *S. aureus* strains contain the RNAIII locus, the yield of RNAIII strongly varies. The regulators involved in the transcriptional control or its stability are not fully understood although it is well known that many transcriptional regulators converged to the *agr* system to modulate the virulence factor production (Figs. 1 and 2). This might confer the advantage to the bacteria to respond to multiple environmental cues in addition to cell density (Priest et al., 2012).

#### 3.3. SprD mediates a crosstalk between pathogenicity island and the core genome to regulate virulence

SprD is the second sRNA emerging as an important player in the regulation of virulence. It is a 142-nts long sRNA expressed from a pathogenic island, PI/ (Chabelskaya et al., 2010; Pichon and Fel-den, 2005) and interacts through base pairings with the *sbi* mRNA, which encodes an immune evasion molecule protecting the bacteria against host immune responses. Interestingly, *sbi* is encoded on the core genome, suggesting that SprD mediates connections between the pathogenic island and the core genome. SprD contains

Table 1  
Summary of experimentally validated sRNAs in *S. aureus*.<sup>a</sup>

Study	Strain used	sRNA discovery methodology (number of in silico predicted sRNAs)	Experimentally validated sRNAs <sup>c</sup>	Experimental validation method	Target and mechanism	Comment
Roberts et al. (2006)	UAMS-1	Gene chip analysis	SSR42	NB, RT-PCR	Spa, hla, hglC, lukF Unknown	
Pichon and Felden (2005) and Sayed et al. (2011)	Mu50 (clonal complex 5)	Bioinformatic predictions (xx)	SprA SprA2, SprA3, SprB, SprC, SprD, SprE, SprF, SprG, SprFG2, SprFG3 4.5S RNA, 6S RNA, RNAlII, tmRNA, RNase P	NB <sup>b</sup> NB <sup>b</sup> NB <sup>b</sup>	ABC transporter (SA2216), possible antisense sRNA Unknown, SprA2 encodes a cytolytic peptide Housekeeping ncRNAs	
Geissmann et al. (2009)	RN6390, Col, Newman, HG001	Bioinformatic predictions and experimental validation	RsaE RsaA, RsaB, RsaC, RsaD, RsaF, RsaG, RsaH, RsaI, RsaJ, RsaK	NB, PE, RACE NB, PE, RACE	Masking of ribosomal binding site for oppB, sucD, SA0873 Unknown	Genetic manipulation demonstrated a role for RsaE in controlling metabolic pathways
Marchais et al. (2009)	N315 (clonal complex 5)	Bioinformatic (NAPP) (189) and Northern analysis.	RsaOA, RsaOB, RsaOC, RsaOD, RsaOE, RsaOF, RsaOG <sup>b</sup>	NB	Unknown	
Nielsen et al. (2011)	N315 (clonal complex 5)	Bioinformatic search for intergenic $\Gamma^B$ consensus sites and experimental validation	SbrA (RsaOO), SbrB, SbrC	NB	$\Gamma^B$ regulated. SbrC interacts with SA0587 (mntC)	sbrA and sbrB potential CDS
Abu-Qatouseh et al. (2010)	–	Cloning and sequencing of cDNAs	Ssr-72, Ssr-80, Ssr-87 Sau-02, Sau-13, Sau-19, Sau-24, Sau-27, Sau-30, Sau-31, Sau-41, Sau-50, Sau-53, Sau-59, Sau-63, Sau-64, Sau-66, Sau-5949, Sau-5971, Sau-6053, Sau-6072	NB <sup>c</sup> NB	Unknown Unknown. Sau-66, putative posttranslational control of antisense gene SA0671	142 candidate sRNA identified
Bohn et al. (2010)	N315 (clonal complex 5)	454 pyrosequencing followed by experimental validation	RsaON (RsaE) <sup>d</sup> RsaOH, RsaOI, RsaOL, RsaOM, RsaOO, RsaOP, RsaOQ, RsaOR, RsaOT, RsaOU, RsaOV, RsaOW, RsaOX	NB NB	Binds opp3A mRNA ribosome-binding site. Overexpression of RsaE reduces central metabolic pathways and increases amino acid pool	30 sRNAs identified, 14 new
Beaume et al. (2010)	N315 (clonal complex 5)	Illumina sequencing	Teg1, Teg4, Teg17, Teg18, Teg19b, Teg21, Teg24, Teg26, Teg28, Teg35, Teg38, Teg42, Teg45, Teg47, Teg55, Teg56, Teg57, Teg60, Teg61, Teg69, Teg70, Teg72, Teg73, Teg76, Teg91, Teg2pl	RT-PCR	Unknown	195 sRNAs predicted by HTS

<sup>a</sup> Validated using Northern blot, RNA extremity mapping, or RT-qPCR.

<sup>b</sup> Also experimentally validated using Northern blot in the study by Abu-Qatouseh et al. (2010).

<sup>c</sup> Originally described by Anderson et al. (2006).

<sup>d</sup> Previously described and validated. NB, Northern blot; PE, primer extension; RACE, random amplification of cDNA ends; RT-PCR, real-time PCR. Several long and stable RNAs (SSR) expressed under particular conditions of growth have been assigned by microarrays Roberts et al. (2006).

<sup>e</sup> A number of different names have been assigned to many of these sRNAs. In this table the initial name from the relevant study has been used. A full list of alternate names can be found in the supplementary table from Felden et al. (2011). In the nearest future, it would be appropriate to adopt a common way to annotate the sRNA genes.

four hairpin structures and the third one interacts with the ribosome binding site of the *sbi* mRNA to form a long duplex that prevent translation initiation in vitro (Fig. 1). As for RNAlII, SprD does not require trans-acting factors like Hfq, but *sbi* mRNA degradation is not promoted by SprD expression. Interestingly enough, the yield of *sbi* mRNA strongly decreased at the stationary phase of growth independently of SprD (Chabelskaya et al., 2010). This data suggested that a second regulatory event might take place. Finally, SprD contributes to *S. aureus* diseases in a mouse model of infection, but this effect was not linked to the repression of Sbi production, indicating that SprD might regulate the expression of other proteins important for infection.

#### 3.4. A pathogenicity-island sRNA attenuates MRSA virulence

In the USA, community acquired methicillin resistant *S. aureus* (CA-MRSA) has been shown to be more virulent than hospital associated MRSA. Recent work discovered that this behavior is at least partly linked to the existence of a transcription product of the *psm-mec* gene which is located in the SCCmec of the hospital acquired

MRSA (HA-MRSA) but not in the equivalent SCCmec locus of CA-MRSA (Kaito et al., 2011, 2013). *Psm-mec* is a bifunctional RNA, which encodes a phenol-soluble modulins (PSM), a cytolytic toxin. In addition, the *psm-mec* RNA inhibits translation of *agrA* mRNA by direct base-pairings, with the coding sequence of the mRNA (Kaito et al., 2013). It remains to be determined how *psm-mec* RNA is able to repress the translation of *agrA* mRNA. It is more likely that duplex formation might recruit a ribonuclease to initiate mRNA degradation. Deletion of the *psm-mec* RNA from HA-MRSA strains increased the expression of *AgrA* which results in an increase of toxin production and of PSM and enhanced virulence in mice (Kaito et al., 2013). Interestingly 25% of 325 clinical MRSA isolated in Japan contained a mutation in the *psm-mec* promoter and 9% did not carry *psm-mec* leading to high level of PSM in the supernatant. Identification of the *psm-mec* mutation could be a novel method to predict virulence properties of MRSA strains. This study shows that another bi-functional RNA is able to affect virulence through *AgrA* inhibition, and illustrates how the acquisition of novel genetic elements and of sRNA alters the behavior of the bacteria.

### 3.5. Another long sRNA contributes to pathogenesis

A subset of RNA molecules, called small stable RNAs (SSRs), has been previously identified by microarrays (Anderson et al., 2006; Roberts et al., 2006). These SSRs are either transcribed or stabilized in response to growth phase, stringent, heat and cold shocks, alkaline and acidic conditions. Among these RNAs, stability of SSR42 is greatly enhanced under stationary phase of growth (Anderson et al., 2006; Morrison et al., 2012b). This 891-long RNA affects erythrocyte lysis, resistance to opsonisation killing, and pathogenesis in a murine model of skin and soft tissue infection (Morrison et al., 2012b). The RNA upregulates genes involved in capsule biosynthesis, and downregulates 80 genes including several cell surface virulence factors. While no direct binding was demonstrated between SSR42 and mRNA encoding virulence determinants, the effect is predicted to be indirect by modulating the expression of a transcriptional regulator, and more experiments would need to be performed to decipher the molecular mechanisms of regulation. The genomic localization of SSR42 does not exclude that the RNA might regulate the upstream and divergent gene through an antisense mechanism, because the two genes are separated only by few nucleotides and recent highthroughput sequencing analysis suggested that the 3' end of SSR42 overlaps with the end of the coding sequence of the upstream gene (Lioliou et al., 2012). This long RNA was co-immunoprecipitated with the endoribonuclease III (RNaseIII) together with an asRNA at the exponential phase of growth in RN6390 (Lioliou et al., 2012). While SSR42 is unstable at the exponential phase of growth (Morrison et al., 2012b), it is tempting to propose that antisense regulation mediated through RNase III would regulate either the stability of SSR42 or of the flanking genes in a growth phase dependent manner. Recent tiling arrays analysis discovered that *S. aureus* genome expressed long antisense RNAs that are rapidly degraded by RNase III (Lasa et al., 2011). Such asRNA regulation is expected to play important regulatory functions (reviewed in Lasa et al., 2012).

### 3.6. Expression of sRNAs in *S. aureus* during infection

Most sRNAs have been identified in a few strains, and their expression profiles during infection in human have been studied for only a subset of sRNAs including RNAIII. The vast majority of clinical isolates from acute infections express RNAIII (Traber et al., 2008), but the level of RNAIII expression varies among isolates (Jelsbak et al., 2010). However, agr-mutant strains, which most probably arose during infection, were isolated from patients (Traber et al., 2008). Interestingly some of the agr-mutant strains have been associated with persistent bacteremia and dormant states, notably for methicillin-resistant *S. aureus* infection (Fowler et al., 2004; Sakoulas et al., 2002). These observations are also supported by the fact that the SCV phenotype is deficient in RNAIII synthesis (Proctor et al., 2006). Thus, the appearance of agr mutations might reflect an adaptation of the pathogen within the infected host.

Recently Song et al. (2012) compared the expression profile of 5 sRNAs (RNAIII, RsaA, RsaE, RsaG, RsaH) of 60 isolates from acute and chronic infections, and from nasal colonization. The expression levels of the sRNAs were measured in the clinical samples and compared with those of the corresponding strains grown in rich culture medium. They showed that all studied sRNAs are expressed in human, and that their expression is strain specific, with consequent variation among strains. Particularly, their levels were highly variable in the abscessed patients, more homogeneous in the cystic fibrosis patients, and highly uniform in the nasal carriers samples. The comparison of the sRNA expression in infection versus colonization specimens suggests that the sRNA expression levels observed in infection would be a consequence of the great

variability of the host response, while colonization is apparently a more uniform scenario.

## 4. RNA-binding proteins

### 4.1. Hfq protein, a mysterious protein in *S. aureus*

There is currently very limited information on staphylococcal RNA-binding proteins that might be involved in sRNA regulation. Hfq is present in most of the staphylococcal genomes, but differentially expressed (Liu et al., 2010). Its role in vivo is still controversial, while an hfq deletion mutant has no significant phenotype (Bohn et al., 2007; Romilly et al., 2012a). Several sRNAs, such as RNAIII or RsaE, and SprD, repress their target mRNAs either by forming long irregular duplexes, or by forming bi-partite binding sites involving distant and unpaired regions of both RNAs. In all cases, fast binding is initiated by G-C base pairs present in loop regions (Fig. 1). Hence, long complementarities or bipartite binding sites might compensate the need of Hfq. The dispensability or requirement of Hfq might have resulted from numerous specific features of Hfq, sRNAs and their RNA duplexes, which still have to be clearly addressed. Although the crystal structures of *S. aureus* Hfq associated with short RNA sequences were the first to be solved (Brennan and Link, 2007), its function still remains mysterious. The protein adopts an homohexamer structure organized into a doughnut with a proximal and distal faces recognizing unpaired uridine-rich and adenine-rich sequences, respectively (e.g., Brennan and Link, 2007; Horstmann et al., 2012; Link et al., 2009; Mikulecky et al., 2004; Someya et al., 2012; Soper et al., 2011). In enterobacteriaceae, the proximal and distal faces of the Hfq hexamer specifically bind sRNA and mRNA targets. Interestingly, a recent study discovered, that a motif rich in arginines (RRER) located on the outer rim of the hexamer, are required for the chaperone activity of Hfq (Panja et al., 2013). Mutations in the arginine patch of *E. coli* Hfq affected sRNA regulation and prevented annealing of the two RNAs even though binding was not affected suggesting that the RRER motif constitutes the active site of the chaperone activity. This arginine patch is replaced by a KANQ motif in *S. aureus* Hfq (Panja et al., 2013). It remains to be analyzed whether this sequence difference explains why Hfq is not required for many sRNAs in *S. aureus*. These different RNA binding sites might explain why *S. aureus* Hfq cannot complement *Salmonella typhimurium* or *E. coli* hfq mutant strains (Rochat et al., 2012; Vecerek et al., 2008).

### 4.2. Ribonucleases as possible co-factors of sRNAs

The role of several ribonucleases has been recently well documented in *S. aureus* virulence. RNase III is responsible to degrade double-stranded RNA formed by base-pairing of RNAIII with its mRNA targets such as spa, coa and rot (Fig. 1). The enzyme recognizes various topologies, rather than specific sequences and cleaves at sites such as imperfect duplexes, helices interrupted by bulged residues and kissing loops between RNAIII and coa or rot mRNAs by examples (Romilly et al., 2012b). Moreover, in two different studies, RNase III has also been clearly identified as a major partner in antisense regulation, beside its known role in mRNA and sRNA turnover and rRNA or tRNA maturation (Lasa et al., 2011; Lioliou et al., 2012). At least four key types of antisense transcription have been defined in *S. aureus* (see Fig. 3): (a) bona-fide short and long antisense RNAs, (b) overlapping 5' or 3'-untranslated region of divergent or convergent mRNAs, (c) overlapping operons and (d) type I toxin-antitoxin modules (Lasa et al., 2012). The biological significance of such overlapping transcription is still unclear, but the role of RNase III in antisense regulation seems to happen most prominently in Gram-positive bacteria such as *S. aur-*



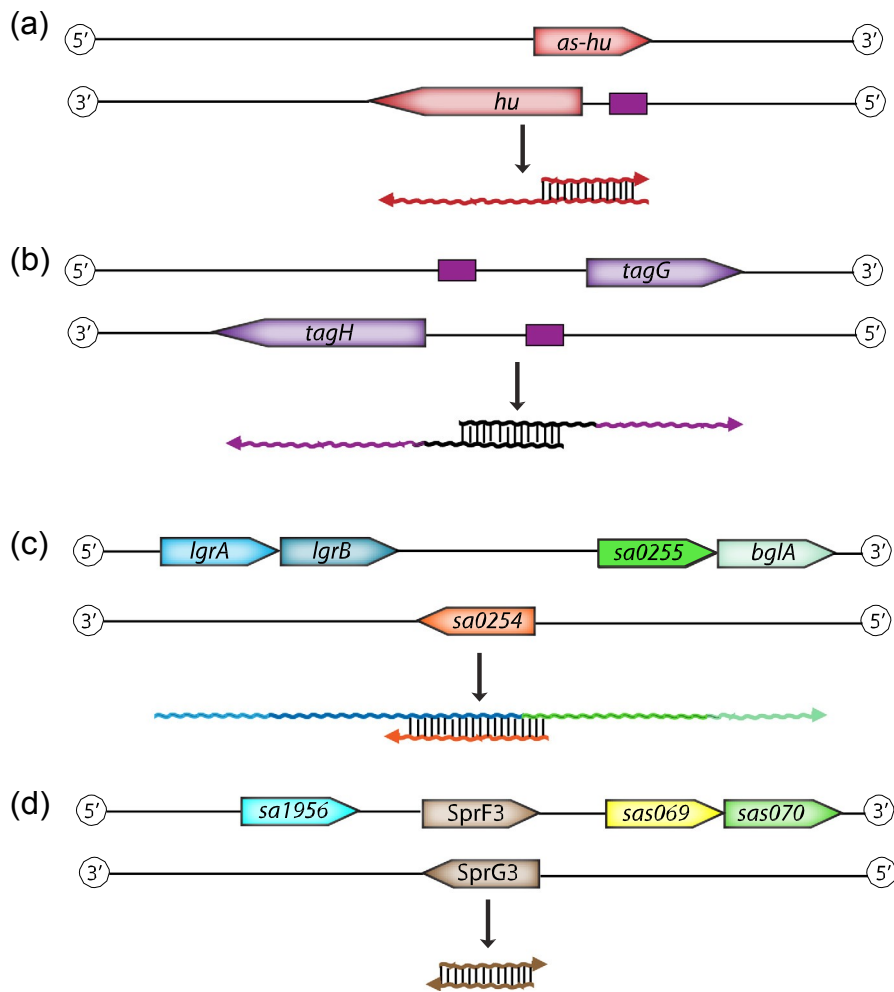


Fig. 3. Four mechanisms to generate overlapping transcription in *S. aureus*. Purple boxes represent the transcriptional start sequences and genes annotations are according to the *S. aureus* N315 genome. (a) Bone fide antisense RNA (asRNA): *hu* mRNA may be rapidly degraded by the dual action of the *hu* asRNA and RNaseIII (Lioliou et al., 2012). (b) Overlapping 5' untranslated region: RNase III cleaves the overlapping 5' UTR of the *tagG/tagH* resulting in the formation of a *tagH*mRNA with a shorten leader (Lioliou et al., 2012). (c) Overlapping operon: *sa0254* is located in the middle of the operon *lgrA**Bsa0255bglA* and is transcribed in opposite direction to the rest of the operon (Lasa et al., 2011). (d) Type I toxin-antitoxin: *SprF3/SprG3* may belong to the group I toxin-antitoxin systems, with *SprG* being the putative toxin. Whether *SprG3* encodes a peptide is currently unknown (Pichon and Felden, 2005). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

eus (Lasa et al., 2011). Indeed, RNase III is responsible to cleave sense-antisense RNAs duplexes leading to accumulation of 20 nucleotides fragments from all over the chromosome. At least 75% of mRNAs were subject to this processing, which reflects the importance of the phenomenon that might function as a general purpose global control on bacterial gene expression.

Regarding the membrane-bound endoribonuclease RNase Y, it cleaves unpaired regions of RNA and was shown to be part of the degradosome-like complex together with RNases J1/J2, PNPase, enolase, phosphofructokinase, and a DEAD box RNA helicase in *S. aureus* (see below, and Roux et al., 2011) (Fig. 2). RNase Y was shown to affect production of virulence factors either through expression of the accessory gene regulatory locus, *agr* or independently (Kaito et al., 2005; Nagata et al., 2008). Recently Marincola et al. (2012) performed a transcriptomic analysis to decipher the effect of RNase Y on global gene regulation. Alongside its major role in the processing of virulence genes, the enzyme is involved in the turnover of two sRNAs *RsaA* and *Sau63* (Fig. 2). Indeed, half-life of both sRNAs was increased more than 5 times in the my mutant compared to the wild type strain. Interestingly, we could also show that RNase III cleaves *RsaA* at two sites in the bulged loop of its 5' hairpin structure and that *RsaA* is more stable in the *mc* mutant strain compared to the wild-type strain (Lioliou et al., 2012) sug-

gesting that RNase III contributes to *RsaA* maturation. Nevertheless, we cannot exclude that like many ncRNAs, *RsaA* regulates expression of its target by an antisense mechanism, and RNase III would be a partner to induce degradation of the duplex between *RsaA* and mRNAs.

The *CshA* DEAD-box helicase was reported to interact with proteins from the degradosome (Roux et al., 2011). It has the typical RNA-dependent ATPase activity of DEAD-box proteins, and plays a role in the steady-state level and stability of *agrA* mRNA, which indirectly affects the level of RNaseIII and the virulence of the bacterium (Oun et al., 2013) (Fig. 2).

#### 4.3. A transcriptional regulator affects RNA stability

Very recently, the pleiotropic transcriptional regulatory protein *SarA* was unexpectedly identified as an RNA binding protein. It altered mRNA turnover by stabilizing few transcripts associated to virulence, like *clpP*, *spa* or *agrA* in the late exponential or stationary phase of growth, suggesting that binding of the protein would protect mRNAs from degradation (Morrison et al., 2012a; Roberts et al., 2006). One cannot exclude that the effect on *agrA* mRNA might be indirect by inhibiting its degradation by *CshA* (see above and Fig. 2). Moreover, using *in vitro* and *in vivo* approaches, *SarA*

protein was found to bind spa mRNA in a cooperative manner. It was also found to destabilize mRNAs or bind to transcripts that mapped to intergenic regions, which might encode sRNAs. Thus, SarA modulates the production of proteins at the mRNA level via both transcriptional and post-transcriptional pathways. The use of UV cross-linking and immunoprecipitation assay (CLIP) would be appropriate to identify direct interaction sites between SarA and RNAs in vivo (Wang et al., 2009).

## 5. Concluding remarks

A rather good knowledge on the sRNA repertoire in specific *S. aureus* strains is now available but significant work is still required to gain a full appreciation of the sRNA functions, of their regulatory networks and their interactions with proteins. The few examples here demonstrate already how sRNAs modulate the expression of key transcriptional factors and how several sRNAs converged to form extensive regulatory networks to regulate virulence factors (Fig. 1). One of the future tasks will be to understand how the different networks are interconnected, and how they influence each other to develop variations in the pattern of virulence gene expression. The use of high-throughput sequencing approaches have revealed hundreds of unexpected transcripts in both sense and antisense orientations, and numerous short transcripts derived from long 5' or 3' UTRs, or internal regions of mRNAs have been detected. Whether they behave as independent regulatory RNAs requires experimental validation. Other sRNAs encode peptides that may play a role in virulence or in growth control. Continued mechanistic and functional studies of individual sRNAs and their machineries will certainly unravel unexpected functions and regulatory mechanisms. For instance, the asRNA SprA1-AS, which regulates the expression of a cytolytic peptide SprA1, binds to the target RNA in a very unusual way (Sayed et al., 2011). Unexpectedly, the interacting region does not involve the 3' end of SprA1-AS, which is fully complementary to the 3' end of the target SprA1 RNA but instead, the active region of the asRNA is in its 5' part that is partially complementary to the ribosome binding site of SprA1. These findings demonstrate that the structures of both RNAs were determinant for the regulatory mechanism.

We illustrate here how *S. aureus* sRNAs interconnect metabolism, stress adaptation, cell-to cell communication and virulence, and how they are able to mediate crosstalks between pathogenicity islands and the core genome. Other avenues can be further explored such as modulation of the bacterial chromosome structure, control of mobility of pathogenicity islands, and host-pathogen interaction. One critical point is the fact that it is difficult to generalize the impact of sRNA regulation in *S. aureus* due to great genetic and phenotypic variability between the strains. In addition, the expression of sRNA in *S. aureus* grown in rich medium culture does not reflect the in vivo conditions. Hence, a major challenge is to understand the contribution of *S. aureus* sRNAs in the human context and to gain knowledge about the sRNA roles and their co-factors during human colonization and infection (Westermann et al., 2012).

*S. aureus* resistance to antimicrobial drugs has become a major health problem. Mechanisms of bacterial defense against antibiotics include multiple strategies such as mutation of the drug target, inactivation or destruction of the agent, inhibition of cell division during exposition to certain antibiotics and repression of transport systems to avoid the antimicrobial entry. Whether sRNAs are involved in the control of *S. aureus* adaptability to antimicrobial agents is still poorly understood. More recently, an aminoglycoside-binding riboswitch that is widely distributed among antibiotic-resistant bacterial pathogens as *S. aureus*, has been discovered. This riboswitch is present in the leader regions of

mRNAs and activates the production of enzymes that confer resistance to aminoglycoside antibiotics when these are present in the environment (Jia et al., 2013). The authors proposed a model of an antibiotic-sensing riboswitch, whose structure will change upon drug binding unmasking the RBS and leading to mRNA translation. This simple mechanism allows the cells to rapidly respond to the antibiotic killing effect and might be exploited by other RNA-binding antibiotics. Conversely, recent promising works have shown that metabolite-binding riboswitches can be exploited to avoid multiple drug resistance. For instance, in *S. aureus*, several non-metabolizable agonists were designed to constitutively switch off the expression of the essential *guaA* gene encoding GMP synthase, which is under the control of a guanine-sensing riboswitch. These compounds inhibit growth and reduce *S. aureus* infection in mice (Mulhbacher et al., 2010a). Hence, manipulating individual sRNA regulatory pathways opens new avenues to explore novel ways to combat *S. aureus* diseases. There are still many exciting discoveries and unanswered questions to be solved on *S. aureus* sRNAs.

## Acknowledgements

We are thankful to all other members of the team for helpful comments on the manuscript. We are grateful to François Vandenesch and to Thomas Geissmann (University of Lyon, INSERM) for helpful discussions and support. This work was supported by the Centre National de la Recherche Scientifique (CNRS, PR), the Agence Nationale de la Recherche and the “Laboratoires d’excellence” (LABEX) NetRNA Grant ANR-10-LABX-36 (PR) in the frame of «Programme d’Investissements d’Avenir», and the National Health and Medical Research Council (NHMRC), Australia (BPH).

## References

- Abu-Qatouseh, L.F., Chinni, S.V., Seggewiß, J., Proctor, R.A., Brosius, J., Rozhdvestvensky, T.S., Peters, G., von Eiff, C., Becker, K., 2010. Identification of differentially expressed small non-protein-coding RNAs in *Staphylococcus aureus* displaying both the normal and the small-colony variant phenotype. *J. Mol. Med.* 88, 565–575.
- Anderson, K.L., Roberts, C., Disz, T., Vonstein, V., Hwang, K., Overbeek, R., Olson, P.D., Projan, S.J., Dunman, P.M., 2006. Characterization of the *Staphylococcus aureus* heat shock, cold shock, stringent, and SOS responses and their effects on log-phase mRNA turnover. *J. Bacteriol.* 188, 6739–6756.
- Bassler, B.L., Losick, R., 2006. Bacterially speaking. *Cell* 125, 237–246.
- Beaume, M., Hernandez, D., Farinelli, L., Deluen, C., Linder, P., Gaspin, C., Romby, P., Schrenzel, J., Francois, P., 2010. Cartography of methicillin-resistant *S. aureus* transcripts: detection, orientation and temporal expression during growth phase and stress conditions. *PLoS One* 5, e10725.
- Benito, Y., Kolb, F.A., Romby, P., Lina, G., Etienne, J., Vandenesch, F., 2000. Probing the structure of RNAIII, the *Staphylococcus aureus* agr regulatory RNA, and identification of the RNA domain involved in repression of protein A expression. *RNA* 6, 668–679.
- Bohn, C., Rigoulay, C., Bouloc, P., 2007. No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*. *BMC Microbiol.* 7, 10.
- Bohn, C., Rigoulay, C., Chabelskaya, S., Sharma, C.M., Marchais, A., Skorski, P., Borezée-Durant, E., Barbet, R., Jacquet, E., Jacq, A., Gautheret, D., Felden, B., Vogel, J., Bouloc, P., 2010. Experimental discovery of small RNAs in *Staphylococcus aureus* reveals a riboregulator of central metabolism. *Nucleic Acids Res.* 38, 6620–6636.
- Boisset, S., Geissmann, T., Huntzinger, E., Fechter, P., Bendridi, N., Possedko, M., Chevalier, C., Helfer, A.C., Benito, Y., Jacquier, A., Gaspin, C., Vandenesch, F., Romby, P., 2007. *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev.* 21, 1353–1366.
- Breaker, R.R., 2009. Riboswitches: from ancient gene-control systems to modern drug targets. *Future Microbiol.* 4, 771–773.
- Breaker, R.R., 2011. Prospects for riboswitch discovery and analysis. *Mol. Cell* 43, 867–879.
- Brennan, R.G., Link, T.M., 2007. Hfq structure, function and ligand binding. *Curr. Opin. Microbiol.* 10, 125–133.
- Caldelari, I., Chao, Y., Romby, P., Vogel, J., 2013. RNA-mediated regulation in pathogenic bacteria. *Cold Spring Harb. Perspect. Med.* 3(9).
- Caldelari, I., Fechter, P., Lioliou, E., Romilly, C., Chevalier, C., Gaspin, C., Romby, P., 2011. A current overview of regulatory RNA in *Staphylococcus aureus*. In: Marchfelder, A., Hess, W. (Eds.), *Regulatory RNAs in Prokaryotes*. Wiley Verlag, Wien and New York.

- Chabelskaya, S., Gaijlot, O., Felden, B., 2010. A *Staphylococcus aureus* small RNA is required for bacterial virulence and regulates the expression of an immune-evasion molecule. *PLoS Pathog.* 6, e1000927.
- Chambers, H.F., DeLeo, F.R., 2009. Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nat. Rev. Microbiol.* 7, 629–641.
- Cheung, A.L., Bayer, A.S., Zhang, G., Gresham, H., Xiong, Y.-Q., 2004. Regulation of virulence determinants in vitro and in vivo in *Staphylococcus aureus*. *FEMS Immunol. Med. Microbiol.* 40, 1–9.
- Chevalier, C., Boisset, S., Romilly, C., Masquida, B., Fechter, P., Geissmann, T., Vandenesch, F., Romby, P., 2010. *Staphylococcus aureus* RNAIII binds to two distant regions of *coa* mRNA to arrest translation and promote mRNA degradation. *PLoS Pathog.* 6, e1000809.
- Chunhua, M., Yu, L., Yaping, G., Jie, D., Qiang, L., Xiaorong, T., Guang, Y., 2012. The expression of *LytM* is down-regulated by RNAIII in *Staphylococcus aureus*. *J. Basic Microbiol.* 52, 636–641. <http://dx.doi.org/10.1002/jobm.201100426>.
- Dambach, M.D., Winkler, W.C., 2009. Expanding roles for metabolite-sensing regulatory RNAs. *Curr. Opin. Microbiol.* 12, 161–169.
- Dunman, P.M., Murphy, E., Haney, S., Palacios, D., Tucker-Kellogg, G., Wu, S., Brown, E.L., Zagursky, R.J., Shlaes, D., Projan, S.J., 2001. Transcription profiling-based identification of *Staphylococcus aureus* genes regulated by the *agr* and/or *sarA* loci. *J. Bacteriol.* 183, 7341–7353.
- Felden, B., Vandenesch, F., Bouloc, P., Romby, P., 2011. The *Staphylococcus aureus* RNome and its commitment to virulence. *PLoS Pathog.* 7, e1002006.
- Fowler Jr., V.G., Sakoulas, G., McIntyre, L.M., Meka, V.G., Arbeit, R.D., Cabell, C.H., Strzyjewski, M.E., Eliopoulos, G.M., Reller, L.B., Corey, G.R., Jones, T., Lucindo, N., Yeaman, M.R., Bayer, A.S., 2004. Persistent bacteremia due to methicillin-resistant *Staphylococcus aureus* infection is associated with *agr* dysfunction and low-level in vitro resistance to thrombin-induced platelet microbicidal protein. *J. Infect. Dis.* 190, 1140–1149.
- Geisinger, E., Adhikari, R.P., Jin, R., Ross, H.F., Novick, R.P., 2006. Inhibition of rot translation by RNAIII, a key feature of *agr* function. *Mol. Microbiol.* 61, 1038–1048.
- Geissmann, T., Chevalier, C., Cros, M.-J., Boisset, S., Fechter, P., Noirot, C., Schrenzel, J., François, P., Vandenesch, F., Gaspin, C., Romby, P., 2009. A search for small noncoding RNAs in *Staphylococcus aureus* reveals a conserved sequence motif for regulation. *Nucleic Acids Res.* 37, 7239–7257.
- Horstmann, N., Orans, J., Valentin-Hansen, P., Shelburne III, S.A., Brennan, R.G., 2012. Structural mechanism of *Staphylococcus aureus* Hfq binding to an RNA A-tract. *Nucleic Acids Res.* 40, 11023–11035.
- Howden, B.P., Beaume, M., Harrison, P.F., Hernandez, D., Schrenzel, J., Seemann, T., François, P., Stinear, T.P., 2013. Analysis of the small RNA transcriptional response in multidrug-resistant *Staphylococcus aureus* after antimicrobial exposure. *Antimicrob. Agents Chemother.* 57, 3864–3874.
- Huntzinger, E., Boisset, S., Saveanu, C., Benito, Y., Geissmann, T., Namane, A., Lina, G., Etienne, J., Ehresmann, B., Ehresmann, C., Jacquier, A., Vandenesch, F., Romby, P., 2005. *Staphylococcus aureus* RNAIII and the endoribonuclease III coordinately regulate *spa* gene expression. *EMBO J.* 24, 824–835.
- Jelsbak, L., Hemmingsen, L., Donat, S., Ohlsen, K., Boye, K., Westh, H., Ingmer, H., Frees, D., 2010. Growth phase-dependent regulation of the global virulence regulator *Rot* in clinical isolates of *Staphylococcus aureus*. *Int. J. Med. Microbiol.* 300, 229–236.
- Jia, X., Zhang, J., Sun, W., He, W., Jiang, H., Chen, D., Murchie, A.I.H., 2013. Riboswitch regulation of aminoglycoside resistance acetyl and adenyl transferases. *Cell* 153, 1419–1420.
- Kaito, C., Kurokawa, K., Matsumoto, Y., Terao, Y., Kawabata, S., Hamada, S., Sekimizu, K., 2005. Silkworm pathogenic bacteria infection model for identification of novel virulence genes. *Mol. Microbiol.* 56, 934–944.
- Kaito, C., Saito, Y., Nagano, G., Ikuo, M., Omae, Y., Hanada, Y., Han, X., Kuwahara-Arai, K., Hishinuma, T., Baba, T., Ito, T., Hiramatsu, K., Sekimizu, K., 2011. Transcription and translation products of the cytolysin gene *psm-mec* on the mobile genetic element SCCmec regulate *Staphylococcus aureus* virulence. *PLoS Pathog.* 7, e1001267.
- Kaito, C., Saito, Y., Ikuo, M., Omae, Y., Mao, H., Nagano, G., Fujiyuki, T., Numata, S., Han, X., Obata, K., Hasegawa, S., Yamaguchi, H., Inokuchi, K., Ito, T., Hiramatsu, K., Sekimizu, K., 2013. Mobile genetic element SCCmec-encoded *psm-mec* RNA suppresses translation of *agrA* and attenuates MRSA virulence. *PLoS Pathog.* 9, e1003269.
- Lasa, I., Toledo-Arana, A., Dobin, A., Villanueva, M., de Mozos, I.R., Vergara-Irigaray, M., Segura, V., Fagegaltier, D., Penadés, J.R., Valle, J., Solano, C., Gingeras, T.R., 2011. Genome-wide antisense transcription drives mRNA processing in bacteria. *PNAS* 108, 20172–20177.
- Lasa, I., Toledo-Arana, A., Gingeras, T.R., 2012. An effort to make sense of antisense transcription in bacteria. *RNA Biol.* 9, 1039–1044.
- Link, T.M., Valentin-Hansen, P., Brennan, R.G., 2009. Structure of *Escherichia coli* Hfq bound to polyribadenylate RNA. *Proc. Natl. Acad. Sci. USA* 106, 19292–19297.
- Lioliou, E., Sharma, C.M., Altuvia, Y., Caldelari, I., Romilly, C., Helfer, A.-C., Margalit, H., Romby, P., 2013. In vivo mapping of RNA–RNA interactions in *Staphylococcus aureus* using the endoribonuclease III. *Methods* 63 (2), 135–143.
- Lioliou, E., Sharma, C.M., Caldelari, I., Helfer, A.-C., Fechter, P., Vandenesch, F., Vogel, J., Romby, P., 2012. Global regulatory functions of the *Staphylococcus aureus* endoribonuclease III in gene expression. *PLoS Genet.* 8, e1002782.
- Liu, Y., Wu, N., Dong, J., Gao, Y., Zhang, X., Mu, C., Shao, N., Yang, G., 2010. Hfq is a global regulator that controls the pathogenicity of *Staphylococcus aureus*. *PLoS One* 5, e13069.
- Liu, Y., Mu, C., Ying, X., Li, W., Wu, N., Dong, J., Gao, Y., Shao, N., Fan, M., Yang, G., 2011. RNAIII activates *map* expression by forming an RNA–RNA complex in *Staphylococcus aureus*. *FEBS Lett.* 585, 899–905.
- Marchais, A., Bohn, C., Bouloc, P., Gautheret, D., 2010. RsaOG, a new staphylococcal family of highly transcribed non-coding RNA. *RNA Biol.* 7, 116–119.
- Marchais, A., Naville, M., Bohn, C., Bouloc, P., Gautheret, D., 2009. Single-pass classification of all noncoding sequences in a bacterial genome using phylogenetic profiles. *Genome Res.* 19, 1084–1092.
- Marincola, G., Schäfer, T., Behler, J., Bernhardt, J., Ohlsen, K., Goerke, C., Wolz, C., 2012. RNaseY of *Staphylococcus aureus* and its role in the activation of virulence genes. *Mol. Microbiol.* 85, 817–832.
- Mikulecky, P.J., Kaw, M.K., Brescia, C.C., Takach, J.C., Sledjeski, D.D., Feig, A.L., 2004. *Escherichia coli* Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs. *Nat. Struct. Mol. Biol.* 11, 1206–1214.
- Morrison, J.M., Anderson, K.L., Beenken, K.E., Smeltzer, M.S., Dunman, P.M., 2012a. The staphylococcal accessory regulator, *SarA*, is an RNA-binding protein that modulates the mRNA turnover properties of late-exponential and stationary phase *Staphylococcus aureus* cells. *Front. Cell. Infect. Microbiol.* 2, 26.
- Morrison, J.M., Miller, E.W., Benson, M.A., Alonzo 3rd, F., Yoong, P., Torres, V.J., Hinrichs, S.H., Dunman, P.M., 2012b. Characterization of SSR42, a novel virulence factor regulatory RNA that contributes to the pathogenesis of a *Staphylococcus aureus* USA300 representative. *J. Bacteriol.* 194, 2924–2938.
- Mulhbachher, J., Brouillette, E., Allard, M., Fortier, L.-C., Malouin, F., Lafontaine, D.A., 2010a. Novel riboswitch ligand analogs as selective inhibitors of guanine-related metabolic pathways. *PLoS Pathog.* 6, e1000865.
- Mulhbachher, J., St-Pierre, P., Lafontaine, D.A., 2010b. Therapeutic applications of ribozymes and riboswitches. *Curr. Opin. Pharmacol.* 10, 551–556.
- Nagata, M., Kaito, C., Sekimizu, K., 2008. Phosphodiesterase activity of CvfA is required for virulence in *Staphylococcus aureus*. *J. Biol. Chem.* 283, 2176–2184.
- Nielsen, J.S., Christiansen, M.H.G., Bonde, M., Gottschalk, S., Frees, D., Thomsen, L.E., Kallipolitis, B.H., 2011. Searching for small  $\sigma$ B-regulated genes in *Staphylococcus aureus*. *Arch. Microbiol.* 193, 23–34.
- Novick, R.P., 2003. Autoinduction and signal transduction in the regulation of staphylococcal virulence. *Mol. Microbiol.* 48, 1429–1449.
- Novick, R.P., Geisinger, E., 2008. Quorum sensing in *Staphylococci*. *Annu. Rev. Genet.* 42, 541–564.
- Novick, R.P., Ross, H.F., Projan, S.J., Kornblum, J., Kreiswirth, B., Moghazeh, S., 1993. Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J.* 12, 3967–3975.
- Oun, S., Redder, P., Didier, J.-P., François, P., Corvaglia, A.-R., Buttazzoni, E., Giraud, C., Girard, M., Schrenzel, J., Linder, P., 2013. The CshA DEAD-box RNA helicase is important for quorum sensing control in *Staphylococcus aureus*. *RNA Biol.* 10, 157–165.
- Panja, S., Schu, D.J., Woodson, S.A., 2013. Conserved arginines on the rim of Hfq catalyze base pair formation and exchange. *Nucleic Acids Res.* 41 (15), 7536–7546.
- Papenfirt, K., Vogel, J., 2010. Regulatory RNA in bacterial pathogens. *Cell Host Microbe* 8, 116–127.
- Pichon, C., Felden, B., 2005. Small RNA genes expressed from *Staphylococcus aureus* genomic and pathogenicity islands with specific expression among pathogenic strains. *PNAS* 102, 14249–14254.
- Priest, N.K., Rudkin, J.K., Feil, E.J., van den Elsen, J.M.H., Cheung, A., Peacock, S.J., Laabei, M., Lucks, D.A., Recker, M., Massey, R.C., 2012. From genotype to phenotype: can systems biology be used to predict *Staphylococcus aureus* virulence? *Nat. Rev. Microbiol.* 10, 791–797.
- Proctor, R.A., von Eiff, C., Kahl, B.C., Becker, K., McNamara, P., Herrmann, M., Peters, G., 2006. Small colony variants: a pathogenic form of bacteria that facilitates persistent and recurrent infections. *Nat. Rev. Microbiol.* 4, 295–305.
- Queck, S.Y., Jameson-Lee, M., Villaruz, A.E., Bach, T.-H.L., Khan, B.A., Sturdevant, D.E., Ricklefs, S.M., Li, M., Otto, M., 2008. RNAIII-independent target gene control by the *agr* quorum-sensing system: insight into the evolution of virulence regulation in *Staphylococcus aureus*. *Mol. Cell* 32, 150–158.
- Roberts, C., Anderson, K.L., Murphy, E., Projan, S.J., Mounts, W., Hurlburt, B., Smeltzer, M., Overbeek, R., Disz, T., Dunman, P.M., 2006. Characterizing the effect of the *Staphylococcus aureus* virulence factor regulator, *SarA*, on log-phase mRNA half-lives. *J. Bacteriol.* 188, 2593–2603.
- Rochat, T., Bouloc, P., Yang, Q., Bossi, L., Figueroa-Bossi, N., 2012. Lack of interchangeability of Hfq-like proteins. *Biochimie* 94, 1554–1559.
- Romby, P., Charpentier, E., 2010. An overview of RNAs with regulatory functions in gram-positive bacteria. *Cell. Mol. Life Sci.* 67, 217–237.
- Romilly, C., Caldelari, I., Parmentier, D., Lioliou, E., Romby, P., Fechter, P., 2012a. Current knowledge on regulatory RNAs and their machineries in *Staphylococcus aureus*. *RNA Biol.* 9, 402–413.
- Romilly, C., Chevalier, C., Marzi, S., Masquida, B., Geissmann, T., Vandenesch, F., Westhof, E., Romby, P., 2012b. Loop–loop interactions involved in antisense regulation are processed by the endoribonuclease III in *Staphylococcus aureus*. *RNA Biol.* 9, 1461–1472.
- Roux, C.M., DeMuth, J.P., Dunman, P.M., 2011. Characterization of components of the *Staphylococcus aureus* mRNA degradosome holoenzyme-like complex. *J. Bacteriol.* 193, 5520–5526.
- Sakoulas, G., Eliopoulos, G.M., Moellering Jr., R.C., Wennersten, C., Venkataraman, L., Novick, R.P., Gold, H.S., 2002. Accessory gene regulator (*agr*) locus in geographically diverse *Staphylococcus aureus* isolates with reduced susceptibility to vancomycin. *Antimicrob. Agents Chemother.* 46, 1492–1502.

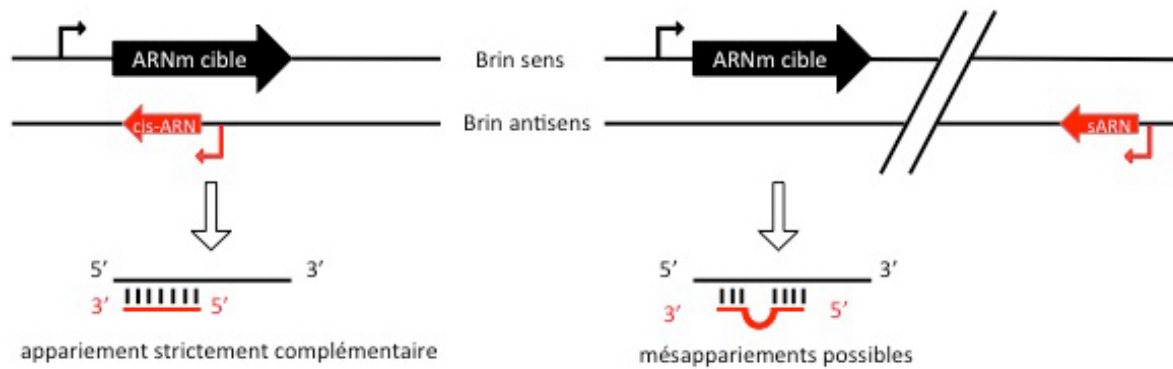
# OBJECTIFS

## 1. Problématique

Les différentes parties de mon introduction ont eu pour but de présenter le staphylocoque doré sous un angle assez large, afin de pouvoir définir facilement ce qui rend ce modèle bactérien très intéressant pour l'étude des mécanismes de régulation par les ARNnc. Comme il l'a été précisé auparavant, onze ARNnc avaient été identifiés en 2009 dans mon équipe et lorsque je suis arrivé au laboratoire, la fonction de deux de ces ARN était en cours d'identification. En premier lieu, la fonction de l'ARN RsaA a été l'objet d'étude du projet de thèse de Cédric Romilly avec qui j'ai collaboré à mon arrivée. Un second ARN, RsaE, a également été étudié dans l'équipe en particulier par Efthimia Lioliou. Dans les domaines de la microbiologie et de la biologie moléculaire, les technologies évoluent très vite. Le développement du séquençage haut débit (ARN-seq) et de la spectrométrie de masse permettent d'accéder à des milliers d'ARN et de protéines, respectivement, exprimés dans les cellules bactériennes. Comme la fonction de la majorité des ARNnc de *S. aureus* découverts dans mon équipe n'avait pas encore été élucidée, et dans une optique de compréhension globale des réseaux de régulation de la bactérie, j'ai combiné ces nouveaux outils technologiques avec des approches déjà existantes pour élucider la fonction de ces ARNnc *in vivo* et à l'échelle du génome.

### a. Cibles primaires et partenaires protéiques : à la recherche d'interactions avec les ARNnc

L'élucidation de la fonction des ARNnc doit passer par la recherche de leurs cibles, qu'elles soient primaires (c'est-à-dire résultant de l'interaction directe de l'ARNnc avec les ARNm cibles ou une protéine) ou secondaires – dans ce cas, l'ARNm n'est pas directement apparié avec l'ARNnc et plusieurs cas de figure peuvent exister. L'ARNm peut, par exemple, être pris dans un complexe avec l'ARNnc qui implique un troisième partenaire (ARN ou protéine), sans être directement en contact avec le premier. La régulation de l'ARNm peut également être contrôlée par un facteur de transcription dont l'expression sera, elle, contrôlée par l'ARNnc (c'est ce qui se passe lors de la régulation transcriptionnelle des gènes *spa* et *hla* assurée par la protéine Rot, dont la traduction est aussi contrôlée par l'ARNIII, § partie 3 introduction). Les ARN régulateurs sont souvent appelé sARN pour « Small non coding RNAs ».



**Figure 10 : Rappel sur les ARN régulateurs chez les bactéries** – On peut distinguer deux grands types d'ARN régulateurs. Les ARN situés en cis, ou cis-ARN, sont codés sur le locus directement opposé à celui de l'ARNm ciblé. Ainsi, l'appariement entre les deux ARN est strictement complémentaire et dans ce cas, l'ARNnc n'a qu'une seule cible. Les ARN qui agissent en trans ou sARN (pour *small non coding* ARN) sont codés sur un locus différent de celui du gène cible. Les appariements entre les deux ARN ne sont pas strictement complémentaires. Au contraire, les mésappariements existent. En découle deux conséquences : un ARNm peut être la cible de plusieurs ARNnc, et un ARNnc peut avoir plusieurs ARNm cibles. RsaA et RsaC (présentés ci-après) sont deux ARNnc de *S. aureus* appartenant à la seconde catégorie.

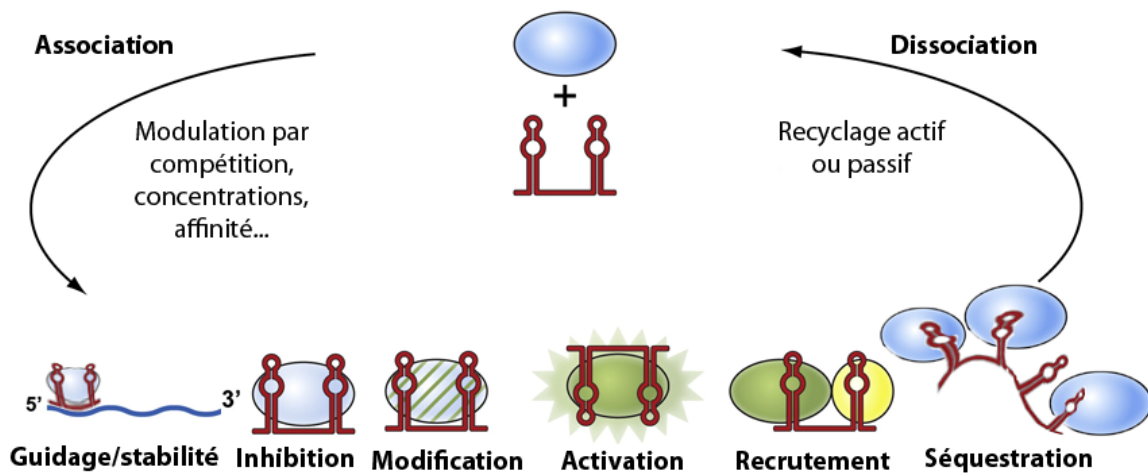
Nous avons fait le postulat que la plupart des ARNnc appelés RsaA-K identifiés chez *S. aureus* dans l'équipe sont des ARN qui agissent en trans (voir la **figure 10** ci-dessus). Par définition, ces ARN régulateurs ne font pas d'appariements strictement complémentaires avec leurs cibles, ce qui implique que celui-ci peut s'hybrider avec plusieurs ARNm et qu'un ARNm peut être régulé par plusieurs sARN. La recherche de cibles potentielles par prédictions *in silico* peut entraîner de nombreux faux positifs ou faux négatifs. En effet, les logiciels utilisés pour réaliser ces prédictions se basent en priorité sur le  $\Delta G$  (énergie libre) des appariements, et plus le  $\Delta G$  est bas, plus l'appariement prédit est stable. De nombreux paramètres sont disponibles en fonction des logiciels, pour prendre en compte les structures secondaires ou l'accessibilité des séquences – c'est notamment le cas du logiciel Inta RNA (Busch *et al.*, 2008). Mais tout ne peut pas être pris en compte par voie informatique : les interactions intermoléculaires, les structures 3D, les interactions à longue-distance, l'intervention de protéines qui débloquent la structure des sARN, la localisation intracellulaire des sARN ou encore l'encombrement moléculaire présent dans la cellule sont autant de facteurs qui peuvent faire s'apparier deux molécules non prédites par l'approche *in silico* ou a contrario prédire un appariement entre deux molécules qui ne se fera jamais *in vivo*. Les approches classiques de constructions de mutants de délétion des gènes d'ARNnc pour dégager des phénotypes – une autre façon d'aller vers la fonction - ou pour appliquer des études comparatives globales des transcriptomes et protéomes se sont avérées puissantes mais souvent non suffisantes pour discriminer les effets directs et indirects.

Le développement d'approches qui permettent de co-purifier les sARN avec leurs cibles est certainement l'un des verrous méthodologiques que l'on doit franchir. Rechercher les cibles permet également de replacer les ARNnc et ces dernières dans les réseaux de régulation de *S. aureus* (voir Tomasini *et al.*, 2014), ce qui permet en définitive d'avoir une meilleure compréhension, et globale, du fonctionnement de la bactérie et des liens qui entremêlent le métabolisme à la virulence.

Une autre question importante se pose chez *S. aureus* au sujet de ces ARNnc : il s'agit des protéines avec lesquelles ceux-ci pourraient potentiellement interagir. Chez les bactéries Gram négatif en effet, de nombreuses protéines se lient aux ARNnc pour les assister dans leur fonction, soit en les guidant vers leurs ARNm cible, ou également en stabilisant les duplexes d'ARN formés (voir **figure 11**). C'est le cas de la protéine Hfq qui est la plus connue des protéines chaperonnes de l'ARN chez les entérobactéries (Vogel & Luisi, 2011). Chez *S. aureus*, cette question est soumise à controverse et l'action de la protéine serait souche-dépendante (Bohn *et al.*, 2007 ; Liu *et al.*, 2010). Toutes les souches de *S. aureus* n'expriment pas le gène *hfq*, et des expériences de complémentation avec le gène de *S. aureus* chez une souche de *Salmonella enterica typhimurium* délétée du gène *hfq* n'ont pas réussi à remettre un phénotype sauvage chez cette dernière (Rochat *et al.*, 2012). Une analyse du mécanisme d'action de Hfq a été comparée chez diverses bactéries dont *S. aureus* (Zheng *et al.*, 2016). Cette analyse a révélé que la protéine Hfq de *S. aureus* n'a pas la capacité de former les appariements ARN-ARN à cause d'un motif rim dégénéré suggérant que cette protéine a évolué vers d'autres fonctions qui restent à découvrir.

Une autre classe de protéines associées à la régulation par les ARNnc sont les ribonucléases. La dégradation des ARN est souvent assurée par la combinaison de plusieurs RNases dont certaines forment un complexe protéique appelé dégradosome. De nombreux ARNnc influencent le devenir des ARNm en s'y liant, ce qui favorise dans un cas le recrutement d'une RNase ou à l'inverse protège l'ARN contre la dégradation. J'ai à ce titre collaboré à l'écriture d'une revue sur les RNases impliquées (Durand *et al.*, 2015) qui est incluse en annexe de cette thèse.

Si le rôle de Hfq chez *S. aureus* n'est pas complètement compris, il n'est pas exclu qu'une ou plusieurs protéines se lient aux ARNnc de la bactérie et les assiste dans cette fonction. A l'inverse il se pourrait également que des protéines soient la cible d'ARN régulateurs comme c'est le cas de l'ARN RsmZ chez *P. aeruginosa* qui séquestre par ses tiges-boucles la protéine RsmA (Bordi *et al.*, 2010). L'étendue des possibilités chez *S. aureus* est grande alors que ce champ est encore vierge de découvertes...



**Figure 11 : Interactions entre ARNnc et protéines et leurs effets** – L'association d'un ARNnc à une protéine peut avoir plusieurs effets sur son activité (inhibition, activation, modification) ou également amener au recrutement d'autres protéines. Dans certains cas, un ARNnc peut séquestrer des protéines par la présence de motifs répétés. Enfin, notamment dans le cas d'Hfq, certaines protéines servent de guide pour un ARNnc et l'amener à son messager cible, ou peuvent stabiliser les duplexes d'ARN. Les conditions d'associations et de dissociations des complexes ARN/protéines sont modulées par différents facteurs. Adapté depuis Storz *et al.*, 2011.

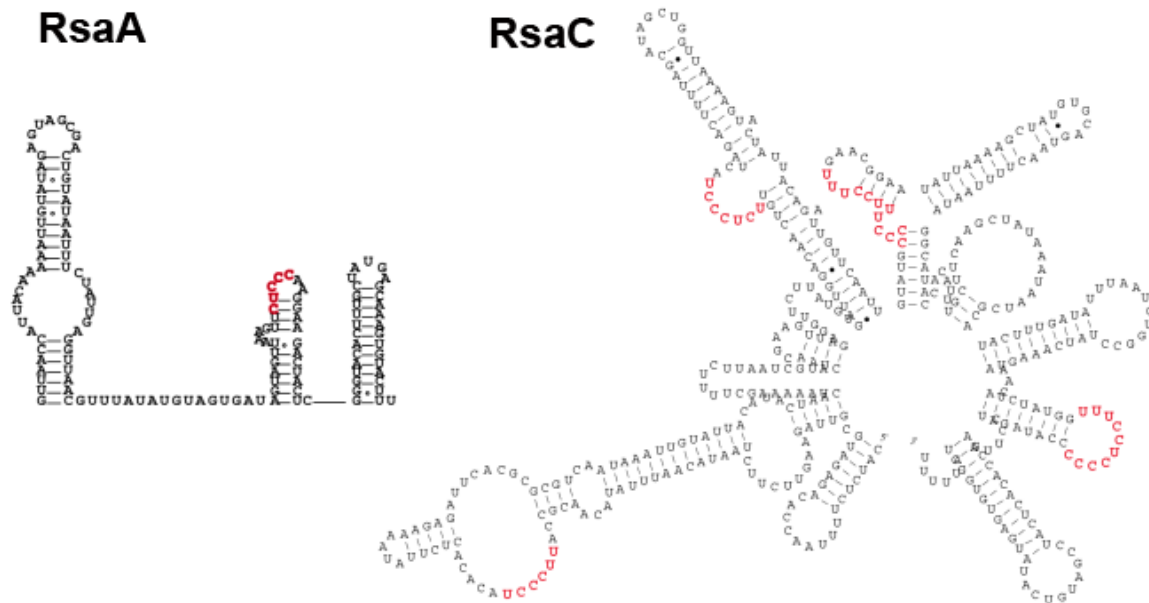
### b. Le choix de deux ARNnc de *S. aureus* comme modèles d'étude

Pour répondre à ces deux grandes problématiques (celle des ARNm cibles, et celles des potentielles protéines interactantes), j'ai utilisé comme modèles deux ARNnc RsaA et RsaC de *S. aureus* qui présentent chacun des caractéristiques particulières (**figure 12**). Tous deux présentent un ou plusieurs motifs riches en cytosines, qui a été montré comme étant une signature possible pour cibler les régions SD des ARNm.

Comme annoncé ci-dessus, j'ai participé à mon arrivée au laboratoire à la validation expérimentale d'une des cibles majeures de RsaA, un ARNnc long de 139 nucléotides et qui répond au facteur  $\sigma^B$  (Geissmann *et al.*, 2009). Ce travail réalisé en collaboration avec C. Romilly, fait l'objet de ma première partie des résultats où j'ai expliqué ma contribution personnelle en introduction du manuscrit publié dans PloS Pathogen.



Ayant caractérisé l'ARNm *mgrA* – qui code le facteur de transcription MgrA comme cible de RsaA, j'ai utilisé ce système comme preuve de concept pour développer de nouvelles approches *in vivo* afin d'étendre le régulon de RsaA mais aussi pour pouvoir les appliquer à tout autre type d'ARN tel que RsaC.



**Figure 12 :** Structure secondaire de RsaA et structure prédite de RsaC selon Geissmann et al, 2009 – Les motifs riches en C sont indiqués en rouge.

RsaC est un ARN bien plus long que le reste des ARNnc de *S. aureus*, ayant une taille (544 nucléotides) qui se rapproche de l'ARNIII (513 nucléotides) et qui possède une structure secondaire prédite complexe (voir **figure 12**). La répétition de plusieurs motifs riches en C, comme observé chez l'ARNIII laisse fortement supposer que cet ARN est aussi capable de réguler la traduction d'ARNm. Cependant sa taille et sa structure secondaire, tout comme la répétition de ces motifs à l'extrémité de tige-boucles en font aussi un très bon candidat d'un ARN qui pourrait séquestrer (ou se lier) à des protéines. Comme la fonction de cet ARN n'était pas encore connue à mon arrivée au laboratoire, l'objectif fut également d'aller vers la recherche des cibles en utilisant une combinaison d'approches globales pour purifier les complexes régulateurs et plus spécifiques par l'utilisation d'une souche mutante pour évaluer les phénotypes.

En définitive, mon projet de thèse a visé à répondre à trois grandes questions.

**Comment aller rechercher de manière efficace les ARNm cibles primaires des ARNnc de *S. aureus* ?**

**Quelles sont les protéines qui interagissent avec les ARNnc de *S. aureus* ?**

**Quelle est la fonction de RsaC ?**

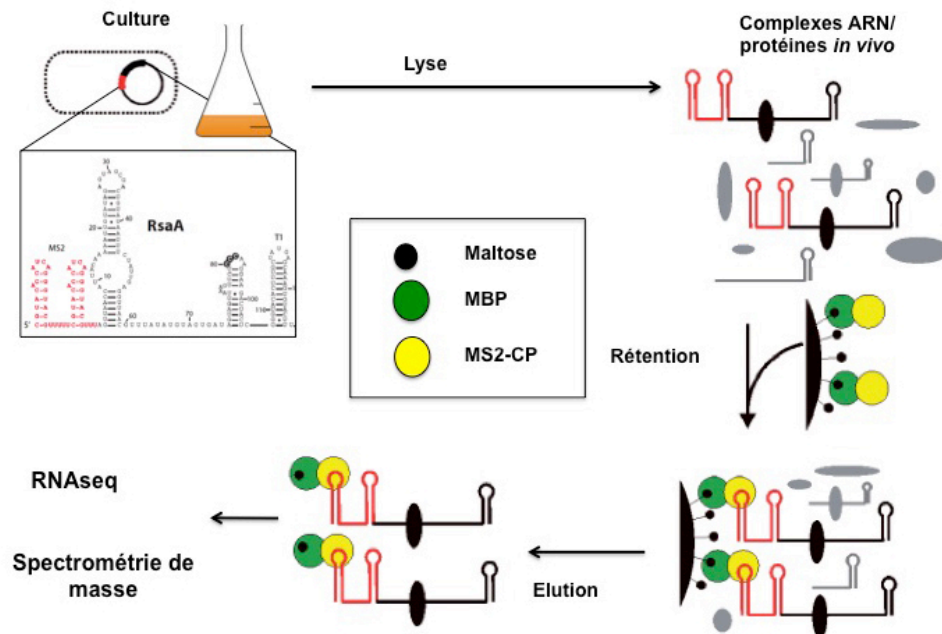
## 2. Démarche expérimentale

Pour répondre aux deux premières questions posées par mon sujet de thèse, j'ai développé au laboratoire deux approches que j'ai adaptées de techniques existantes, et qui ont été utilisées pour la première fois chez *S. aureus* dans un contexte *in vivo*.

La première approche est tirée des travaux de Said *et al.* (2009). Elle consiste en une chromatographie d'affinité qui utilise un ARN modifié à son extrémité 5' par l'ajout d'une étiquette ribonucléique, appelée MS2. L'ARN (ARN-MS2) est exprimé *in vivo* à partir d'un plasmide transformé dans une souche de *S. aureus* où le gène de l'ARNnc a été au préalable délété. Au cours de la culture bactérienne, l'ARN-MS2 va se lier à ses ARNm cibles et éventuellement à diverses protéines. Les cellules sont ensuite lysées et l'extrait brut qui est récupéré est passé sur une chromatographie d'affinité constituée d'une matrice d'amylose sur laquelle est lié de manière covalente du maltose auquel est accroché une protéine de fusion MBP-CP (*maltose binding protein* fusionnée à la protéine de capsid du phage MS2, qui reconnaît spécifiquement l'étiquette MS2 de l'ARN-MS2). Ainsi, les complexes ARN-protéine sont retenus sur la colonne et peuvent ensuite être élués par ajout de maltose en excès. Les ARN et protéines sont ensuite purifiés et identifiés par une approche de séquençage ARN-seq haut débit et de spectrométrie de masse, respectivement. Le principe est expliqué schématiquement dans la **figure 13**. Cette approche a été récemment appliquée avec succès chez *E. coli* (Lalaouna *et al.*, 2015). Elle a été nommée MAPS pour « MS2-Affinity Purification coupled with RNA Sequencing ».

Alors que l'approche MAPS est spécifique d'un ARN donné, nous avons également besoin de méthodes qui soient plus globales et permettent d'aller rechercher les interactants des ARNnc à grande échelle. J'en viens donc à la seconde approche que j'ai développée au laboratoire, tirée des travaux de Wassarman & Storz (2010). Elle est basée sur une méthode de séparation des complexes ribonucléoprotéiques par une centrifugation d'extraits bruts sur un gradient de glycérol. Dans notre cas, les souches d'intérêt de *S. aureus* sont cultivées en conditions standard de laboratoire et les extraits bruts sont préparés à partir des lysats de cellules. Les extraits sont déposés sur un

gradient de glycérol 10-40% et la centrifugation est opérée pendant 16 heures à 36000 rpm et 4°C. Le gradient est séparé en différentes fractions sur chacune desquelles les ARN et protéines sont séquencés séparément par ARN-seq et spectrométrie de masse, respectivement.



**Figure 13 : Représentation schématique de l'approche MAPS** – La structure MS2 est présentée en rouge, l'ARN RsaA est ici montré comme exemple. Adapté de Said *et al.*, 2009.

Enfin pour déterminer la fonction de l'ARN RsaC, la première étape était d'obtenir un mutant *rsaC* dans diverses souches de *S. aureus*, ainsi que les complémentations, pour ensuite mener des approches de protéomique ainsi que des tests phénotypiques, tout en utilisant également l'approche MAPS que j'ai appliquée à cet ARN.

# RESULTATS

# Partie 1 – Validation de MgrA comme cible primaire de RsaA

Comme je l'ai précisé précédemment, en collaboration avec Cédric Romilly j'ai contribué à la caractérisation d'une cible majeure de l'ARNnc RsaA, au mécanisme de régulation, et à l'étude de son impact sur la physiologie de *S. aureus*. J'ai en particulier construit un gène rapporteur pour valider *in vivo* une des cibles primaires de RsaA, l'ARNm *mgrA*, et montrer l'importance de deux sites d'interactions pour la régulation de l'expression de cette cible.

## 1. Validation *in vivo* de la régulation de MgrA par RsaA par diverses approches différentielles de spectrométrie de masse

Pour caractériser la fonction de RsaA, Cédric Romilly a tout d'abord réalisé une approche protéomique différentielle réalisée à partir d'extraits protéiques préparés à partir de la souche sauvage HG001, de la souche délété du gène *rsaA* (réalisée par remplacement allélique par nos collaborateurs, F. Vandenesch, Lyon) et de cette même souche mutante complémentée avec un plasmide qui exprime le gène *rsaA* sous le contrôle du promoteur P3 de l'ARNIII. Celui-ci permet une accumulation de RsaA en phase stationnaire de croissance, avec un taux proche de celui de RsaA exprimé à partir du chromosome. L'approche protéomique comparative est basée sur le marquage des protéines par des fluorophores spécifiques, chacun correspondant à une souche. Un mélange des deux extraits protéiques a été marqué avec un troisième fluorochrome pour servir de contrôle interne – ainsi on s'assure que la quantité totale de protéines dans chaque extrait est la même. Les extraits marqués ont ensuite été mélangés puis séparés par électrophorèse à deux dimensions (DiGE). La première séparation se fait en fonction du point isoélectrique (pI) et la seconde en fonction du poids moléculaire. Après migration, les différentes protéines marquées sont révélées sur le gel et quantifiées (un spot correspond souvent à une protéine). La superposition des différentes couleurs permet de visualiser rapidement les protéines dont l'expression varie en fonction de la présence ou de l'absence de l'ARN. Les spots présentant des expressions différentielles sont découpés pour caractériser la nature de la protéine par spectrométrie de masse. Parmi les protéines identifiées au cours de cette approche, au nombre d'une centaine, il a été remarqué qu'un certain nombre d'entre elles font partie du régulon d'un même facteur de transcription, MgrA. La protéine MgrA n'a en revanche pas été détectée par

l'approche DiGE, qui présente des limites pour détecter les protéines de faible poids moléculaire, faiblement exprimées et à pI basiques comme c'est le cas des facteurs de transcription (dont MgrA est un représentant ; sa taille est de 17kDa).

Des prédictions d'appariements de paires de bases entre RsaA et les ARNm ont été effectuées par divers outils informatiques (IntaRNA, targetRNA) et dans tous les cas l'ARNm *mgrA* est ressorti en tête de liste. En accord avec cette étude, des expériences de retard sur gel (voir Romilly et al., 2014) ont montré que RsaA est capable de former un complexe stable avec l'ARNm *mgrA*. Comme la protéine MgrA n'avait pu être détectée en DiGE, il fallait donc trouver une méthode alternative pour valider la régulation de l'expression de cette protéine *in vivo* par RsaA.

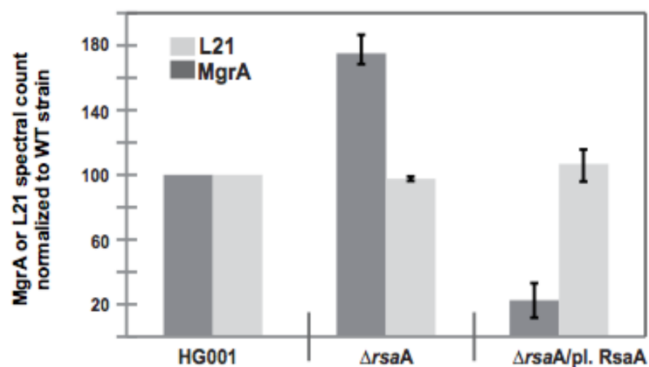
Deux approches ont été menées en parallèle. Dans un premier temps, j'ai utilisé une nouvelle approche de protéomique quantitative sur des extraits bruts non marqués. Ces expériences ont été réalisées sur la souche HG001 dérivée de la souche de laboratoire RN8325-4 mais dans laquelle la mutation dans le gène *rsbU* a été restaurée pour permettre l'expression du facteur  $\sigma^B$ . En effet, ce facteur est requis pour induire l'expression de RsaA. J'ai modifié le plasmide pCN51 (Charpentier et al., 2004) qui porte le gène *rsaA* sous le contrôle de son promoteur endogène que j'ai transformé dans la souche mutante  $\Delta$ *rsaA* pour obtenir la souche complémentée. Les souches HG001 (sauvage) et  $\Delta$ *rsaA* ont également été transformées avec le même plasmide vide. En effet, la présence d'un plasmide et l'utilisation d'un antibiotique pour maintenir le plasmide dans la bactérie peuvent modifier la physiologie de la bactérie. En transformant la souche témoin avec le même plasmide vide, je m'assure donc que les seules différences que j'observe sont dues à l'expression de RsaA porté par le plasmide. Les souches ont été mises en culture en condition standard pendant 6 heures puisque c'est à ce moment de la phase de croissance que RsaA est le plus exprimé (Geissmann *et al.*, 2009). J'ai récupéré pour chaque culture les extraits protéiques totaux qui ont été ensuite séquencés par la spectrométrie de masse (plateforme protéomique, IBMC). Pour chaque condition, l'expérience a été réalisée en triplicats biologiques (trois colonies différentes mises en culture). La plateforme de spectrométrie de masse venait d'acquérir un nouveau séquenceur équipé d'une chromatographie liquide, pour séquencer directement les protéines à partir d'un échantillon sans passer par une étape de gel en 2D. J'ai mené avec Philippe Hamman, responsable de ladite plateforme, cette approche de séquençage direct sur échantillons, de façon pionnière dans l'équipe. Comme dans toute expérience de spectrométrie de masse, les protéines contenues dans les échantillons sont en premier lieu digérées par la trypsine. Les peptides obtenus sont ensuite

injectés dans le séquenceur et séparés par chromatographie liquide. Les peptides sont ensuite ionisés et analysés par MALDI-TOF (LC MSMS). Pour chaque peptide, un spectre lui est attribué et le comptage d'un ensemble de spectres attribués à l'ensemble des peptides générés par la trypsine permet donc d'établir l'abondance de la protéine à laquelle ce *spectral count* correspond.

**A**

WT1	WT2	WT3	$\Delta RsaA1$	$\Delta RsaA2$	$\Delta RsaA3$	Compl1	Compl2	Compl3
17	23	21	31	39	29	6	4	5

**B**



**Figure 14 : RsaA inhibe l'expression de MgrA in vivo** – A. Nombre de spectres comptés pour MgrA dans chacun des extraits protéiques totaux. B. Représentation graphique de l'abondance relative de MgrA chez les souches HG001 mutantes et complémentées par rapport à la souche sauvage.

Nous avons tout d'abord analysé les spectres des peptides obtenus pour MgrA dans les trois souches, en utilisant comme contrôle la protéine ribosomique L21 dont l'expression doit être inchangée dans les trois souches (WT, mutée et complémentée). Les résultats de cette expérience, présentés en **figure 14**. Le taux de protéine MgrA a été quantifiée par la méthode du *spectral count* et normalisée sur la condition sauvage (**figure 14**). Les résultats montrent clairement une accumulation de la protéine en absence de RsaA alors que dans la souche complémentée l'expression de cette dernière est fortement inhibée. Cette inhibition est plus forte dans la souche complémentée que dans la souche WT car le taux de RsaA est plus abondant dans la souche complémentée puisque le plasmide est présent en plusieurs copies (nombre néanmoins modéré).

Ces résultats ont été corrélés à l'analyse directe du taux de MgrA par une approche de Western-blot. Dans cette deuxième approche protéomique, une étiquette FLAG a été introduite dans le chromosome pour obtenir une protéine marquée à son extrémité C-terminale MgrA-FLAG dans la souche sauvage, mutante, et complémentée. La protéine a été détectée en utilisant des anticorps spécifiques de l'étiquette FLAG. Ces résultats sont présentés dans la Figure 1 de la publication de Romilly *et al.* (2014). L'ensemble de cette étude protéomique confirme une inhibition de la synthèse de MgrA dépendante de RsaA.

Cette régulation pourrait être potentiellement indirecte, mais les expériences de retard sur gel et de cartographie en solution réalisées par Cédric Romilly vont clairement dans le sens d'une régulation médiée par des interactions ARN-ARN. Bien entendu, si nous avons choisi de nous concentrer sur MgrA dans cette étude, le séquençage des culots de protéines des trois souches a apporté un nombre de données importantes sur plus de 1000 protéines. Ces données seront notamment discutées dans la seconde partie des Résultats et dans la Discussion de cette thèse.

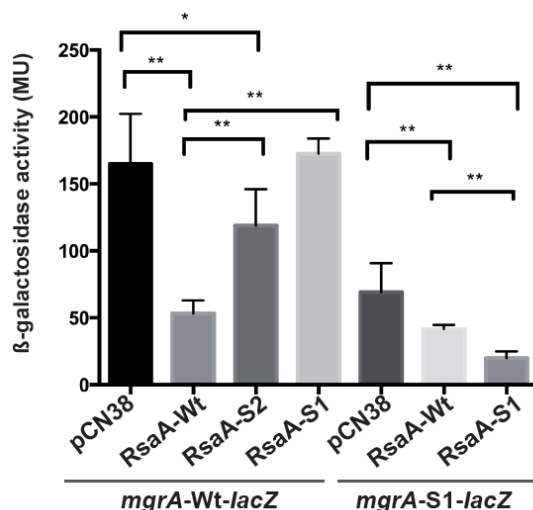
## 2. Validation *in vivo* de l'importance des deux sites S1 et S2 de RsaA pour la régulation de MgrA

Nous nous sommes ensuite attachés à déterminer la mécanistique de la régulation. Cédric Romilly a mis en place plusieurs approches *in vitro*. Il a tout d'abord démontré que RsaA et MgrA interagissent directement ensemble par des expériences de retard sur gel, et déterminer les régions impliquées dans l'interaction par cartographie en solution. Deux sites distants ont été caractérisés. Le site S1 comprend le motif riche en C de RsaA qui interagit avec la séquence SD de l'ARNm *mgrA*. Le site S2, est quant à lui localisé dans la boucle apicale d'une des tiges-boucles de RsaA, qui forme une interaction de type boucle-boucle avec une séquence complémentaire localisée dans la partie codante de l'ARNm *mgrA*. Cédric a notamment montré par retard sur gel que les deux sites sont importants pour l'interaction *in vitro* entre les deux ARN. Des mutations ont été introduites dans les sites S1 et S2 de RsaA. Chaque mutation conduit à la perte de la formation du complexe. Des mutations compensatoires réalisées sur l'ARNm *mgrA*, restaure l'interaction.

Pour démontrer que les deux sites sont importants *in vivo* pour la régulation de l'expression de MgrA, j'ai analysé l'expression d'une fusion dans lequel le gène *lacZ* a été fusionné à l'extrémité 5' avec la région de régulation de *mgrA*. J'ai ainsi transformé trois souches de *S. aureus*  $\Delta$ *rsaA* exprimant la version sauvage de RsaA sur le plasmide pCN38 (Charpentier *et al.*, 2004) ainsi que les versions RsaA-S1 et RsaA-S2, qui sont les versions mutées dans les deux sites d'interaction en question. Une souche contrôle a été transformée avec le plasmide pCN38 ne portant pas d'insert. Chacune des souches a également été transformée avec un plasmide pLUG220 – construit par Cédric - portant la fusion *mgrA-lacZ*. J'ai utilisé deux constructions différentes, une incluant la région régulatrice sauvage de MgrA, et une seconde qui porte une mutation au niveau du site S1 (au niveau de la séquence SD). J'ai ensuite mis en culture chacune des souches portant les différentes combinaisons de plasmides. Les extraits bruts ont été préparés pour mesurer l'activité



de la  $\beta$ -galactosidase, cette activité étant sous la dépendance de la région régulatrice de *mgrA*. Les expériences ont été reproduites trois fois et une analyse statistique par test Anova a été réalisée. Comme le montre la **figure 15**, on observe un fort niveau d'activité de la  $\beta$ -galactosidase en absence de RsaA. Cette activité est fortement diminuée lorsque la version sauvage de RsaA est exprimée, ce qui traduit une inhibition de l'expression de MgrA. En revanche, les versions mutantes RsaA-S1 et RsaA-S2 ne sont plus capables d'inhiber la synthèse de la fusion. La perte d'activité est partielle pour RsaA-S2 alors qu'elle est totale avec RsaA-S1, ce qui montre que les deux sites sont importants pour la régulation. Cependant c'est le site S1 qui fait, pour rappel, intervenir le motif riche en C de RsaA et la séquence SD sur MgrA, qui est prédominant pour la régulation. Ces résultats montrent que la régulation *in vivo* est bien dépendante d'une interaction ARN-ARN.



**Figure 15 : Les sites S1 et S2 sont importants pour la régulation de MgrA par RsaA *in vivo*** – L'activité enzymatique d'une protéine de fusion MgrA-lacZ est mesurée en l'absence et en présence de RsaA Wt et de ses variants RsaA-S1 et RsaA-S2. Les résultats ont été comparés statistiquement deux-à-deux par un test Anova. \* p < 0,05 et \*\* p < 0,005.

Dans la seconde expérience (**figure 15**, à droite) utilisant la fusion mutée au niveau du site S1, on observe que le niveau basal d'activité de la  $\beta$ -galactosidase est beaucoup plus faible, ce qui s'explique par le fait que la mutation proche de la séquence SD de *mgrA* entraîne une diminution de la reconnaissance de l'ARNm par les ribosomes. L'expression de RsaA inhibe faiblement la synthèse de la fusion mutée. En revanche, la mutation compensatoire dans RsaA (RsaA-S1) permet d'amplifier l'inhibition. Ainsi cette expérience montre que les deux sites contribuent à la répression bien que le site S1 soit prédominant. Ces résultats se corréleront parfaitement aux expériences de gel retard.

Des expériences de toeprint (réalisées par Cédric) ont par ailleurs montré que RsaA inhibe la formation du complexe d'initiation (Figure 3, Romilly et al., 2014). Par ailleurs, la mesure du taux de l'ARNm *mgrA* par des expériences de RT-qPCR a montré que celui-ci s'accumule dans la souche mutée  $\Delta$ rsaA suggérant que cet ARNm est dégradé suite à l'inhibition de la traduction. Il se peut que cette dégradation soit initiée par la RNase III qui coupe l'hybride formé entre *mgrA* et RsaA.

Basés sur nos résultats, en collaboration avec l'équipe de F. Vandenesch, nous avons étudié l'impact de cette régulation sur les interactions entre le pathogène et l'hôte. Du fait que MgrA est un activateur majeur de la formation de la capsule et un inhibiteur des biofilms, nous avons analysé l'impact de RsaA sur ces deux phénotypes caractéristiques. Les résultats ont démontré pour la première fois que RsaA favorise la formation des biofilms et les infections de type chronique dans un modèle murin. En revanche celui-ci agit comme un atténuateur des infections aiguës en réprimant la synthèse de la capsule.

**Ainsi, les expériences que j'ai réalisées ont permis de valider *in vivo* une cible primaire de l'ARNnc RsaA, démontrer que le contrôle de *mgrA* par RsaA s'effectue au niveau post-transcriptionnel, et illustrer l'importance des deux sites d'interactions avec l'ARNm *mgrA*. Elles ont été de fait incluses dans une publication scientifique dont je suis auteur en 3<sup>ème</sup> position. Ladite publication est à retrouver ci-après.**

# A Non-Coding RNA Promotes Bacterial Persistence and Decreases Virulence by Regulating a Regulator in *Staphylococcus aureus*

Cédric Romilly<sup>1,2</sup>, Claire Lays<sup>2,3,4,5,6</sup>, Arnaud Tomasini<sup>1</sup>, Isabelle Caldelari<sup>1</sup>, Yvonne Benito<sup>2,3,4,5,6,7</sup>, Philippe Hamann<sup>8</sup>, Thomas Geissmann<sup>2,3,4,5,6</sup>, Sandrine Boisset<sup>2,3,4,5,6,7</sup>, Pascale Romby<sup>1\*</sup>, François Vandenesch<sup>2,3,4,5,6,7\*</sup>

**1** Architecture et Réactivité de l'ARN, Université de Strasbourg, CNRS, IBMC, Strasbourg, France, **2** CIRI, International Center for Infectiology Research, Université de Lyon, Lyon, France, **3** Inserm U1111, Lyon, France, **4** Ecole Normale Supérieure de Lyon, Lyon, France, **5** Université Lyon 1, Lyon, France, **6** CNRS, UMR5308, Lyon, France, **7** Hospices Civils de Lyon, Lyon, France, **8** Plateforme Protéomique IBMC, Strasbourg, France

## Abstract

*Staphylococcus aureus* produces a high number of RNAs for which the functions are poorly understood. Several non-coding RNAs carry a C-rich sequence suggesting that they regulate mRNAs at the post-transcriptional level. We demonstrate that the Sigma B-dependent RsaA RNA represses the synthesis of the global transcriptional regulator MgrA by forming an imperfect duplex with the Shine and Dalgarno sequence and a loop-loop interaction within the coding region of the target mRNA. These two recognition sites are required for translation repression. Consequently, RsaA causes enhanced production of biofilm and a decreased synthesis of capsule formation in several strain backgrounds. These phenotypes led to a decreased protection of *S. aureus* against opsonophagocytic killing by polymorphonuclear leukocytes compared to the mutant strains lacking RsaA. Mice animal models showed that RsaA attenuates the severity of acute systemic infections and enhances chronic catheter infection. RsaA takes part in a regulatory network that contributes to the complex interactions of *S. aureus* with the host immune system to moderate invasiveness and favour chronic infections. It is the first example of a conserved small RNA in *S. aureus* functioning as a virulence suppressor of acute infections. Because *S. aureus* is essentially a human commensal, we propose that RsaA has been positively selected through evolution to support commensalism and saprophytic interactions with the host.

Citation: Romilly C, Lays C, Tomasini A, Caldelari I, Benito Y, et al. (2014) A Non-Coding RNA Promotes Bacterial Persistence and Decreases Virulence by Regulating a Regulator in *Staphylococcus aureus*. *PLoS Pathog* 10(3): e1003979. doi:10.1371/journal.ppat.1003979

Editor: Michael S. Gilmore, Harvard Medical School, United States of America

Received September 20, 2013; Accepted January 21, 2014; Published March 20, 2014

Copyright: © 2014 Romilly et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the Centre National de la Recherche Scientifique (CNRS; PR CR AT IC), the Institut National de la Santé et de la Recherche Médicale (INSERM; FV CL TG SB YB), the Agence Nationale de la Recherche (ANR10-Pathogenomics-ARMSA; FV PR), the "Laboratoires d'excellence" (LABEX) NetRNA grant ANR-10-LABX-36 (PR) in the frame of "Programme d'Investissements d'Avenir." CR was supported by fellowships from the CNRS, DGA (Délégation Générale pour l'Armement) and FRM (Fondation de la Recherche Médicale). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

\* E-mail: p.romby@ibmc-cnrs.unistra.fr (PR); francois.vandenesch@univ-lyon1.fr (FV)

□ Current address: Department of Cell and Molecular Biology, microbiology program, Biomedical Center, Uppsala University, Uppsala, Sweden.

– These authors contributed equally to this work.

## Introduction

*Staphylococcus aureus* is an opportunistic pathogen that has evolved complex regulatory circuits allowing rapid adaptation of cell growth in response to its diverse hosts and ecological niches. Present in a large proportion of the population as a commensal of skin and nose, the bacteria is also responsible for a large range of hospital-acquired and community infections [1]. A successful infection by *S. aureus* largely depends on the coordinated and sequential expression of a multitude of virulence factors and accessory genes. Over the last decade, it has been established that *S. aureus* genes are regulated at many different levels by a variety of trans-acting regulators, which act in a coordinated manner [2,3]. Among them, RNAs are now recognized as important players in virulence and many physiological and adaptive responses [4,5]. The first regulatory RNA that was discovered in 1993 is RNAIII, the main intracellular effector of the

quorum sensing agr system [6]. This multi-functional regulatory RNA binds to several target mRNAs to regulate their translation and decay [7–11]. Later on, several teams have experimentally identified a large number of small RNAs (sRNA) that are issued from the core genome and from mobile and accessory elements (e.g., [12,13]). These sRNAs include cis-acting regulatory regions of mRNAs (the so-called riboswitches), cis-encoding antisense RNAs (asRNA), and non-coding RNAs (ncRNAs). In addition, sRNAs carrying small open reading frames (sORF) have been recently identified [14–17] and one of them was shown to express a small cytolytic peptide [17]. The functional and mechanistic studies of trans-acting sRNAs are still lagging behind their discovery, but some of them are known to sense population density and various environmental changes, modify cell surface properties, adjust the bacterial metabolism during cell growth, regulate virulence gene expression, and respond to antibiotic treatment [4,18].

## Author Summary

*Staphylococcus aureus* is a commensal and an opportunistic pathogen that causes a large range of community and hospital-acquired infections. The bacteria produce an array of virulence factors, the expression of which is regulated by a set of regulators including proteins and RNAs. In recent years, a large number of small non-coding RNAs encoded by the *S. aureus* genome have been identified but determination of their function is still lagging behind. This study shows that RsaA, a staphylococcal conserved non-coding RNA, operates at the post-transcriptional level by repressing the translation of the master regulatory protein MgrA. The repression is based on a direct interaction of RsaA with the ribosome binding site of mgrA mRNA. Through MgrA regulation, RsaA activates biofilm formation and inhibits capsule synthesis. Using appropriate animal models, we showed that RsaA acts as a suppressor of virulence because the deletion of its gene increases the invasiveness of *S. aureus* in the mice sepsis model. RsaA is thus part of complex regulatory network that modify the interactions of *S. aureus* with the eukaryotic immune system. These findings illustrate how small RNAs can have a major impact in bacterial biology.

We have previously identified a group of sRNAs that carry a conserved and unpaired UCCC sequence suggesting that they act through a common regulatory mechanism [12–14,19–23]. This sequence motif is located in several hairpin loops of *S. aureus* RNAIII and was demonstrated to be the seed sequence, which recognizes the ribosome binding sites of mRNA targets to repress translation [5]. Here, using a combination of *in vivo* and *in vitro* approaches, we have elucidated the function and the mechanism of action of one of these sRNAs called RsaA [20]. We show that RsaA, a Sigma B ( $\sigma^B$ )-dependent sRNA, represses the translation of mgrA mRNA, which codes for a master regulator of transcription [24–28]. Through MgrA regulation, RsaA activates biofilm formation and inhibits capsule synthesis. Experimental animal models showed that RsaA attenuates the severity of acute systemic infections and enhances chronic catheter infection. These data indicate that RsaA is part of a regulatory circuit that contributes to the complex interactions of *S. aureus* with the eukaryotic immune system.

## Results

### RsaA affects the synthesis of MgrA and of proteins regulated by MgrA

To assess the function of RsaA, we first constructed an HG001-derived strain, in which the *rsaA* gene was deleted and replaced by *aphA-3* (aminoglycoside 39-phosphotransferase, kanamycin resistance). We have verified that RsaA expression in HG001, which is  $\sigma^B$ -dependent, was maximal at the late-exponential phase of growth and was abolished in the mutant *DrsaA* strain. In BHI medium, growth was not affected by *rsaA* deletion. We then analyzed the effect of the *DrsaA* mutation on protein synthesis using comparative proteomics performed on cytosolic protein extracts prepared from wild-type (HG001) and mutant *DrsaA* strains grown in BHI medium to late-exponential growth phase. Protein samples were pre-labeled with two different fluorescent dyes (Cy3 and Cy5) and were separated by a two-dimensional difference gel electrophoresis (DiGE). Even though only a small proportion of cytosolic proteins (around 100 proteins) were

analyzed, we identified more than 10 proteins with significantly altered expression patterns (Figures 1A, S1). Among the proteins whose synthesis was significantly increased in the mutant strain, was SpoVG, a protein involved in capsule synthesis [29]. Other proteins involved in intermediary metabolism and in virulence were also affected by the *DrsaA* mutation. Several of these proteins belong to the regulon of the pleiotropic regulator MgrA [25,26] but the effect of RsaA was opposite to that of MgrA.

RsaA contains an unpaired and conserved C-rich sequence, suitable for interaction with ribosome binding sites (RBS) of mRNAs, as previously shown for several other regulatory RNAs in *S. aureus* [12–14,19–23]. Using IntaRNA [30] and targetRNA [31], we searched for potential intermolecular base pairings between the ribosome binding sites (RBS) of mRNAs (30 nucleotides (nts) downstream and 20 nts upstream of the AUG codon) and the conserved unpaired region of RsaA. Interestingly, the best mRNA candidate obtained by the two programs was mgrA mRNA (Table S1). In addition, two mRNAs encoding SpoVG and glycyl-tRNA synthetase, whose protein levels were decreased in an RsaA-dependent fashion, were predicted to form base pairings with RsaA (Figure 1B).

Because MgrA protein could not be detected by the differential proteomic analysis, two alternative methods were used to analyze whether the synthesis of MgrA was under the control of RsaA. Protein extracts, prepared from the WT strain (HG001), the *DrsaA*-HG001 mutant strain, and the same strain complemented with a plasmid expressing RsaA under its own promoter were analyzed by Western blot using monoclonal antibodies (generously provided by A. Cheung) and by LC-MS-MS coupled to spectral count analysis (Figure 1C, D). The data obtained by the two methods were well correlated and showed that the yield of MgrA protein was decreased by two-fold in the WT strain as compared to the *DrsaA* mutant strain while the repression was even more pronounced (about 10-fold) in the mutant strain expressing RsaA from a plasmid. This effect was specific since the yield of the ribosomal protein L21 remained identical in the various strains (Figure 1C).

We have previously shown that *S. aureus* RNAIII, which represses translation of target mRNAs, also affects their turnover. Therefore, we analyzed the steady-state level of mgrA mRNA using quantitative RT-PCR analysis on total RNA extracts prepared from the WT (HG001 and Becker) strains, the *DrsaA* mutant strains, and the same mutant strains transformed with a plasmid expressing RsaA (Figure 1E). The data showed that the level of mgrA mRNA was strongly enhanced in the *DrsaA* mutant strains while complementation caused an even greater decrease in mRNA yield than in the isogenic WT strains (Figure 1E). Because mgrA is transcribed in two transcripts [27], which initiate at  $\geq 124$  or  $\geq 302$  nts upstream of the initiation codon, northern experiments were also performed (Figure 1F). The level of the longest mRNA (from P2 promoter) decreased significantly at the late exponential phase of growth in both the WT and *DrsaA* derivative of HG001, indicating that the growth phase-dependent regulation of the P2 promoter is not under the control of RsaA (Figure 1F). Conversely, the level of the second and smaller transcript (from P1 promoter) was strongly enhanced in the mutant strain lacking the *rsaA* gene (*DrsaA*). These data suggested that the mgrA mRNA transcribed from the P1 promoter is under the control of RsaA while the mRNA formed from the P2 promoter is regulated independently of RsaA.

All in all, these data advocated that RsaA has a significant effect on the *S. aureus* proteome most probably through indirect interactions via the regulation of mgrA.

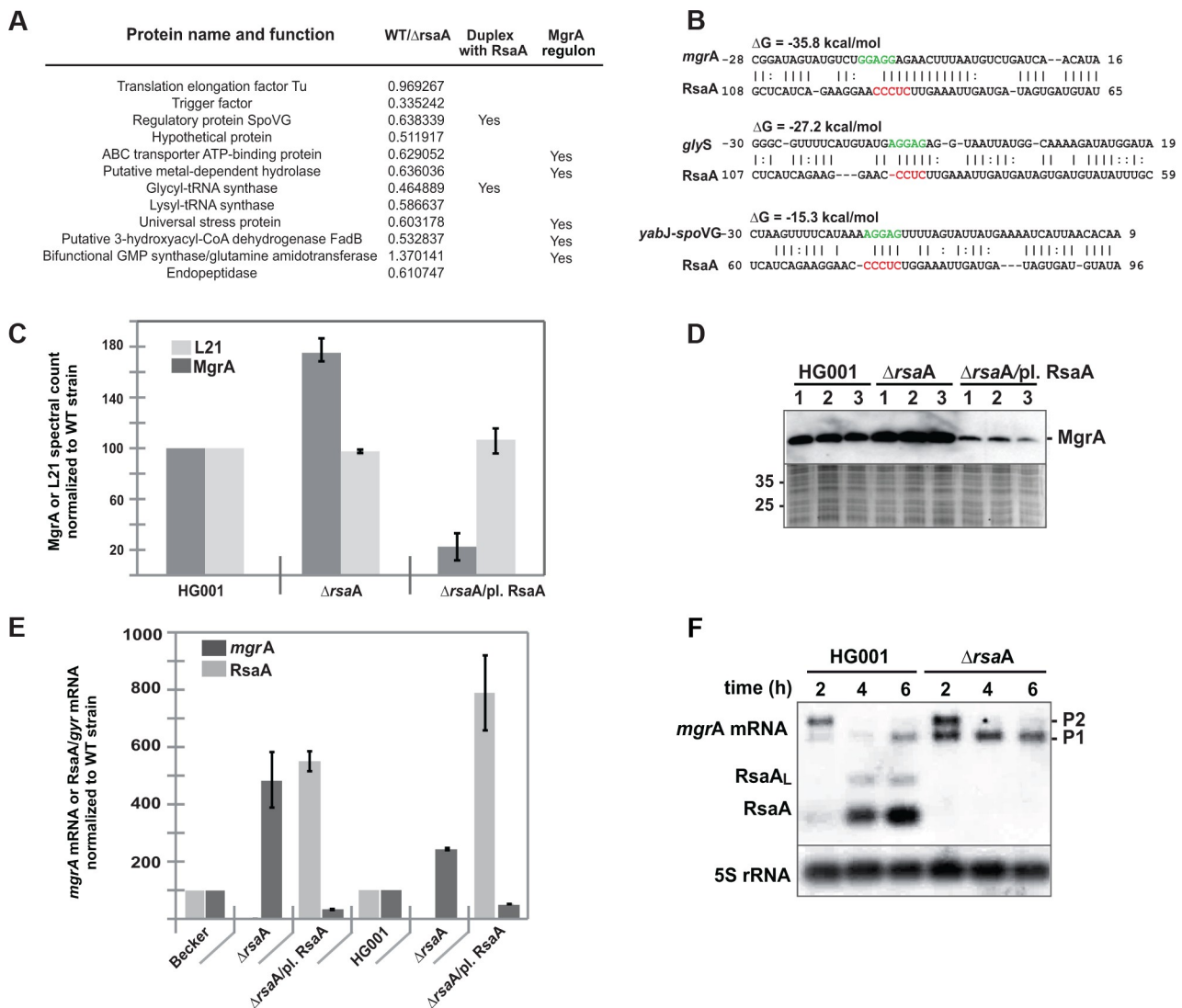


Figure 1. RsaA regulates the synthesis of several metabolic enzymes. (A) 2-D fluorescence difference gel electrophoresis (DiGE) performed on HG001 and  $\Delta$ rsaA-HG001 strains reveals proteins whose synthesis is affected by the deletion of RsaA. The proteins were identified by mass spectrometry analysis. Ratios correspond to the quantification obtained for HG001 versus  $\Delta$ rsaA strain (data from four independent experiments). In the last column are given the proteins, which belong to the MgrA regulon as shown by Luong et al. [26]. Statistical differences ( $p > 0.05$ , except for the bifunctional GMP synthase :  $p > 0.01$ ) between the two strains were obtained for each protein. (B) Potential base pairings between RsaA and potential mRNA targets are shown. The Shine and Dalgarno sequence (SD) of target mRNAs are shown in green while the conserved C rich motif in RsaA is in red. Other predicted RsaA-mRNA complexes are shown in Table S1. (C) Relative MgrA protein expression level was defined by LC-MS-MS followed by a spectral counting strategy. Protein extracts were prepared from HG001 (WT),  $\Delta$ rsaA mutant strain, and the same strain complemented with a plasmid expressing RsaA. In parallel to MgrA, we analyzed ribosomal protein L21 as internal control of constitutively expressed protein. The spectral count was normalized to the value determined for the WT strain. (D) Western blot performed with monoclonal antibodies against MgrA performed on protein extracts prepared from HG001,  $\Delta$ rsaA mutant strain and the same strain complemented with a plasmid expressing RsaA. As a control, a gel run in parallel with the same samples was stained with Coomassie blue to verify that each lane contained comparable amounts of protein. Protein size markers were run in parallel. (E) The expression of RsaA and of *mgrA* mRNA was quantified and normalized to the level of *gyrB* mRNA expression from total RNA extracts prepared from in vitro culture to the mid- or late-exponential phase of HG001 (WT), the  $\Delta$ rsaA mutant strain, and the same strain complemented with a plasmid expressing RsaA. (F) Northern experiments showing the steady state levels of *mgrA* mRNA in HG001 and in the  $\Delta$ rsaA mutant strains. Total RNAs were prepared after 2, 4 and 6 h of culture in BHI at 37°C. We used two different RNA probes to visualize the expression of RsaA and of *mgrA* mRNA. We have previously shown that a longer transcript including RsaA was also expressed [20]. This longer form of RsaA is known to be processed by RNase Y [37].  
 doi:10.1371/journal.ppat.1003979.g001

### RsaA binds to *mgrA* mRNA at two distinct regions

The predicted base pairing between RsaA (nts 65 to 108) and *mgrA* mRNA (nts 228 to +16) suggested that the RsaA-dependent repression of *mgrA* mRNA was governed by direct RsaA-mRNA pairing. To map more precisely the regions of interaction, we performed footprinting experiments using enzymatic and chemical

probing. The regions of interaction between *mgrA* mRNA and RsaA were probed using RNase T1 (specific for unpaired guanines), RNase T2 (specific for unpaired nucleotides with a preference for adenines), RNase V1 and the endoribonuclease III (specific for double stranded regions), and lead(II)-induced cleavages (specific for unpaired nucleotides). We also used

dimethylsulfate, which methylates adenines at the N1 position and cytosines at the N3 position to probe the structure of *mgrA* mRNA and its interactions with RsaA. The cleavage sites were analyzed using 5' end labeled RsaA while for *mgrA* mRNA, the cleavages and chemical modifications were analyzed by primer extension using reverse transcriptase. Experiments are shown in Figures 2A–B.

The probing data obtained with RsaA alone correlated well with the existence of three hairpin loops (Figure S2A) as previously published [12–14,19–23]. Indeed, single-strand specific RNase cleavages were mainly located in the apical loops and unpaired regions while RNase V1 cuts were predominantly found in the helices of RsaA. Furthermore, RNase III specifically cleaved helix H1, in agreement with the observation that RsaA decay is mediated by RNase III *in vivo* [23]. Enzymatic hydrolysis was then performed on 5' end labeled RsaA incubated with increasing concentrations of *mgrA* mRNA. Binding of *mgrA* mRNA affected the cleavages at two distant regions of RsaA. Strong protections were observed close to the C-rich motif of RsaA, i.e. at A83–85 against RNase T2, at U82–A94 against lead(II)-induced cleavages, and at G86 against RNase T1 (Figure 2C). Concomitantly, several enhanced RNase V1 cuts were found at U88–C89, and two strong

RNase III cleavages were observed at U82 and U87 of RsaA (Figure 2A). Unexpectedly, binding of *mgrA* mRNA also induced strong protections against single-strand specific probes in the hairpin loop H1 at residues A27 to G33 of RsaA revealing the existence of a second RNA binding site (Figure 2A, C).

The same experiments were performed on a *mgrA* fragment encompassing 124 nts upstream of the initiation codon and 258 nts of the coding sequence (Figure 2B). The enzymatic cleavages obtained on *mgrA* mRNA alone supported the existence of the long hairpin IV within the coding sequence as well as the stem-loop III structure, which partially sequestered the SD sequence (Figure S1B). The guanines of the Shine and Dalgarno (SD) sequence were indeed weakly accessible towards RNases T1 and T2. Binding of RsaA caused changes in the RNase cleavage patterns at two distant regions of the mRNA, namely around the RBS and in the hairpin loop IV located in the coding sequence (Figure 2C). In particular, strong RNase V1 cleavages occurred at positions C13 to A17, C70–C72 and C75, and two strong RNase III cleavages were observed at A+1 and C-5 of *mgrA* mRNA (Figure S2C). Concomitantly, RsaA-induced protections against RNase T2 were observed in the hairpin loop IV (at A67, A68, A74) and at nucleotide A-9 close to the SD sequence, and strong

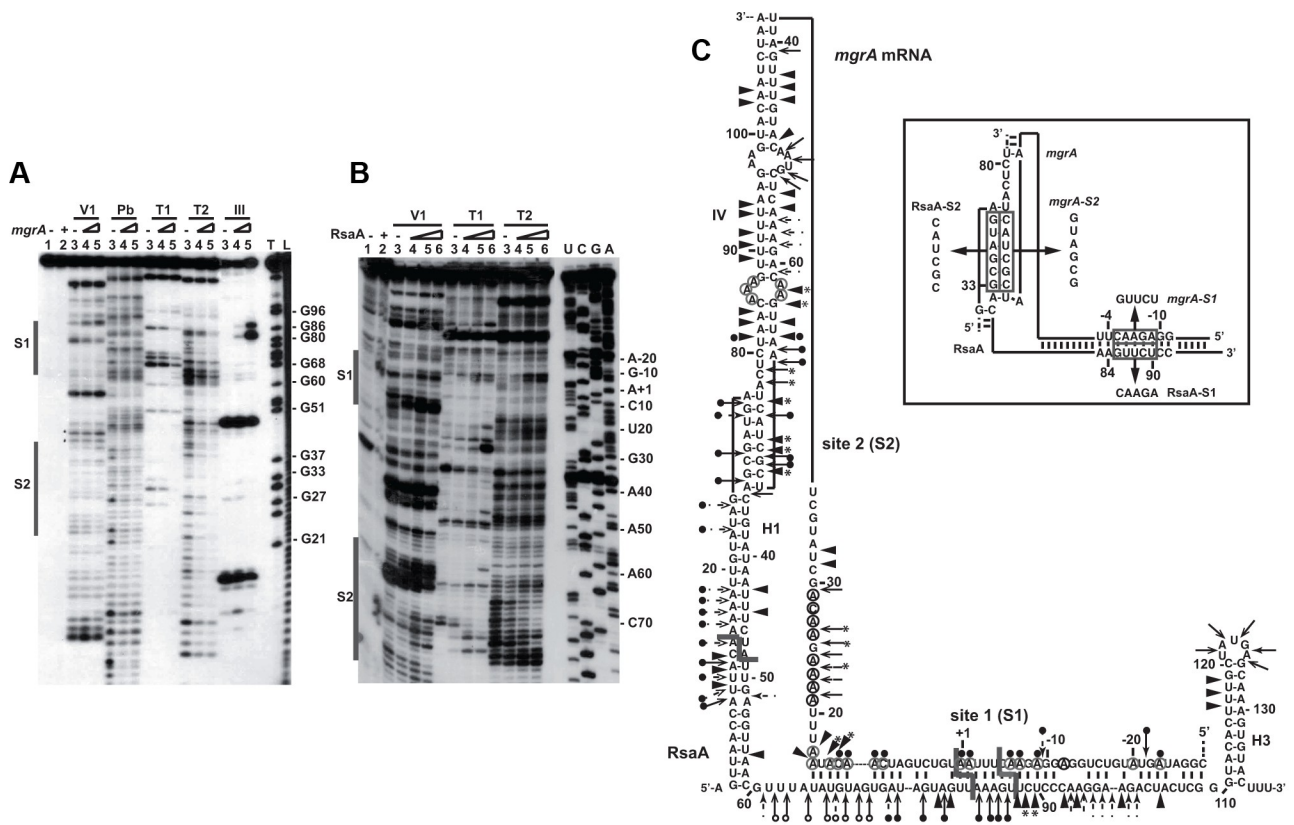


Figure 2. Probing the structure of the inhibitory RsaA-*mgrA* mRNA complex using enzymes and lead(II)-induced cleavages. (A) Reactions performed with 5'-end-labeled RsaA alone (lane 3) or in the presence of an excess of *mgrA* mRNA (lane 4, 25 nM; lane 5, 50 nM; lane 6, 250 nM). Lanes 1, 2: incubation controls on free RNA or bound to *mgrA* mRNA, respectively. The experiments were performed with RNase V1 (V1), lead(II) induced cleavages (Pb), RNase T1 (T1), RNase T2 (T2) and the endoribonuclease III (RNase III). Lanes T, L: RNase T1 under denaturing conditions and alkaline ladders, respectively. The two regions protected by *mgrA* are shown by bars. (B) Enzymatic reactions performed on *mgrA* mRNA alone (lane 3) or bound to increasing concentrations of RsaA (lane 4, 25 nM; lane 5, 50 nM; lane 6, 250 nM). Same legend as in A. Lanes U, C, G, A: dideoxy-sequencing reactions. (C) Structure of the RsaA-*mgrA* mRNA complex. Summary of the enzymatic cleavages on RsaA, *mgrA*, and the complex. Enzymatic cleavages are given as follows: RNase T1 (black and plain arrow), RNase V1 (black and plain arrowhead), RNase T2 (arrow). Circled adenines and cytosines in *mgrA* mRNA are reactive towards dimethylsulfate modification. The RNase III cleavages led to RNA fragments, which terminated by a two-nucleotide 3' overhang. Reactivity changes induced by the binding of RsaA are indicated as follows: black circles denote strong protection, enhancements and new RNase V1 cleavages are represented by asterisks. The mutations performed in RsaA and *mgrA* mRNA, are shown in the insert. doi:10.1371/journal.ppat.1003979.g002

reduction of reactivities of adenines at N1 against dimethylsulfate were located at A+1, A-1, A-6, A-7 and A-9 in the RBS (Figures 2B, 2C). Interestingly, the hairpin loop IV carries the unpaired nucleotides UCGCUACU, which are fully complementary to the hairpin loop I of RsaA (AGUAGCGA) (Figure 2C).

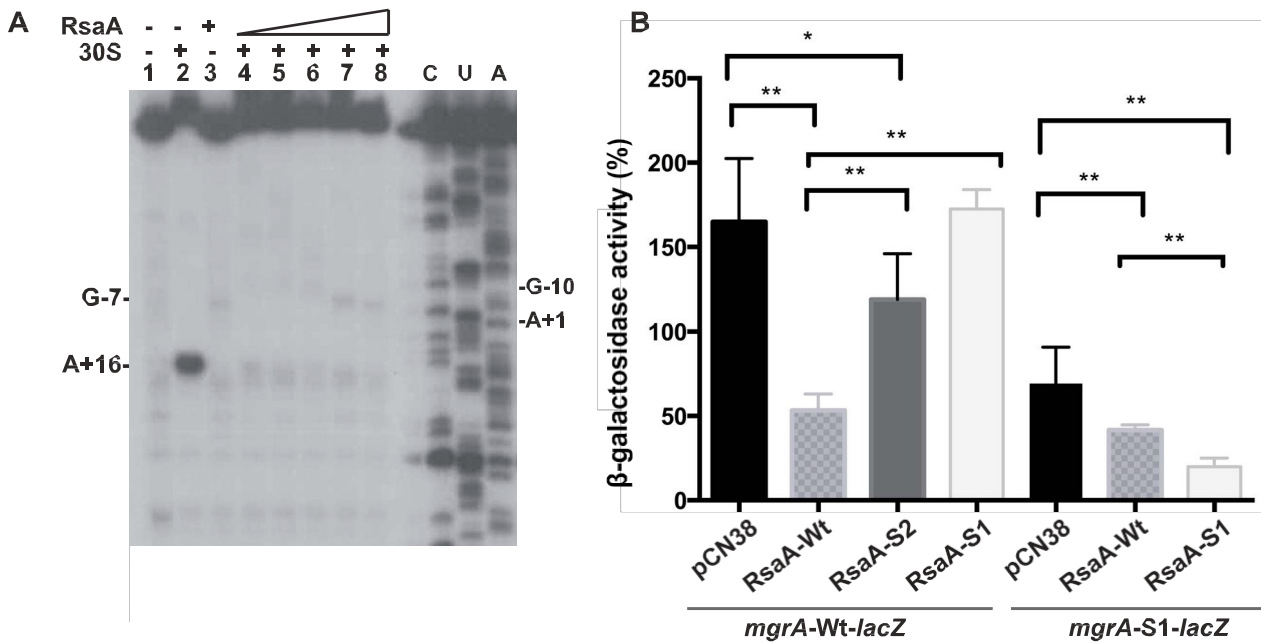
All together, these data revealed that the mgrA mRNA-RsaA complex is composed of a bipartite site, which implies the formation of an imperfect duplex sequestering the RBS of mgrA and a loop-loop interaction. Furthermore, the long imperfect duplex creates a specific site for the RNase III.

### RsaA inhibits ribosome binding in vitro and regulates mgrA translation in vivo

Structure probing strongly suggested that RsaA would repress translation initiation of mgrA. Toeprinting assays were first used to monitor the effect of RsaA binding on the formation of the initiation ribosomal complex involving the 30S ribosomal subunits, the initiator tRNA<sup>Met</sup>, and mgrA mRNA. Formation of the ternary complex, which blocked the elongation of a cDNA primer by reverse transcriptase, produced a toeprint at A+16 (Figure 3A). Binding of RsaA completely abolished the toeprint at a ratio of 1:1 (RsaA:mgrA mRNA). These data indicate that the RsaA-mgrA mRNA complex is rapidly formed and sufficiently stable to prevent the formation of the ribosomal initiation complex.

To further validate the in vivo relevance of RsaA-dependent repression of mgrA mRNA, we analyzed the expression of a reporter construct in *S. aureus* HG001-DrsaA strain expressing the wild-type (WT) RsaA. Mutations were introduced in both mgrA and RsaA at the two binding sites (S1 and S2), and were designed

to disrupt or restore base pairings (Figure 2C). The entire leader regulatory region of the mgrA gene produced from P1 (124 nts upstream of the initiation codon) including 258 nts of the coding sequence was cloned in-frame with the lacZ gene into the pTCV-lac shuttle vector (Table S2). This construct is under the control of a constitutive promoter (PrpoB).  $\beta$ -galactosidase activity assays were first performed in the *S. aureus* DrsaA mutant strain transformed with the pCN38 vector expressing WT RsaA from its own promoter. The  $\beta$ -galactosidase activity was 4-fold lower in the DrsaA mutant strain expressing WT RsaA than in the same strain transformed with the control vector (Figure 3B). To probe the importance of the first binding site, we introduced mutations in the stem-loop structure H2 of RsaA where 86GUUCU90 was substituted by CAAGA (RsaA-S1) to weaken the interaction with the RBS of mgrA mRNA (Figure 2C). The  $\beta$ -galactosidase activity was reproducibly enhanced 3-fold in the DrsaA mutant strain expressing RsaA-S1 as compared to the DrsaA strain expressing WT RsaA (Figure 3B). Thus, disruption of the predicted base pairings alleviated the repression of the reporter mgrA-WT-lacZ gene. We then introduced changes in the RBS of mgrA-lacZ reporter gene where -9AGAAC-5 was replaced by UCUUG to restore base pairing with RsaA-S1. This mutation weakened the strength of the SD-aSD helix since the first adenine of the SD sequence was substituted by a guanine. Indeed, the resulting mutated mgrA-S1-lacZ fusion produced less  $\beta$ -galactosidase than the WT fusion (Figure 3B). The expression of WT RsaA reduced only slightly the  $\beta$ -galactosidase synthesis while RsaA-S1 caused a significant reduction of the  $\beta$ -galactosidase synthesis from the fusion mgrA-S1-lacZ by a reproducible 4-fold factor (Figure 3B).



**Figure 3. RsaA-dependent repression of mgrA mRNA translation.** (A) Effect of RsaA on the formation of the initiation ribosomal complex. Lane 1: incubation control of mgrA mRNA; lane 2: formation of the initiation ribosomal complex containing mgrA mRNA, the initiator tRNA, and the 30S subunits; lane 3: incubation control of mgrA mRNA bound to RsaA; lanes 4 to 8: formation of the initiation ribosomal complex in the presence of increasing concentrations of RsaA: 20, 50, 150, 500 nM and 1 mM, respectively; lanes C, U, A: sequencing ladders. (B)  $\beta$ -galactosidase activity detected from different gene fusions.  $\beta$ -galactosidase activity was measured from PrpoB-mgrA (+1/+258):lacZ fusion or from PrpoB-mgrA-S1 (+1/+258):lacZ fusion expressed in HG001-DrsaA strain. The experiments were carried out in DrsaA-HG001 mutant strain transformed with the plasmid pCN38 containing no insert (pCN38) or expressing the WTRsaA, or its two mutant derivatives RsaA-S1 (85GUUCU90 changed to CAAGA) and RsaA-S2 (28GUAGCG33 changed to CAUCGC). The  $\beta$ -galactosidase activity was normalized for total cell density and is represented as a percentage of the uninhibited control. The results represented a mean of four independent experiments. Statistical differences \* ( $p > 0.05$ ) and \*\* ( $p > 0.005$ ) between each construct were obtained by using Anova test. doi:10.1371/journal.ppat.1003979.g003

The contribution of the second binding site in regulation was also analyzed. Mutations were introduced in the hairpin loop H1 of RsaA (RsaA-S2, 28GUAGCG33 changed to CAUCGC) to destabilize the loop-loop interaction. The data showed that the  $\beta$ -galactosidase synthesis obtained from *mgrA*-WT-lacZ was strongly enhanced in the *DrsA* strain expressing RsaA-S2 as compared to the same strain expressing the WT RsaA. Hence, disruption of the loop-loop interaction significantly alleviated repression of *mgrA* (Figure 3B).

Taken together, these data showed that RsaA represses the synthesis of MgrA at the post-transcriptional level through the formation of an RsaA-mRNA duplex, which sequesters the RBS of *mgrA* mRNA. Furthermore, the two distant regions of interaction contribute to the repression of MgrA synthesis.

### The two binding sites are essential for RsaA-dependent translation inhibition of *mgrA*

The effects of mutations that disrupt or restore base pairings were further analyzed for their ability to form RNA-RNA interactions and their ability to inhibit *mgrA* translation. The formation of the RsaA-mRNA complex was followed using gel retardation assays (Figure 4B). In vitro 5' end-labeled RsaA was incubated with increasing concentrations of *mgrA* mRNA. This experiment showed that *mgrA* mRNA binds to RsaA with a  $K_d$  value of around 20 nM. The initial rate of RsaA binding was estimated from a time-course analysis and resulted in an association rate constant of  $5.6 \times 10^5 \text{ M}^{-21} \text{ s}^{-21}$  indicating that the complex is rapidly formed. Mutations of the RBS of *mgrA* (*mgrA*-S1) had a strong effect on binding to WT RsaA while

compensatory mutations performed in RsaA-S1 restored the binding albeit with less efficiency ( $K_d$  value around 50 nM). Disruption of the loop-loop interaction caused a strong effect on the binding of *mgrA*-S2 to WT RsaA (Figure 4B). Therefore, both regions of interaction contribute to the stability of the RsaA-mRNA complex indicating that they act in a cooperative manner.

We then performed in vitro translation assays (PUREsystem) using the whole *mgrA* mRNA (expressed either from the P1 or the P2 promoter) carrying the sequence encoding a FLAG tag to visualize the protein by Western blot using FLAG-specific monoclonal antibodies. Addition of WT RsaA specifically repressed the translation of both *mgrA* mRNAs (Figure 4C). These data were well correlated with the in vivo reporter assays (Figure 3B). The same experiment was performed with mutated versions of *mgrA* mRNA (*mgrA*-S1, *mgrA*-S2) expressed from the P1 promoter. Consistent with the binding experiments, the addition of WT RsaA had no effect on the translation of *mgrA*-S1 while RsaA-S1 efficiently inhibited the translation of *mgrA*-S1 (Figure 4C). Furthermore, destabilization of the loop-loop interaction had a significant effect on translation repression because higher concentrations of WT RsaA were required to achieve inhibition with *mgrA*-S2 compared to WT *mgrA* (Figure 4C).

All in all, these data show that RsaA inhibits initiation of translation of *mgrA* and that the two regions of interaction contribute to binding and regulation.

### Deletion of RsaA reduces biofilm formation

We then analyzed whether RsaA is able to counteract the action of the master regulator of transcription MgrA. The effect of RsaA

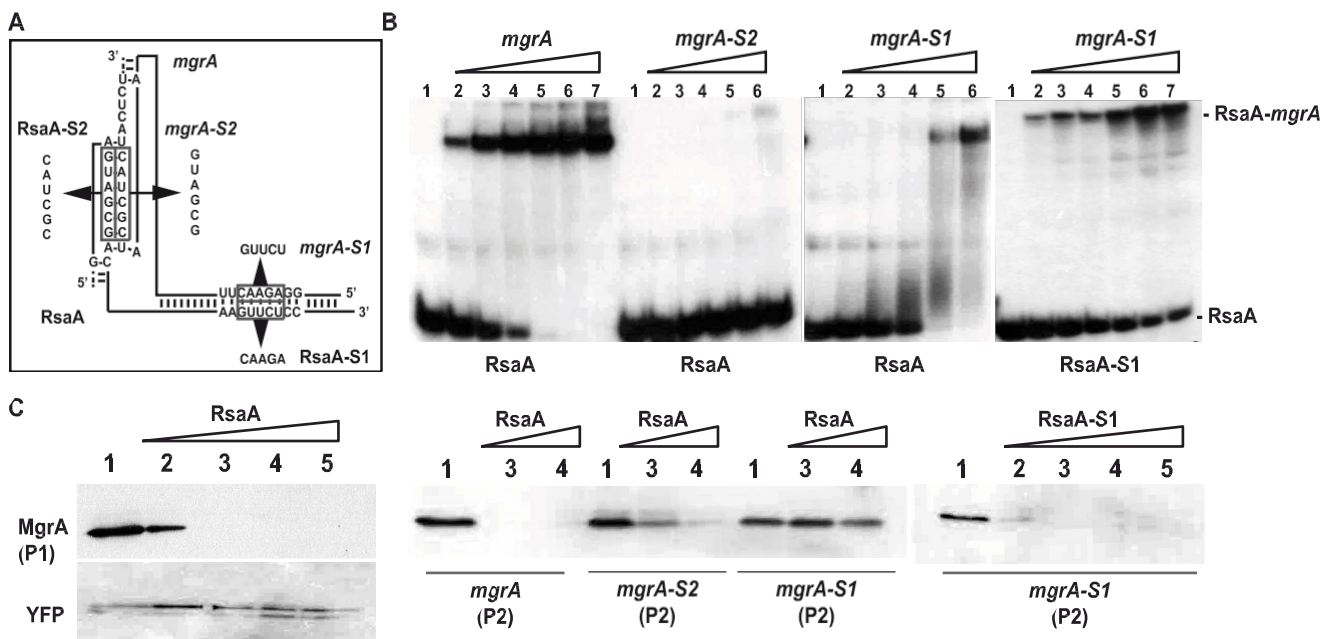


Figure 4. Contribution of the two binding sites of RsaA-mgrA duplex on translation repression. (A) Schematic drawings showing the two sites of interaction in the RsaA-mgrA complex. The various mutations are indicated. (B) Gel retardation assays to follow the formation of RsaA-mgrA complex. The 5' end-labeled wild-type (WT) RsaA or the mutant RsaA-S2 were incubated with increasing concentrations of either the WT *mgrA* mRNA or its two mutant derivatives (*mgrA*-S1, *mgrA*-S2): lane 2, 5 nM; lane 3, 10 nM; lane 4, 20 nM; lane 5, 50 nM; lane 6, 100 nM; lane 7, 500 nM. Lane 1: free 5' end-labeled RsaA. The bands corresponding to the free RsaA and to the complexes are specified. (C) In vitro translation assays using the PURESYSTEM. The reactions were performed with 10 pmol of WT *mgrA* or its two mutant derivatives (*mgrA*-S1, *mgrA*-S2) alone (lane 1) or in the presence of increasing quantities of WT RsaA or RsaA-S1 (lane 2: 5, lane 3: 10, lane 4: 30 and lane 5: 50 pmoles). The two mRNAs produced from the P1 or the P2 promoter were also analyzed. The same experiment was performed with YFP protein as the negative control. The proteins were separated on a SDS-PAGE 10% and were revealed using FLAG-specific antibodies.

doi:10.1371/journal.ppat.1003979.g004



was monitored on biofilm and capsule formation, which are two processes known to be regulated by MgrA [24,26,32,33].

The biofilm Ring Test was used to follow the kinetics of early biofilm formation. This approach focuses on adhesion and aggregation of bacteria to a substrate (Figure 5A). We analyzed various strain backgrounds, namely RN6390 ( $s^{B2}$ ), HG001 ( $s^{B+}$ ), Newman and Becker (clinical isolates,  $s^{B+}$ ) as well as their *rsaA* mutant and complemented counterparts. For all genetic backgrounds, the *rsaA* inactivation induced a delay in the kinetic of biofilm formation but this delay was variable according to the genetic background (Figure 5A). The delay in the kinetics of early biofilm formation of mutant vs WT strains was equal to 10 min for RN6390, 30 min for HG001, 1 h 30 for Becker and 2 h 10 for Newman. The *rsaA* complementation restored fully (RN6390 and HG001) or partially (Becker and Newman) the parental phenotype. The Becker DmgrA strain, generously provided by C. Lee, was used as a control. As expected [32], the inactivation of *mgrA* reduced the delay of early biofilm formation by about 1 h 30 in comparison to the WT strain. The double mutant DmgrA-DrsA Becker strain did not further reduce the delay of biofilm formation, supporting that RsaA regulated the biofilm formation via *mgrA* inhibition. Of note, RsaA plasmid complementation of the DmgrA-DrsA Becker strain did not change the biofilm phenotype obtained with the double mutant.

Quantification of mature biofilm after 24 h was also performed for all strains (Figure 5B). The level of mature biofilm was extremely abundant for HG001, intermediate for Becker and RN6390, and moderate for Newman. The inactivation of *rsaA* was associated with a significant decrease in biofilm thickness in the three strain backgrounds. Conversely, the plasmid complementation restored the parental WT phenotype and biofilm formation was even slightly higher due to the overexpression of RsaA. As a negative control, transformation with the empty plasmid showed no effect (Figure 5B). Furthermore, *mgrA* inactivation in the Becker strain induced a high increase in mature biofilm formation, as previously published [32]. As expected, the double mutant DrsA-DmgrA Becker strain, and the same strain complemented with a plasmid expressing RsaA, produced a phenotype similar to that of the single DmgrA Becker strain.

The *icaADBC* operon encoding the extracellular polysaccharide adhesin PIA/PNAG (polysaccharide intracellular adhesion/poly-N-acetylglucosamine) is responsible for some of the biofilm-positive phenotypes in *S. aureus* [34]. However, *S. aureus* can produce an alternative *ica*-independent biofilm via the expression of *mgrA* [32]. To rule out an involvement of *icaADBC* in the RsaA-dependent biofilm formation, we quantified PIA/PNAG production using an anti-PIA/PNAG specific antibody in the three strain backgrounds. We saw that the inactivation of *rsaA* had no effect on PIA/PNAG production for the equivalent quantity of bacterial lysates (Figure S3).

In summary, these data suggested that RsaA is acting upstream of MgrA and activates biofilm formation through the inhibition of MgrA synthesis.

### RsaA inhibits capsule production and prevents opsonization of *S. aureus*

To study whether RsaA was involved in capsule regulation, we quantified capsule production after an overnight culture using anti-CP5 or anti-CP8 specific antibodies. Using equivalent amounts of bacterial lysates, we found that the Becker strain (capsular polysaccharide type 8, CP8) produced more capsular polysaccharides than Newman and HG001 (capsular polysaccharide type 5, CP5). The inactivation of *rsaA* induced a significant increase in CP8 biosynthesis in the Becker strain and in CP5 in the

Newman strain while only a slight increase in CP5 was detectable for HG001 (Figure 6A). The expression of RsaA from a plasmid restored the parental WT phenotype for all strains. The same results were obtained after 6, 9 and 12 h of culture (data not shown). To verify that the regulation of capsule production by RsaA resulted from *mgrA* repression, we again analyzed the phenotype of the double mutant DrsA-DmgrA in the Becker strain background. Compared to the WT, the level of CP8 production was considerably reduced in the single DmgrA mutant strain as well as in the double mutant DrsA-DmgrA strain. Noticeably, we reproducibly found that the complementation of the double mutant with a plasmid expressing RsaA further reduced the level of CP8 production, suggesting the existence of an *mgrA*-independent repression effect of capsule production. This additional effect might be attributed to the RsaA-dependent inhibition of SpoVG (Figures 1, S1), a protein involved in the capsule formation [29].

To illustrate the accumulation of capsular polysaccharides in DrsA mutant strains, indirect fluorescence was performed to investigate the capsule formation of the Becker strain and its derivatives after 7 h of growth in BHI. Strikingly, the capsule was highly labeled only in the DrsA mutant Becker strain contrasting with the WT parental and complemented strains (Figure 6B). To verify the specificity of the secondary antibody, the same experiment was performed without the primary CP8 antibody, and as expected, all strains were only labeled in green (data not shown).

Because capsular polysaccharides protect *S. aureus* against opsonophagocytic killing by polymorphonuclear leukocytes (PMNs) [35,36], we measured the percentage of bacteria killed after contact with serum and PMNs in strains expressing or deleted for RsaA. The data were not statistically significant because a large variability of bacterial survival was observed due to inter-PMNs donor differences. However, in the mutant strains, fewer bacteria were killed than in the parental strains expressing RsaA (particularly in the Becker strain background) in each replicate (Figure S4).

These data showed that RsaA significantly represses capsule synthesis in all strain backgrounds and renders *S. aureus* more sensitive to opsonophagocytic killing by PMNs.

### RsaA attenuates virulence in mice infectious models

Because RsaA impacts both biofilm and capsular formation, we selected two animal models in mice that were considered relevant with respect to these processes, namely, the catheter infection model (for biofilm) and the intraperitoneal infection model (for the capsule).

For the *in vivo* catheter infection model, sections of polyurethane catheter (1 cm) were implanted sub-cutaneously into C57BL/6 mice and filled with about  $4 \times 10^4$  bacteria/25 ml (Becker WT and DrsA mutant strains) in PBS (6 mice/group). The catheters were implanted for 3, 6 or 10 days. No difference in weight was observed between the mice infected by the WT and mutant strains (data not shown). Both strains led to a local infection with swelling but the size of the swelling was not significantly different between the studied strains (data not shown). At day 3, the concentration of colony-forming units observed on the catheters was significantly higher in mice infected with the WT strain than those infected by the DrsA mutant strain ( $p = 0.0173$ ) (Figure 7A). Between days 6 and 10, this difference was no longer observed in mice infected with WT and mutant strains, likely due to the saturation of the growth within the catheters (Figure 7A). However, between days 3, 6 and 10, more colony-forming units were observed in the surrounding tissues of the mice infected with the WT strain while

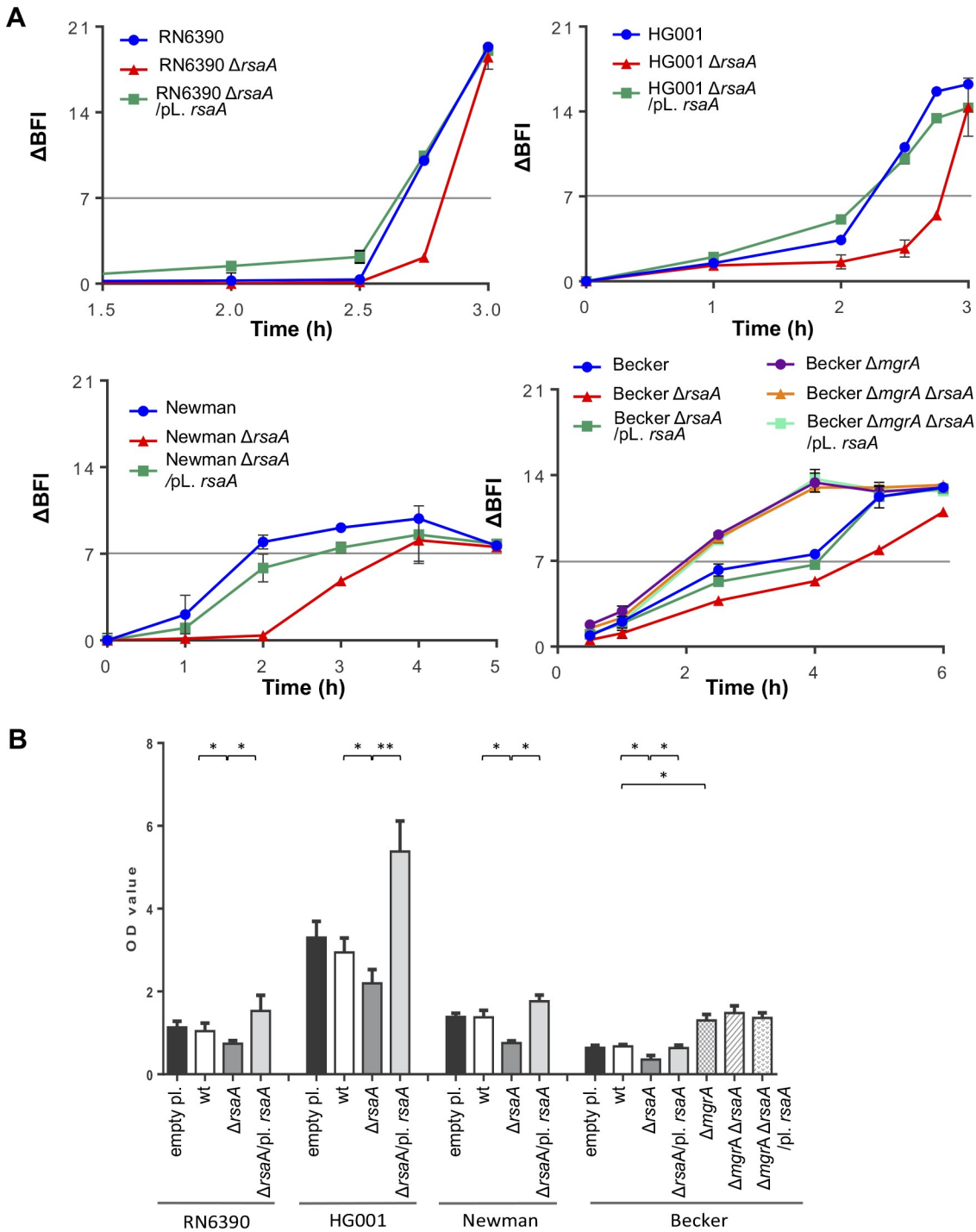


Figure 5. Effect of RsaA on biofilm formation. (A) Kinetics of the early phase of biofilm formation in different *S. aureus* backgrounds with the Biofilm Ring Test at different incubation times. Standardized bacterial cultures were incubated in 96 well micro-titer plate in the presence of magnetic beads at 37°C. A plate reader measures the number of beads in each well before and after magnetization. The beads are gathered in the center of the well if no biofilm is formed whereas the beads are dispersed if they are immobilized into the biofilm. The Biofilm Forming Index (BFI) corresponds to the mobility of beads under magnet action. Results are expressed as the DBFI corresponding to the difference between a control well (without bacteria) and the studied sample. A value of DBFI  $\geq$  7 means that the biofilm is formed. Data obtained for each strain were shown with a color code as given in the figure: blue for the parental wild-type (RN6390, Newman, Becker) strains, red for the *DrsaA* mutant strain, and green for the complemented strain expressing RsaA from a plasmid. (B) Analysis of the mature biofilm formation with crystal violet staining. Overnight bacterial cultures were standardized and incubated 24 h at 37°C in rich medium supplemented with glucose 0.25%. The mature biofilm was stained by crystal violet 0.1%, released with acetic acid 33%, and quantified by OD<sub>570 nm</sub>. The mean and the standard deviations are presented on the histogram. Statistical differences \* ( $p < 0.05$ ) and \*\* ( $p < 0.005$ ) between each strains were obtained by using Anova test. doi:10.1371/journal.ppat.1003979.g005

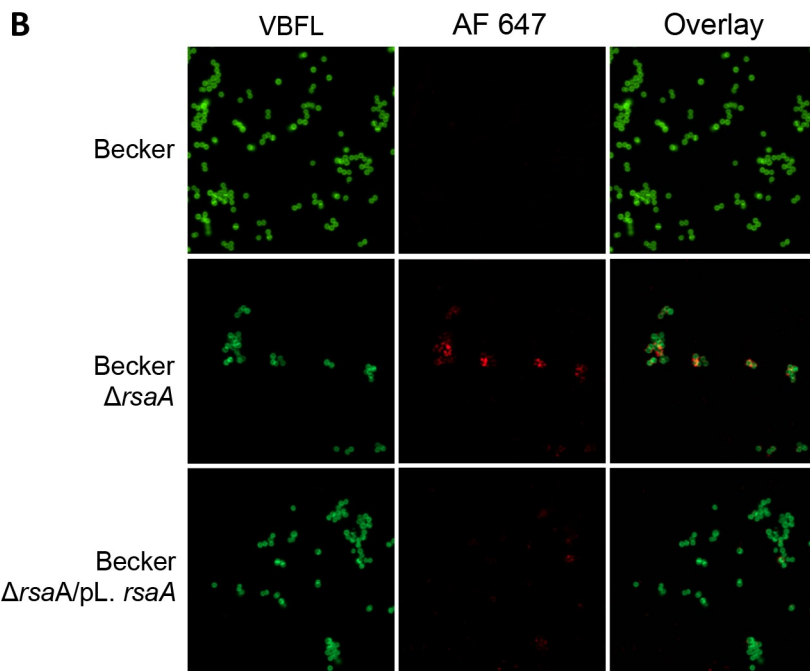
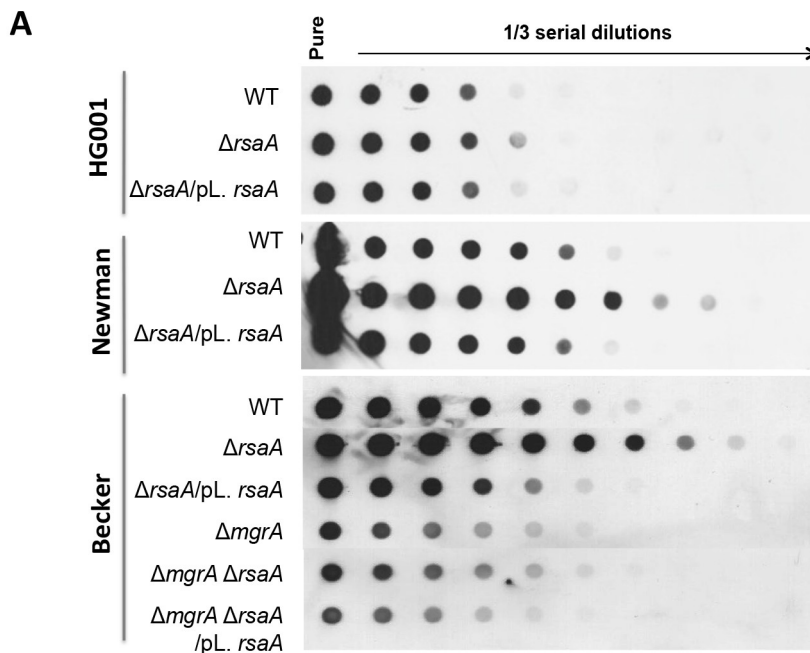


Figure 6. Effect of RsaA on capsule formation. (A) Dot-blot analysis of capsular polysaccharide accumulation in wild-type strains and its corresponding DrsaA mutant and complemented strains after growth in BHI for 16 h. Standardized bacterial suspensions were treated with lysostaphine, DNase I and proteinase K, and 1/3 serial dilutions were spotted onto nitrocellulose membranes. Capsular polysaccharides were detected with an anti-CP5 (HG001 and Newman) or CP8 (Becker) antisera. The DrsaA mutant produced a higher level of capsular polysaccharides and the DmgrA mutant produced a lower level of capsular polysaccharides. The inactivation of both *rsaA* and *mgrA* revealed that the repression of capsule production by *rsaA* achieved via the repression of *mgrA*. (B) Capsule production in Becker strain and its derivatives grown for 7 h in BHI at 37°C. CP8 expression was determined by indirect immunofluorescence of WT strain, the DrsaA mutant strain (LUG2010) and the mutant strain complemented with a plasmid expressing RsaA (LUG2032). The bacteria were labeled with Bodipy-vancomycin and capsule was labeled with an anti-CP8 antibody and a rabbit anti-IgG coupled with a fluorochrome AlexaFluor 647.  
doi:10.1371/journal.ppat.1003979.g006

the number of colony-forming units of the mutant DrsaA strain remained unchanged (Figure 7B).

For the bacteremia model, we included the two major capsular types of *S. aureus* (type 5 and 8), namely Newman (type 5) and

Becker (type 8), their DrsaA mutant derivatives, as well as their complemented counterparts (studied at the early time point only). Mice were challenged with sub-lethal doses of each strain by intraperitoneal injection ( $1 \times 10^6$  bacteria/500  $\mu$ l). Animals challenged

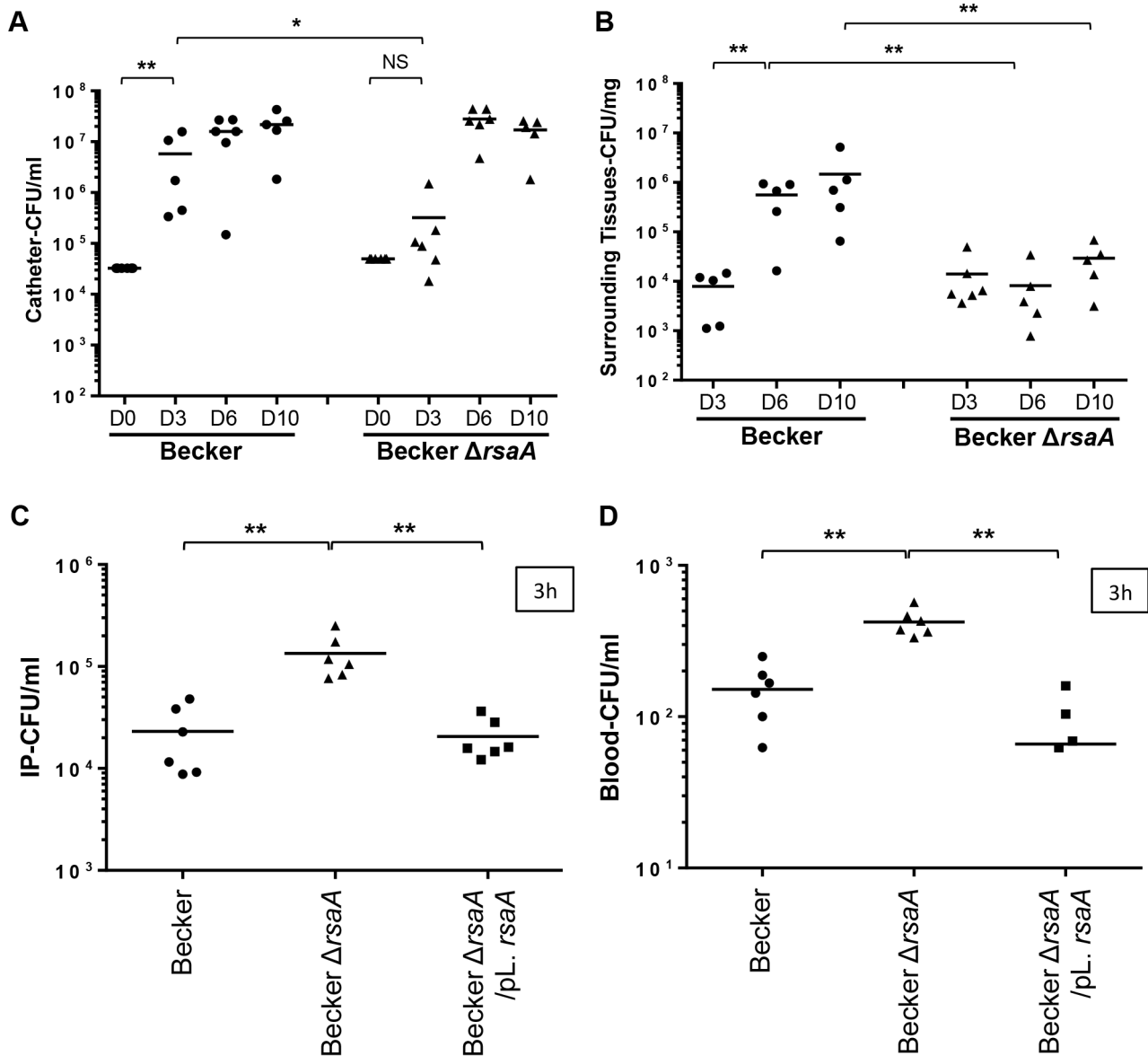


Figure 7. The impact of RsaA on virulence in vivo animal models. Two in vivo infections models were developed: a catheter infection model (to mimic a chronic infection) and a bacteremia model (to mimic an acute infection). Quantification of bacteria into catheter (A) and in the surrounding tissues (B) after 3, 6 and 10 days post-infection. After euthanasia of mice, the catheters and surrounding tissues were removed and processed as described in Materials and Methods. Bacteria were enumerated by serial dilution and plate counting on agar plates. The results were expressed in number of CFU/ml. Each point designates the number of bacteria for each mouse and the black line represents the mean of each tested condition. The second model was performed by intraoperitonnal (IP) injection of  $1.6 \times 10^6$  bacteria. At 3 h post-infection, viable bacteria present into IP (C) and in blood (D) were enumerated by serial dilution and plate counting on agar plates. The results were expressed in number of CFU/ml. Statistical differences \* ( $p < 0.05$ ) and \*\* ( $p < 0.005$ ) between each groups were obtained by using Mann Whitney test. doi:10.1371/journal.ppat.1003979.g007

with the strongly encapsulated strains (Becker-DrsaA) showed a significantly ( $p < 0.005$ ) higher bacteremia level at 1 h and 3 h after inoculation compared with animals infected with the parental and complemented strains (both less encapsulated) (Figures 7C and S5A). Furthermore, after 1 h and 3 h, the blood concentrations of DrsaA bacteria were significantly higher ( $p < 0.05$ ) than those of parental and complemented strains (Figures 7D and S5B). The same results were obtained with Newman background (Figures S5C–F). Evolution of the infection was followed by collecting the mice spleen for bacterial numeration. After 1 and 2 days, the bacterial counts in the spleen were generally higher in the DrsaA

mutant strain- vs WT strain-infected animals for both Becker (Figure S5H) and Newman (Figure S5G) backgrounds. However, this difference only reached significance in the spleen after 1 day with Newman strain and after 2 days with Becker strain.

Our data obtained with both animal models suggested that RsaA diminishes invasiveness but favors chronic infections and local colonization.

RsaA is expressed by clinical isolates of *S. aureus*

The observation that RsaA impacted oppositely biofilm formation and invasiveness raised the question as to whether

clinical isolates would express differential levels of RsaA depending on the clinical situation in which they had been sampled. To this end, a series of 18 isolates reflecting the commensal status of *S. aureus* (nasal colonization) as well as *S. aureus* diseases in various forms (chronic lung infection in cystic fibrosis patients, non-invasive acute skin and soft tissue infections, severe invasive disease illustrated by community-acquired pneumonia and infective endocarditis), were tested for RsaA expression at mid and late-exponential phases by RT-qPCR. The results revealed no specific correlation between RsaA expression levels and invasiveness, both at the mid and late-exponential state of growth (Figure S6A). Of note, all reference strains tested (HG001, Becker, Newman, USA300 LAC and SF8300, and the European CA-MRSA ST80) showed equivalent levels of RsaA production (Figure S6B).

## Discussion

### RsaA regulates the synthesis of the master global regulator MgrA

The virulence of *S. aureus* is in part conferred by numerous effector proteins that interact with the host. The levels of these virulence factors are controlled by a set of more than 20 transcriptional regulatory proteins [2]. However, *S. aureus* strains are genetically highly diverse and it is well recognized that their regulatory networks are not uniform. In order to determine the virulence traits of a defined strain, we must understand how the regulators interact with each other and how novel regulators such as non-coding RNAs influence these networks.

In this work, we show that the small non-coding RNA RsaA functionally links two global regulators, the  $s^B$  and MgrA proteins. The transcription of RsaA was shown to be under the control of  $s^B$  protein [20], while its fate was determined by two endoribonucleases, the double-strand specific RNase III [23] and RNase Y [37]. Our data indicate that the *mgrA* mRNA is repressed at the stationary phase of growth in an RsaA-dependent and RsaA-independent manner. The *mgrA* mRNA is transcribed from two promoters, and the mRNA with a longer 5'UTR was no longer produced at the late-exponential phase when RsaA was produced (Figure 1F). Whether the regulation is mediated by a protein- or another sRNA remains to be studied. Using a combination of approaches *in vivo* and *in vitro*, we showed that the primary effect of RsaA is to repress the translation of *mgrA* mRNA by preventing the formation of the ribosomal initiation complex. The steady state level of the mRNA was also substantially enhanced in the *DrsaA* strain suggesting subsequent degradation of the repressed mRNA. A reduction of *mgrA* mRNA from the P1 promoter was also observed at the post-exponential phase of growth in various  $s^{B+}$  strains, and the levels of *mgrA* mRNA were found to be higher in RN6390 ( $s^{B2}$ ) than in SH1000 ( $s^{B+}$ ) [38]. This was correlated *in vivo* with the level of MgrA protein, which was found to be higher in the *DrsaA* mutant strain than in the isogenic WT strain (Figures 1C, D). *In vitro* binding assays showed that RsaA binds to *mgrA* mRNA with a fast association rate constant, which is essential for efficient repression because the sRNA has to bind to its target mRNA within a short time frame before the stable ribosomal initiation complex is formed [39]. Two stretches of highly conserved and unpaired nucleotides of RsaA bind to two distant regions of *mgrA*, a C-rich motif of RsaA masks the SD sequence of *mgrA* while a loop-loop interaction occurs between two hairpin loop structures found in the 5' end of RsaA and in the coding sequence of *mgrA* (Figure 8A). Mutational analysis revealed that the two RsaA binding sites are required for optimal binding and *mgrA* translation repression *in vivo* (Figures 3 and 4).

It is worth noting that RsaA and RNAIII share very similar regulatory mechanisms [11,40] suggesting that the importance of bi-partite RNA-RNA binding sites in translation repression appears to be more general than previously expected in *S. aureus* (Figure 8A). In both cases, their mRNA targets are characterized by two regulatory modules: the SD sequence interacts with a C-rich motif within the regulatory RNAs to prevent ribosome binding and to create a specific RNase III binding site, while a hairpin motif constitutes either a second RNA binding site to stabilize the RNA-RNA interaction or to promote a specific binding site for RNase III [5,10]. For the RsaA-*mgrA* complex, the second interaction site involves a loop-loop interaction, which greatly enhances the stability of the complex and the efficiency of translation repression (Figure 4). RNase III-dependent cleavages were observed at the site of interaction formed between the RBS of *mgrA* and the C-rich motif of RsaA. Note that *mgrA* mRNA and RsaA were co-immunoprecipitated with RNase III and the half-life of RsaA was significantly enhanced in the *Drc* mutant strain [23]. As it was shown for RNAIII, it is tempting to propose that RNase III also contributes to the degradation of the translationally repressed *mgrA* mRNA. In addition, since MgrA positively autoregulates its own transcription [27], RsaA-dependent repression of *mgrA* translation might indirectly inhibit *mgrA* transcription.

### RsaA-dependent regulatory circuit might be linked to the quorum sensing system

By regulating the translation of MgrA, RsaA participates in a complex regulatory network involving two global regulators, which modulate the levels of RNAIII, thus affecting biofilm formation and capsule synthesis (Figure 8). RsaA expression is positively regulated by the alternative  $s^B$  factor [20], an observation that is well correlated with our data. Indeed, *rsaA* inactivation affected less strongly the kinetics of biofilm development in RN6390 ( $s^{B2}$ , 10 min) than in HG001 ( $s^{B+}$ , 30 min) due to the fact that the sRNA was not strongly expressed in RN6390 [20]. RsaA synthesis follows the same pattern as  $s^B$ , which is activated at a basal level at the exponential phase and strongly enhanced at the stationary phase [41]. Furthermore,  $s^B$  affected the transcription of the virulence genes in an opposite manner to RNAIII, and this effect would be in part the consequence of lower RNAIII levels in strains harbouring a fully active  $s^B$  [41].  $s^B$  was also shown to enhance resistance to oxidative and UV stresses, biofilm, and capsule formation. The  $s^B$ -dependent activation of biofilm formation occurs via the *icaADBC* operon [42,43]. We show that RsaA does not perturb the production of the PIA/PNAG polysaccharides produced by the *icaADBC* operon (Figure S2), but affects pathways that are predominantly dependent on *mgrA*.  $s^B$  is also known to activate the capsule synthesis via the two-component system ArlRS and *yabJ-spoVG* [33]. It is worth noting that *yabJ-spoVG* is a potential mRNA target of RsaA whose expression was altered in the proteome analysis, and for which base pairings with RsaA have been predicted (Figure 1). The main target of RsaA, which is the transcription factor MgrA, regulates more than 350 genes [26]. MgrA is activated by the *arlRS* operon, whose expression is controlled by  $s^B$  [26]. Taken together, the expression of MgrA is under the control of an incoherent feed-forward loop, involving a positive loop mediated through ArlRS and a negative loop mediated through RsaA (Figure 8). In addition, a recent study revealed that the DNA binding activity of MgrA is modulated through specific phosphorylation by the eukaryotic like kinase phosphatase pair Stk1-Stp1 [44]. Hence, the expression of this master regulator is controlled through multiple pathways. MgrA is central to the regulation of biofilm and represses the production of two out of the three components of the biofilm, i.e. surface proteins

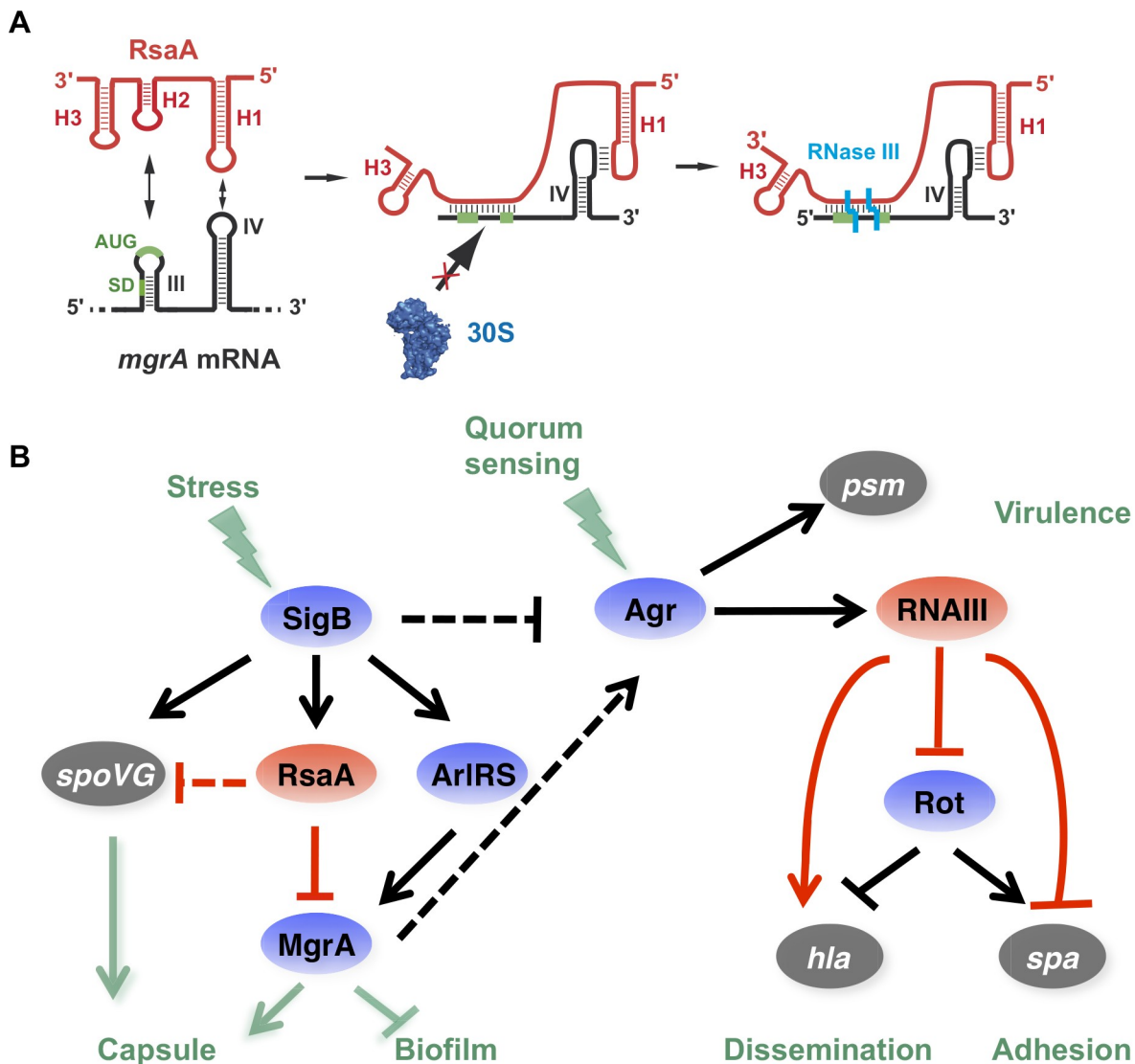


Figure 8. RsaA and its regulatory circuits. (A) Schematic drawing summarizing the regulatory mechanism. RsaA binds to *mgrA* mRNA and inhibits translation by preventing the 30S subunit binding, and recruits RNase III to induce simultaneous degradation of both RNAs. (B) RsaA is activated by  $s^B$  and in turn represses *mgrA* translation. RsaA is thus indirectly linked to RNAIII regulatory networks because MgrA activates *agrACDB* expression while  $s^B$  represses it [32,41]. Arrows are for activation, bars for repression. In blue are the transcriptional protein regulators, in red the regulatory RNAs and in grey the virulence factors. Red lines corresponded to post-transcriptional regulation and black lines to transcriptional regulation. The regulatory events for which direct regulation is not yet demonstrated are shown by dotted lines. doi:10.1371/journal.ppat.1003979.g008

and exogenous DNA release [32]. Indeed, MgrA negatively regulates the DNA release by suppressing *cidA* expression and activating *lrgAB* encoding a peptidoglycan hydrolase and an inhibitor of murein hydrolase activity, respectively [32]. Through its synergistic action on these two loci, MgrA inhibits cell autolysis and the release of DNA within the biofilm [32]. The second control exerted by MgrA is to activate the *agr* system, in a quorum sensing-independent pathway, to inhibit the synthesis of surface proteins essential for adhesion [25,26]. Although *mgrA* mutations significantly reduced RNAIII levels [25,27], the protein also bypasses the quorum sensing system by directly activating the *hla* gene and repressing *spa* gene expression via *sarS*. As with the *agr* system, MgrA also has a positive effect on capsule synthesis [26], and in a rat infective endocarditis model, MgrA was shown to be the major regulator of capsule formation *in vivo* [28]. In correlation with these studies, we showed that inactivation of *rsaA* alleviates

the repression of *mgrA* to increase capsule synthesis (Figure 6). Therefore, we propose that RsaA functionally links the global regulators  $s^B$  and *mgrA*, with potential consequences on the temporal expression of virulence determinants ([25,26]; Figure 8).

We demonstrate that RsaA activates biofilm formation (Figure 5) and represses capsule synthesis (Figure 6) most probably via the repression of *mgrA* translation. The regulation of these two phenotypes has been extensively studied in pathogenic bacteria demonstrating a complex interdependency. Studies in *Staphylococcus haemolyticus* and *Streptococcus pneumoniae* highlighted that encapsulated strains produced little or no biofilm, whereas non-encapsulated strains produced a strong biofilm [45]. Moreover, inactivation of the *cap* operon of *S. haemolyticus* JCS1435 (sharing 57% of sequence similarity with *S. aureus cap*) induced biofilm formation [46] suggesting that the capsule would partially mask or inhibit the surface proteins essential for biofilm formation. In our

study, we demonstrated this imbalance in favour of biofilm production for the HG001 strain that is minimally encapsulated but produces lots of biofilm; or in favour of the capsule synthesis for Becker and Newman strains that produce small amounts of biofilm but abundant capsule (Figures 5 and 6). The mechanistic details of this imbalance are not totally understood and may certainly involve other factors besides MgrA and RsaA.

### RsaA favors chronic infection and attenuates invasiveness

The pathogenic bacterial chromosome is composed of the core genome plus accessory elements. The accessory genome contains genes that have been horizontally acquired through mobile genetic elements (plasmids, phages, pathogenicity islands). The acquisition of these elements, which often encode virulence determinants and antibiotic resistance, has been accompanied by subsequent adaptation of the bacteria to successfully incorporate these novel elements [47]. Some pathogenic bacteria harbour genes that can hinder the expression of acquired virulence factors, and certain genes must be inactivated for the expression of full pathogenicity [48]. Other bacterial genes lead to hypervirulence when they are inactivated [49]. This family of genes was identified in several pathogenic bacteria such as *Listeria monocytogenes* [50] or *Salmonella enterica* [51]. For instance, in *S. enterica* the leucine-responsive regulatory protein (Lrp) represses key virulence factors involved in the invasion of host cells [51]. These genes encode proteins, but recently an antisense RNA in *S. typhimurium* was shown to kinetically control the virulence determinant MgtC, which is required for defense against the host immune system [52]. Deletion of this antisense RNA causes a hypervirulent phenotype. In the present study, we demonstrate that RsaA, a non-essential ncRNA present in all *Staphylococcus* species [20] reduces the virulence of *S. aureus*. Indeed, the deletion of *rsaA* increases the invasiveness of the bacteria in the mice sepsis model and resistance to opsonophagocytic killing. Conversely, RsaA promotes biofilm formation and catheter colonization in mice, potentially promoting harmful colonization and chronic infections in human. Noticeably, quantification of RsaA expression from human clinical isolates revealed that Rsa is functional in all clinical isolates of *S. aureus* studied so far [53] but is not differentially expressed depending on the level of invasiveness ranging from commensalism to deep-seated infections (Figure S6). Of note, we also showed previously by direct quantification from clinical samples that RsaA is expressed in all clinical samples including nasal secretion [53]. Since *Staphylococci* are mainly present as commensal bacteria of warm-blooded organisms (e.g. mammals and birds), the functionality of RsaA as a suppressor of virulence may have been positively selected through evolution because it is favorable for commensalism and saprophytic interactions with the host. We thus postulate that mutations leading to RsaA deficiency would be counter selective, perhaps because hypervirulence would be deleterious to the human host that is the natural reservoir of this bacterium.

## Materials and Methods

### Ethics statement

All mouse protocols were carried out in strict accordance with the Directive 2010/63/EU revising Directive 86/609/EEC on the protection of animals used for scientific purposes. This directive was translated in the French regulation as the Décret nu2013-118 of Feb 2013 under the jurisdiction of the Ministry of Education, Research and Technology. The protocols were approved by the CECCAPP (Comité d'Evaluation Commun au Centre Léon Bérard, à l'Animalerie de transit de l'ENS, au PBES et au laboratoire P4) under

identification numbers ENS\_2011\_033 and ENS\_2012\_051 for biofilm and bacteremia models, respectively.

### Strains, plasmids and growth conditions

The *Staphylococcus aureus* strains, clinical isolates, and plasmids used in this study are listed in Table S2. *Escherichia coli* strain DH5a (Promega, Madison, USA) was used as a host strain for plasmid constructions. *S. aureus* strain RN4220 [54] was used as the recipient in electroporation of the constructed plasmids. *Staphylococci* were grown either on GP agar plates (1% peptone, 0.5% yeast extract, 0.5% NaCl, 0.1% glucose, 1.7% agar, pH 7.2) or in brain-heart infusion (BHI) supplemented with erythromycin (5 mg/ml) or chloramphenicol (10 mg/ml) when appropriate. *E. coli* strains were cultivated in Luria-Bertani medium (1% peptone, 0.5% yeast extract, 1% NaCl). Total DNA and plasmid DNA were prepared using DNAeasy tissue Kit Qiagen and Qiaprep Miniprep respectively (Qiagen, Valencia, USA). Transformation of *E. coli* DH5a was performed by thermic shock and *S. aureus* strains were transformed by electroporation (Bio-Rad gene Pulser).

The deletions/replacements *DrsaA/aphA-3* mutants of *S. aureus* RN6390 (LUG1450), HG001 (LUG1630), Newman (LUG1680), Becker (LUG2010) and CYL1040 (LUG2009) strains were obtained using pMAD [55]. The kanamycin resistance gene *aphA-3* was cloned in pMAD between two DNA fragments corresponding to the chromosomal regions upstream and downstream of the *rsaA* coding sequence using primers *rsaA-76/rsaA-1120* and *rsaA-1246/rsaA-2194*, respectively (Table S3). The resulting plasmid (pLUG754) was electroporated into RN4220 recipient strain and then transferred to RN6390, HG001, Newman, Becker and CYL1040. Growth at non-permissive temperature (44uC) was followed by several subcultures at 30uC and 37uC to favor double crossing over as previously described [10]. To complement these mutant strains, the *rsaA* gene was amplified using *rsaA-76/rsaA-2194* oligonucleotides and inserted into pLUG274 forming pLUG959. Complementation of mutant strains was performed using pE194, pCN51 or pCN38 carrying *rsaA* gene (Table S2). The CYL1040 was generously provided by CY Lee [24].

### Preparation of RNAs

PCR fragments containing RsaA and its derivatives (RsaA-S1, RsaA-S2) were cloned into pUC18 vector, and then were transcribed *in vitro* using T7 RNA polymerase after plasmid linearization. The WT *mgrA* mRNA (2124 to +258) and its derivative (*mgrA-S1*, *mgrA-S2*) fragments were transcribed from PCR fragments (Table S3). The transcribed RNAs were purified by 8% polyacrylamide-8 M urea gel electrophoresis. After elution in 0.5 M ammonium acetate/1 mM EDTA buffer, the RNAs were precipitated twice with ethanol. The 5' end-labeling of dephosphorylated RNA or DNA oligonucleotides was performed with T4 polynucleotide kinase and [<sup>32</sup>P]ATP. Before use, RNAs were renatured by incubation at 90uC for 2 min in the absence of magnesium and salt, 1 min on ice, followed by an incubation step at 37uC for 15 min in TMN buffer (20 mM Tris-acetate pH 7.5, 10 mM magnesium-acetate, 150 mM Na-acetate).

### Relative quantification of sRNA from *S. aureus* by RT-PCR

RNA extraction was performed from mid (OD<sub>600</sub> = 0.5) or post-exponential phase (OD<sub>600</sub> = 6) grown bacteria using RNeasy Plus mini kit (QIAGEN) as described [53]. One mg of total RNA was transcribed into cDNA using Reverse transcription system (Promega). Subsequently, the cDNA was used as a template for the real-time PCR amplification using a Realplex2 instrument (Eppendorf) and the specific primers shown in Table S1. The

amplification products were detected using SYBR Green. The relative amounts of amplicons for each gene were determined using quantitative PCR relative to an internal standard (*gyrB* encoding Gyrase B subunit) as described [53]. The expression levels of the sRNA were expressed as n-fold differences relative to the calibrator.

#### In vivo $\beta$ -galactosidase assays

Translation fusions were constructed with plasmid pLUG220, a derivative of pTCV-lac, a low-copy-number promoter-less lacZ vector (Table S2). The whole leader region of *mgrA* (nts 2124 to +258) mRNA (from P1 promoter) including 258 nts of the coding sequence, was cloned downstream the *rpoB* promoter in frame with lacZ.  $\beta$ -galactosidase activity was measured four times on duplicate cultures with the Enzyme Assay System (Promega).

#### Northern blots

Electrophoresis of total RNA (10 mg) was done on a 1% agarose gel containing 25 mM guanidine thiocyanate and vacuum transfer to nylon membrane. Hybridizations with specific digoxigenin-labeled RNA probes complementary to *mgrA* mRNA or RsaA and luminescent detection were carried out as described previously [10]. For all experiments, we verified the quantity of 5S rRNA using a digoxigenin-labeled RNA probe.

#### Gel retardation assays

Binding rate constant of RNIII-*mgrA* mRNA complex was measured as described previously [10]. Binding of 59 end-labeled *mgrA* mRNA to a ten-fold excess of unlabeled RsaA or its derivatives was performed at 37°C in TMN buffer. Samples were withdrawn at various time points (0–10 min), added to gel application buffer and loaded onto a native 5% polyacrylamide gel under non denaturing conditions. The gel was run at 4uC and constant voltage (300 V) for 3 h and subsequently dried. Bands corresponding to the RsaA-*mgrA* mRNA complex and free RsaA, respectively, were quantified using the SAFA algorithm. For determination of the dissociation rate constant of RsaA-*mgrA* mRNA complex, end-labeled *mgrA* mRNA was incubated with an increased concentrations of WT RsaA or its variants for 15 min at 37°C in TMN buffer. All experiments were done four times.

#### Proteomics

Growth conditions, preparation of protein extracts, and separation of proteins using two-dimensional (2-D) gel electrophoresis concerted with protein identification by mass spectrometry approach have been described in [20]. Experimental details are given in the supplementary materials (Supplementary Protocol S1).

Relative MgrA protein expression level was defined by a spectral counting strategy (see Supplementary Protocol S1). Triplicate protein extracts from HG001 (WT), DrsaA mutant strain (LUG1630), and the same strain complemented with a plasmid expressing RsaA, were analyzed in separate LC/MS experiments and the MS/MS spectra number were compared for each protein. In parallel to MgrA, we analyzed several ribosomal proteins (L21, L22, L24) as internal controls of constitutively expressed proteins.

#### Western blot analysis

Cell extracts from the post-exponential ( $OD_{600} = 4$ ) phase of growth were prepared from various staphylococcal strains as described in Supplementary Protocol S1. The concentration of total proteins from clear lysates was determined using the Bio-Rad

protein estimation kit and BSA as the standard. Equal amounts of total cellular proteins were separated on 15% polyacrylamide-SDS gels and transferred onto PVDF membranes. Blots were incubated at 20°C with an appropriate dilution (1:1000) of anti-MgrA-specific monoclonal for 1 h (generous gift from A. Cheung, USA), followed by another 1 h incubation with a 1:2500 dilution of goat anti-mouse peroxidase (HRP) conjugate. Prestained protein standards were used for molecular mass estimations, and the gel was stained by Coomassie blue to verify that the quantity of proteins was homogenous in each sample. Each experiment was repeated at least three times with different samples.

#### RNA structure probing using enzymes

RNIII-*mgrA* mRNA formation was carried out at 37°C for 15 min in TMN buffer. Enzymatic hydrolysis was performed either on cold *mgrA* (1 pmole) or 59 end-labeled RsaA (50000 cpm) in 10 ml of TMN, in the presence of 1 mg carrier tRNA at 37°C for 5 min: RNase T1 (0.0025 units), RNase V1 (0.5 units), RNase T2 (0.05 U), RNase III (500 nM). Lead(II) induced cleavages were performed on 59 end-labeled RsaA at 20°C in 20 ml of TMN buffer containing 2 mg of carrier tRNA with 2.5 ml of different concentrations of lead(II)-acetate (12, 40, and 80 mM) for 5 min at 37°C. The reactions were stopped with 5 ml of EDTA 0.1 M for lead(II)-induced cleavages and all other reactions were stopped by phenol extraction followed by RNA precipitation. The end-labeled RsaA fragments were sized on 12% polyacrylamide/8 M urea slab gels. For *mgrA* mRNA, the enzymatic cleavages were detected by primer extension with reverse transcriptase [56].

#### Toeprinting assays

The preparation of the *S. aureus* 30S subunits, the formation of a simplified translational initiation complex with mRNA, and the extension inhibition conditions were described according to [57]. Experimental details are given in the supplementary materials (Supplementary Protocol S1).

#### In vitro transcription-translation assays

The in vitro translation assays were carried out using the full-length WT *mgrA* or *mgrA*-S1 or *mgrA*-S2 mRNAs and the PURESYSTEMII classic kit (Cosmo Bio Co, Japon). All constructs carry an additional sequence corresponding to the Flag peptide, which was inserted at the 59 end of *mgrA*. The reaction was done at 37°C for 1 h in the presence of 25 ml of the commercial solution A and 10 ml of the commercial solution B in the presence of 10 pmoles of mRNA. Experiments were also carried out in the presence of increasing concentrations of WT or mutant RsaA (5 to 30 pmoles). The tubes were placed on ice to stop the reactions, and the proteins were detected by western blot using antibodies against the FLAG tag.

#### Biofilm Crystal Violet Staining and Biofilm Ring Test

The biofilm formation was detected by tissue culture plate method that allowed semi-quantitative measurement of biofilm formation as described previously by Heilmann et al. [58]. All tests were carried out three times and average OD value was calculated for each strain and for negative controls. Experimental details are given in the supplementary materials (Supplementary Protocol S1).

The kinetics of early biofilm was evaluated by Biofilm Ring test (BRT) (BioFilm Control, Saint Beuzire, France) as described [59]. This assay is based on the immobilization of magnetic beads embedded by bacterial aggregates with enough strength to overcome the magnetic attraction forces applied on them. Thus,



in the absence of sessile cells, all the beads are gathered by a magnet centered under the well, giving an easily detectable red spot [59]. Three experiments with two repeats (2 wells per slide) were performed per strain and per incubation time. Experimental details are given in the supplementary materials (Supplementary Protocol S1).

#### Quantification of PIA/PNAG production

The PIA/PNAG polysaccharide of biofilm was quantified by the immuno-slot-blotting method with rabbit anti-PIA/PNAG antiserum as described by Cuçarella et al. [60].

#### Quantification and visualization of the capsular polysaccharides

Capsular polysaccharides type 5 (CP5) or 8 (CP8) were quantified by the immuno-slot-blotting method with rabbit anti-CP5 or anti-CP8 antisera, essentially as described by Luong et al. [24]. CP8 production was determined by indirect immunofluorescence. After culture for 8 h in BHI medium at 37°C, bacteria were spotted in diagnostic slides. The slide was incubated with rabbit anti-CP8 antiserum for 1 h at 37°C in humid chamber. After 2 washes, the slide was incubated with secondary antibody, AlexaFluor 647 goat anti-rabbit immunoglobulin G (Invitrogen, Paisley, UK). Total bacteria were labeled with a 1:1 mixture of vancomycin and vancomycin-Bodipy FL (VBFL) fluorochrome (Invitrogen) at a concentration of 0.8 mg/ml for 15 min. Images were analyzed with confocal microscopy Leica SP5 X (Leica, Solms, Germany) and treated with Leica software: LAS AF Lite (Leica Application Suite 2.6.0).

CP5 or CP8 were quantified by the immuno-slot-blotting method with rabbit anti-CP5 or anti-CP8 antisera, essentially as described by Luong et al. [24]. Experimental details are given in the supplementary material (Supplementary Protocol S1).

#### Model of catheter infection in mice

The catheter infection model was performed essentially as described by [61]. C57Bl/6 female mice (11–14 weeks old) were purchased from Charles River Laboratories (Charles River, Wilmington, USA) and maintained in pathogen-free conditions at the “Plateau de Biologie Expérimentale de la Souris” (PBES, Ecole Normale Supérieure de Lyon, Lyon, France). This study and procedure were approved by the “Comité Rhône-Alpes d’Éthique pour l’Expérimentation Animale” (CECCAPP, Lyon, France).

Mice were anesthetized with 2–3% of isofluran. After depilating and disinfecting the back of the mouse, a sterile catheter was implanted subcutaneously on the back of each mouse. Twenty-five microliters of PBS containing approximately  $4.6 \times 10^4$  bacteria (Becker or Becker-DrSA) were then injected into the catheter. The wound was sewed by a clip. Each day, mice were weighed and the size of the swelling was measured with an electronic caliper. On days 3, 6 and 10, mice were euthanized by cervical dislocation, the catheter and surrounding tissues were carefully removed and transferred separately in sterile tubes for bacterial quantification by plate counting as described. Experimental details are given in the supplementary materials (Supplementary Protocol S1).

#### Bacteremia model

CD-1 female mice (8 to 10 weeks old) were obtained from Charles River Laboratories and maintained in pathogen-free conditions at the PBES. This study and procedure were approved by the CECCAPP. Groups of six mice were inoculated intraperitoneally with  $5.6 \times 10^6$  CFU/mouse in 0.5 ml of PBS. After 1 and 3 h, mice were sacrificed and 4 ml of PBS were intraperitoneally

injected to recover the surviving bacteria. Blood samples were also collected in heparinized tubes from intracardiac puncture. Spleen and kidneys were collected after 24 and 48 h. After addition of 1 ml of NaCl 0.9%, organs were crushed with a homogenizer T18 basic. Bacterial quantification of all samples was performed by plate counting.

#### Statistical analysis

All the data were expressed as mean  $\pm$  SD. The statistical analyses were performed using GraphPad Prism 6 software. A one-way analysis of variance was realized to test the results obtained in mature biofilm quantification. Other data were analyzed using Mann-Whitney test. A value of  $p > 0.05$ ,  $> 0.005$  or  $> 0.0005$  was considered to be statistically significant (\*), highly significant (\*\*), or extremely significant (\*\*\*), respectively.

#### Supporting Information

Figure S1 Effect of RsaA on the *S. aureus* proteome. (A) 2-D fluorescence difference gel electrophoresis (DiGE) performed on HG001 and DrSA-HG001 strains. Total protein extracts were prepared from cultures performed at 37°C stopped at the stationary phase of growth, and labeled with different cyanines. Yellow spot, for unchanged proteins; green spot, protein synthesis activated in strain expressing RsaA; red spot, protein synthesis repressed in strain expressing RsaA. (B) Interpretation of the LC-MS-MS data. A Spectral Counting (SpC) strategy was carried out using the Mascot identification results and Proteinscape 3.1 package. The total number of proteins and spectra are given for each assay. To normalize data, a correction factor was applied for each condition according to the average spectra number for all samples. Three independent experiments were done. In the second table, the normalized spectral counts SpC for MgrA and for the ribosomal protein L21 (RL21) are given and the data have been normalized to the wild-type (WT) strain. STD: A statistical analysis was performed to provide the mean and the standard deviations, which are represented on the histogram. The data were done on crude extracts prepared from late-exponential growth of the WT (HG001), the DrSA mutant strain (DrSA), and DrSA mutant strain complemented with a plasmid expressing RsaA under its own promoter (DrSA/pl. rSA). (EPS)

Figure S2 Secondary structures of RsaA and of mgrA mRNA. (A) RsaA secondary structure. The conserved nucleotides are shown in red. The two regions of interaction (sites 1 and 2) are boxed. (B) Structure of the leader region of mgrA mRNA expressed from the P2 promoter. The Shine and Dalgarno sequence (SD) and the AUG codon are in red. (C) RNase III cleavage assays on mgrA mRNA. Lanes 1–2: incubation controls performed on mgrA mRNA in the absence (lane 1) or in the presence of 250 nM RsaA (lane 2); lanes 3–5: RNase III (0.65 mM) cleavage assays on mgrA in the presence of increasing concentrations of RsaA (25 nM, lane 4; 50 nM, lane 5; 250 nM, lane 6); lane 7: RNase III (0.65 mM) cleavage assay performed on mgrA bound to RsaA (250 nM) in a buffer containing  $\text{Ca}^{2+}$  instead of  $\text{Mg}^{2+}$ . Lanes 8–12: same legend as for lanes 3–7, respectively except that the concentration of RNase III was 0.35 mM. Lanes U, C, G, A: sequencing ladders. (EPS)

Figure S3 Effect on RsaA on ica-dependent biofilm formation. Dot-blot analysis of PIA/PNAG polysaccharide accumulation in wild-type strains and their corresponding DrSA mutant strains and the same strains complemented with a plasmid

expressing RsaA. Growth was performed in BHI medium containing 0.25% glucose for 16 h. Standardized bacterial suspensions were treated with proteinase K, and 1/5 serial dilutions were spotted on nitrocellulose membrane. PIA/PNAG production was detected with an anti-PIA/PNAG antibody. (EPS)

**Figure S4 RsaA affects opsonophagocytosis.** Mortality of bacteria after opsonophagocytosis assay. Each *S. aureus* strain was mixed with the same quantity of human serum and PMNs for 30 min at 37°C. Sample dilutions were made in sterile deionized water, and bacterial killing was estimated by plating the diluted samples in duplicate on GP. The cell killing was defined as the reduction in CFU/ml after 30 min compared with that at time zero. The graphic represents the % of mortality for each strain of 3 to 5 individual experiments. Three different *S. aureus* backgrounds (HG001, Newman, Becker) were analyzed, as well as their corresponding *DrsA* mutant strains and the same mutant strains complemented with a plasmid expressing RsaA. (TIFF)

**Figure S5 The impact of RsaA on virulence in in vivo bacteremia model.** An in vivo bacteremia model has been used to mimic an acute infection with Becker and Newman background. (A, C, D)  $5 \times 10^6$  bacteria in 500 µl were injected into peritoneal cavity. After 1 h (C) and 3 h (A, D), viable bacteria into peritoneal cavity were removed, enumerated by serial dilutions and plate counting on agar plates. (B, E, F) Blood samples were also collected in heparinated from intra-cardiac puncture and bacteria were enumerated by serial dilutions and plate counting on agar plates. (G, H) After 24 and 48 h, spleen was collected in the two strain backgrounds (Becker and Newman) and their corresponding *DrsA* mutant strains, and bacteria were enumerated. The results were expressed in number of CFU/ml. Each point designates the number of bacteria for each mouse and the black line represents the mean of each tested condition. Statistical differences \* ( $p > 0.05$ ), \*\* ( $p > 0.005$ ) and \*\*\* ( $p > 0.0005$ ) between each groups were obtained by using Mann Whitney test. (TIFF)

**Figure S6 Expression of RsaA in reference strains and clinical isolates.** The expression of RsaA was quantified and

normalized to the level of *gyrB* mRNA expression after in vitro culture to the mid- or late-exponential phase for 16 clinical isolates (A) and 6 reference strains (B). (EPS)

Protocol S1 Supplementary Material and Methods. (DOC)

Table S1 In silico prediction of base pairings between RsaA and mRNAs. (DOCX)

Table S2 Strains and plasmids. (DOCX)

Table S3 Primers used in this study. (DOCX)

## Acknowledgments

We thank Eric Westhof, Frédéric Laurent, Gerard Lina and Thomas Henry for helpful discussions and support. We are thankful to Redmond Smyth and Elise Rattigan for critical reading of the manuscript, and to Efthymia Lioliou and Stefano Marzi for discussions. We are grateful to Chia Lee for providing strain Becker and Becker *DmgrA*, and rabbit anti-CP5 and anti-CP8 antisera, to A. Cheung for providing monoclonal antibodies against *MgrA*, to Inigo Lasa for providing rabbit anti-PIA/PNAG antiserum, to Franck DeLeo and Binh An Diep for providing strain LAC and SF8300 respectively, to Efthymia Lioliou for providing purified RNase III, to Pierre Fechter for advices on the DiGE approach, to Johana Chicher and Laurianne Kuhn for the mass spectrometry analysis, to Anne-Catherine Helfer Le Foll and Florence Couzon for technical help, to Sabine Bourdeau-Patot and Marie Teixeira for help regarding the animal model and animal facility, for Michele Bes for help regarding clinical isolates, and to Anabelle Bouchardon and Chantal Thenevon (CECIL) for technical help concerning microscopy imaging and FACS.

## Author Contributions

Conceived and designed the experiments: CR CL AT IC YB PH TG SB PR FV. Performed the experiments: CR CL AT YB TG SB PH. Analyzed the data: CR CL AT IC TG SB PH PR FV. Wrote the paper: CR CL TG SB PR FV.

## References

- Chambers HF, Deleo FR (2009) Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nat Rev Microbiol* 7: 629–641.
- Priest NK, Rudkin JK, Feil EJ, van den Elsen JM, Cheung A, et al. (2012) From genotype to phenotype: can systems biology be used to predict *Staphylococcus aureus* virulence? *Nat Rev Microbiol* 10: 791–797.
- Novick RP, Geisinger E (2008) Quorum sensing in staphylococci. *Annu Rev Genet* 42: 541–564.
- Felden B, Vandenesch F, Bouloc P, Romby P (2011) The *Staphylococcus aureus* RNome and its commitment to virulence. *PLoS Pathog* 7: e1002006.
- Romilly C, Caldelari I, Parmentier D, Lioliou E, Romby P, et al. (2012) Current knowledge on regulatory RNAs and their machineries in *Staphylococcus aureus*. *RNA Biol* 9: 402–413.
- Novick RP, Ross HF, Projan SJ, Kornblum J, Kreiswirth B, et al. (1993) Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J* 12: 3967–3975.
- Morfeldt E, Taylor D, von Gabain A, Arvidson S (1995) Activation of alpha-toxin translation in *Staphylococcus aureus* by the trans-encoded antisense RNA, RNAlII. *EMBO J* 14: 4569–4577.
- Geisinger E, Adhikari RP, Jin R, Ross HF, Novick RP (2006) Inhibition of rot translation by RNAlII, a key feature of *agr* function. *Mol Microbiol* 61: 1038–1048.
- Huntzinger E, Boisset S, Saveanu C, Benito Y, Geissmann T, et al. (2005) *Staphylococcus aureus* RNAlII and the endoribonuclease III coordinately regulate *spa* gene expression. *EMBO J* 24: 824–835.
- Boisset S, Geissmann T, Huntzinger E, Fechter P, Bendridi N, et al. (2007) *Staphylococcus aureus* RNAlII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev* 21: 1353–1366.
- Chevalier C, Boisset S, Romilly C, Masquida B, Fechter P, et al. (2010) *Staphylococcus aureus* RNAlII binds to two distant regions of *coa* mRNA to arrest translation and promote mRNA degradation. *PLoS Pathog* 6: e1000809.
- Pichon C, Felden B (2005) Small RNA genes expressed from *Staphylococcus aureus* genomic and pathogenicity islands with specific expression among pathogenic strains. *Proc Natl Acad Sci USA* 102: 14249–14254.
- Roberts C, Anderson KL, Murphy E, Projan SJ, Mounts W, et al. (2006) Characterizing the effect of the *Staphylococcus aureus* virulence factor regulator, SarA, on log-phase mRNA half-lives. *J Bacteriol* 188: 2593–2603.
- Beaume M, Hernandez D, Farinelli L, Deluen C, Linder P, et al. (2010) Cartography of methicillin-resistant *S. aureus* transcripts: detection, orientation and temporal expression during growth phase and stress conditions. *PLoS One* 5: e10725.
- Nielsen JS, Larsen MH, Lillebaek EM, Bergholm TM, Christiansen MH, et al. (2011) A small RNA controls expression of the chitinase ChiA in *Listeria monocytogenes*. *PLoS One* 6: e19019.
- Sayed N, Jousset A, Felden B (2012) A cis-antisense RNA acts in trans in *Staphylococcus aureus* to control translation of a human cytolysin peptide. *Nat Struct Mol Biol* 19: 105–112.
- Sayed N, Nonin-Lecomte S, Rety S, Felden B (2012) Functional and structural insights of a *Staphylococcus aureus* apoptotic-like membrane peptide from a toxin-antitoxin module. *J Biol Chem* 287: 43454–43463.
- Howden BP, Beaume M, Harrison PF, Hernandez D, Schrenzel J, et al. (2013) Analysis of the small RNA transcriptional response in multidrug-resistant *Staphylococcus aureus* after antimicrobial exposure. *Antimicrob Agents Chemother* 57: 3864–3874.
- Abu-Qatouseh LF, Chinni SV, Seggewiss J, Proctor RA, Brosius J, et al. (2010) Identification of differentially expressed small non-protein-coding RNAs in

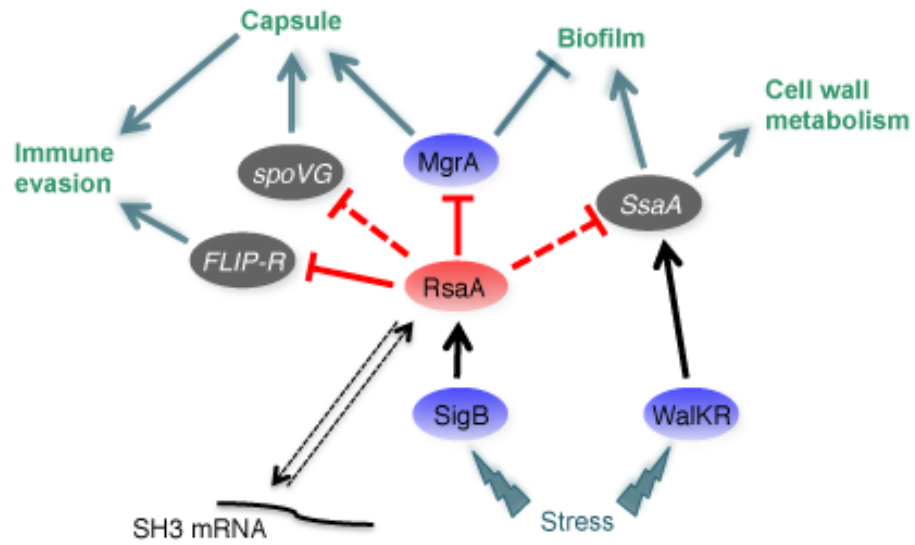
- Staphylococcus aureus* displaying both the normal and the small-colony variant phenotype. *J Mol Med (Berl)* 88: 565–575.
20. Geissmann T, Marzi S, Romby P (2009) The role of mRNA structure in translational control in bacteria. *RNA Biol* 6: 153–160.
  21. Bohn C, Rigoulay C, Chabelskaya S, Sharma CM, Marchais A, et al. (2010) Experimental discovery of small RNAs in *Staphylococcus aureus* reveals a riboregulator of central metabolism. *Nucleic Acids Res* 38: 6620–6636.
  22. Lasa I, Toledo-Arana A, Dobin A, Villanueva M, de los Mozos IR, et al. (2011) Genome-wide antisense transcription drives mRNA processing in bacteria. *Proc Natl Acad Sci USA* 108: 20172–20177.
  23. Lioliou E, Sharma CM, Caldeleri I, Helfer AC, Fechter P, et al. (2012) Global regulatory functions of the *Staphylococcus aureus* endoribonuclease III in gene expression. *PLoS Genet* 8: e1002782.
  24. Luong TT, Newell SW, Lee CY (2003) Mgr, a novel global regulator in *Staphylococcus aureus*. *J Bacteriol* 185: 3703–3710.
  25. Ingavale S, van Wamel W, Luong TT, Lee CY, Cheung AL (2005) Rat/MgrA, a regulator of autolysis, is a regulator of virulence genes in *Staphylococcus aureus*. *Infect Immun* 73: 1423–1431.
  26. Luong TT, Dunman PM, Murphy E, Projan SJ, Lee CY (2006) Transcription Profiling of the mgrA Regulon in *Staphylococcus aureus*. *J Bacteriol* 188: 1899–1910.
  27. Ingavale SS, Van Wamel W, Cheung AL (2003) Characterization of RAT, an autolysis regulator in *Staphylococcus aureus*. *Mol Microbiol* 48: 1451–1466.
  28. Gupta RK, Alba J, Xiong YQ, Bayer AS, Lee CY (2013) MgrA activates expression of capsule genes, but not the alpha-toxin gene in experimental *Staphylococcus aureus* endocarditis. *J Infect Dis*
  29. Meier S, Goerke C, Wolz C, Seidl K, Homerova D, et al. (2007) sigmaB and the sigmaB-dependent arlRS and yabJ-spoVG loci affect capsule formation in *Staphylococcus aureus*. *Infect Immun* 75: 4562–4571.
  30. Busch A, Richter AS, Backofen R (2008) IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* 24: 2849–2856.
  31. Tjaden B (2008) TargetRNA: a tool for predicting targets of small RNA action in bacteria. *Nucleic Acids Res* 36: W109–13.
  32. Trotonda MP, Tamber S, Memmi G, Cheung AL (2008) MgrA represses biofilm formation in *Staphylococcus aureus*. *Infect Immun* 76: 5645–5654.
  33. Luong TT, Lee CY (2006) The arl locus positively regulates *Staphylococcus aureus* type 5 capsule via an mgrA-dependent pathway. *Microbiology* 152: 3123–3131.
  34. Vuong C, Voyich JM, Fischer ER, Braughton KR, Whitney AR, et al. (2004) Polysaccharide intercellular adhesin (PIA) protects *Staphylococcus epidermidis* against major components of the human innate immune system. *Cell Microbiol* 6: 269–275.
  35. Thakker M, Park JS, Carey V, Lee JC (1998) *Staphylococcus aureus* serotype 5 capsular polysaccharide is antiphagocytic and enhances bacterial virulence in a murine bacteremia model. *Infect Immun* 66: 5183–5189.
  36. Xu S, Arbeit RD, Lee JC (1992) Phagocytic killing of encapsulated and microencapsulated *Staphylococcus aureus* by human polymorphonuclear leukocytes. *Infect Immun* 60: 1358–1362.
  37. Marincola G, Schafer T, Behler J, Bernhardt J, Ohlsen K, et al. (2012) RNase Y of *Staphylococcus aureus* and its role in the activation of virulence genes. *Mol Microbiol* 85: 817–832.
  38. Ballal A, Manna AC (2009) Expression of the sarA family of genes in different strains of *Staphylococcus aureus*. *Microbiology* 155: 2342–2352.
  39. Wagner EG, Altuvia S, Romby P (2002) Antisense RNAs in bacteria and their genetic elements. *Adv Genet* 46: 361–398.
  40. Lioliou E, Sharma CM, Altuvia Y, Caldeleri I, Romilly C, et al. (2013) In vivo mapping of RNA-RNA interactions in *Staphylococcus aureus* using the endoribonuclease III. *Methods* 63: 135–143.
  41. Bischoff M, Dunman P, Kormanec J, Macapagal D, Murphy E, et al. (2004) Microarray-based analysis of the *Staphylococcus aureus* sigmaB regulon. *J Bacteriol* 186: 4085–4099.
  42. Rashid MH, Rumbaugh K, Passador L, Davies DG, Hamood AN, et al. (2000) Polyphosphate kinase is essential for biofilm development, quorum sensing, and virulence of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci USA* 97: 9636–9641.
  43. Valle J, Toledo-Arana A, Berasain C, Ghigo JM, Amorena B, et al. (2003) SarA and not sigmaB is essential for biofilm development by *Staphylococcus aureus*. *Mol Microbiol* 48: 1075–1087.
  44. Sun F, Ding Y, Ji Q, Liang Z, Deng X, et al. (2012) Protein cysteine phosphorylation of SarA/MgrA family transcriptional regulators mediates bacterial virulence and antibiotic resistance. *Proc Natl Acad Sci USA* 109: 15461–15466.
  45. Moscoso M, Garcia E, Lopez R (2006) Biofilm formation by *Streptococcus pneumoniae*: role of choline, extracellular DNA, and capsular polysaccharide in microbial accretion. *J Bacteriol* 188: 7785–7795.
  46. Flahaut S, Vinogradov E, Kelley KA, Brennan S, Hiramatsu K, et al. (2008) Structural and biological characterization of a capsular polysaccharide produced by *Staphylococcus haemolyticus*. *J Bacteriol* 190: 1649–1657.
  47. Novick RP, Christie GE, Penades JR (2010) The phage-related chromosomal islands of Gram-positive bacteria. *Nat Rev Microbiol* 8: 541–551.
  48. Maurelli AT (2007) Black holes, antivirulence genes, and gene inactivation in the evolution of bacterial pathogens. *FEMS Microbiol Lett* 267: 1–8.
  49. Bliven KA, Maurelli AT (2012) Antivirulence genes: insights into pathogen evolution through gene loss. *Infect Immun* 80: 4061–4070.
  50. Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, et al. (2009) The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* 459: 950–956.
  51. Baek CH, Wang S, Roland KL, Curtiss Jr (2009) Leucine-responsive regulatory protein (Lrp) acts as a virulence repressor in *Salmonella enterica* serovar Typhimurium. *J Bacteriol* 191: 1278–1292.
  52. Lee EJ, Groisman EA (2010) An antisense RNA that governs the expression kinetics of a multifunctional virulence gene. *Mol Microbiol* 76: 1020–1033.
  53. Song J, Lays C, Vandenesch F, Benito Y, Bes M, et al. (2012) The expression of small regulatory RNAs in clinical samples reflects the different life styles of *Staphylococcus aureus* in colonization vs. infection. *PLoS One* 7: e37294.
  54. Kreiswirth BN, Lofdahl S, Betley MJ, O'Reilly M, Schlievert PM, et al. (1983) The toxic shock syndrome exotoxin structural gene is not detectably transmitted by a prophage. *Nature* 305: 709–712.
  55. Arnaud M, Chastanet A, Debarbouille M (2004) New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. *Appl Environ Microbiol* 70: 6887–6891.
  56. Chevalier C, Geissmann T, Helfer AC, Romby P (2009) Probing mRNA structure and sRNA-mRNA interactions in bacteria using enzymes and lead(II). *Methods Mol Biol* 540: 215–232.
  57. Fechter P, Chevalier C, Yusupova G, Yusupov M, Romby P, et al. (2009) Ribosomal initiation complexes probed by toeprinting and effect of trans-acting translational regulators in bacteria. *Methods Mol Biol* 540: 247–263.
  58. Heilmann C, Schweitzer O, Gerke C, Vanittanakom N, Mack D, et al. (1996) Molecular basis of intercellular adhesion in the biofilm-forming *Staphylococcus epidermidis*. *Mol Microbiol* 20: 1083–1091.
  59. Chavant P, Gaillard-Martinie B, Talon R, Hebraud M, Bernardi T (2007) A new device for rapid evaluation of biofilm formation potential by bacteria. *J Microbiol Methods* 68: 605–612.
  60. Cucarella C, Solano C, Valle J, Amorena B, Lasa I, et al. (2001) Bap, a *Staphylococcus aureus* surface protein involved in biofilm formation. *J Bacteriol* 183: 2888–2896.
  61. Kristian SA, Birkenstock TA, Sauder U, Mack D, Gotz F, et al. (2008) Biofilm formation induces C3a release and protects *Staphylococcus aureus* from IgG and complement deposition and from neutrophil-dependent killing. *J Infect Dis* 197: 1028–1035.

## Partie 2 – Application de l’approche MAPS à RsaA : à la recherche des autres ARNm cibles

L'expression du virulon de *Staphylococcus aureus* est régulée par des connexions complexes entre les régulateurs de transcription et post-transcriptionnel, y compris les protéines et les petits ARN non codants (ARNnc). La plupart des ARNnc, dont les fonctions ont été résolues, régule l'expression des gènes au niveau post-transcriptionnel. Cependant, la caractérisation des cibles directes dans les bactéries Gram-positives demeure un défi à relever.

Dans cet article, j'ai appliqué et adapté une méthode de purification des complexes régulateurs *in vivo* par une approche de chromatographie d'affinité couplée au séquençage de l'ARN (MAPS) pour déterminer la fonction de l'ARNnc RsaA de *S. aureus*. Cette méthode permet la purification des complexes régulateurs à partir de cellules exprimant tout type d'ARN sans tenir compte du mécanisme de régulation. Dans cette étude, j'ai caractérisé de nouveaux ARNm significativement enrichis avec RsaA. Nous avons montré que, outre MgrA, plusieurs ARNm codent pour une famille d'enzymes impliquées dans le métabolisme du peptidoglycane et la protéine sécrétée FLIPR qui contrecarre la réponse immunitaire de l'hôte en ciblant un récepteur des neutrophiles et macrophages.

En utilisant une combinaison d'approches expérimentales, j'ai validé ces ARNm comme cibles directes de RsaA. Le motif riche en C de RsaA semble être l'un des déterminants principal qui reconnaît le site de liaison du ribosome de plusieurs des ARNm cibles. L'analyse protéomique différentielle quantitative a révélé que RsaA active plusieurs protéines de surface de manière indirecte. Cette étude est en accord avec le fait que RsaA favorise les infections chroniques, en induisant la formation des biofilms et en inhibant la formation de la capsule et le phénomène de « clumping » (voir figure). J'ai ainsi montré que l'approche MAPS peut être appliqué aux bactéries Gram-positives pour l'identification du targetome des ARNnc.



**Figure 16 :** Représentation schématique d'une partie du réseau de régulation de RsaA - L'expression de RsaA est dépendante du facteur SigmaB impliqué dans la réponse au stress. Sa cible principale est MgrA et j'ai démontré que les protéines FLIP-R et SsaA sont également régulées par l'ARN. Dans le cas de FLIP-R (codé par le gène *fliR*) la région d'interaction a été caractérisée. La régulation de ces protéines met RsaA en position de levier entre la formation de capsule et la formation du biofilm chez *S. aureus*.

Cette étude est présentée sous la forme d'un article qui sera soumis prochainement.

**The RNA targetome of *Staphylococcus aureus* non coding RNA RsaA: impact on cell surface properties and defence mechanisms**

Tomasini A., Moreau K., Geissmann T., Vandenesch F., Romby P.\*, Marzi S., and Caldelari I.\*

<sup>1</sup>Architecture et Réactivité de l'ARN, Université de Strasbourg, CNRS, IBMC, 15 Rue René Descartes 67084 Strasbourg, France

<sup>2</sup>Univ Lyon, Université Claude Bernard Lyon1, Ecole Normale Supérieure de Lyon, INSERM U1111, CNRS UMR 5308, CIRI, F-69008, Lyon, France

\*To whom correspondence should be addressed. Tel: 33(0) 388417068; Fax: 33(0) 3886022188; Email: [i.caldelari@ibmc-cnrs.unistra.fr](mailto:i.caldelari@ibmc-cnrs.unistra.fr); [p.romby@ibmc-cnrs.unistra.fr](mailto:p.romby@ibmc-cnrs.unistra.fr)

## INTRODUCTION

The importance of post-transcriptional regulation in bacteria is now widely recognized for several adaptive cellular processes such as iron regulation, biofilm formation, quorum sensing and virulence (e.g., Caldelari *et al*, 2013). These regulations are essentially achieved through a combination of proteins (RNA binding proteins, ribonucleases) and small non coding RNAs. Pioneered studies performed in *Escherichia coli* and closely related bacteria have revealed that most of the sRNAs regulate gene expression through direct mRNA binding (e.g., Storz *et al.*, 2011; Barquist and Vogel, 2015). Because many sRNAs regulate transcriptional/post-transcriptional regulatory proteins, they induce downstream effects rendering challenging the identification of their primary targets. In addition, target predictions are difficult since no general rules have emerged and the basepairing interactions can be rather short and non contiguous but they often involved conserved and unpaired regions of the sRNAs (reviewed in Wagner and Romby, 2015, Updegrove *et al.*, 2015). Hence, several approaches have been developed over the years to identify the primary sRNA targets in order to characterize their regulatory networks and their functional impact on cell physiology (Vogel and Wagner, 2007; Barquist and Vogel, 2015; Lalaouna and Massé, 2015). In *E. coli* and closely related bacteria, analysis of the rapid changes in mRNA levels after pulse induction of the analyzed sRNA has been for a long time one of the most successful methods (Vogel and Wagner, 2007). Besides, other approaches were used to assess the sRNAs that do not alter mRNA levels such as differential proteomic approaches (Udekwu *et al.*, 2005), ribosome profiling (Wang *et al.*, 2015), or purification of regulatory complexes using protein factors as baits (Chao *et al.*, 2012; Tree *et al.*, 2014). More recently, affinity purification of *in vivo* regulatory complexes coupled with high throughput RNA sequencing methodology (Lalouna and Massé, 2015; Lalaouna *et al.*, 2015; Carrier *et al.*, 2016) has proven to get access to the whole targetome of any type of RNAs in *E. coli*. In this work, we have adapted the MAPS strategy to *S. aureus* sRNAs.

*S. aureus* is a Gram-positive opportunistic pathogen that is well adapted to diverse hosts and ecological niches. The bacterium is present in a third of the population on skin and in nostrils as commensal bacteria, but is also responsible for a wide range of infections from minor skin abscesses to life-threatening diseases such as septicaemia or toxic shock syndrome. Mainly recognized as extracellular pathogen, it

can nevertheless invade professional and non-professional phagocytes leading to chronic infections that became difficult to treat (Surmann et al., 2014). *S. aureus* produces a plethora of virulence factors including secreted toxins, exopolysaccharides, the major component of the biofilm matrix, capsule and several surface proteins contributing to host adhesion. Transcriptional factors, two components systems, and RNA regulators fine-tune the virulence expression programs (Tomasini et al., 2014). The discovery of RNAIII, the major RNA regulator of virulence in *S. aureus*, by Novick *et al* (1993) has opened new ways of research. Today, more than 100 sRNAs have been described and the functions of many of them are still unknown (reviewed in Felden et al., 2011). Several studies revealed that RNAIII and other sRNAs contain conserved UCCC sequences that play the roles of seed sequences to target the ribosome-binding sites (RBS) of target mRNAs (*i.e.* Huntzinger et al., 2005; Geisinger et al., 2006; Boisset et al., 2007). It appears that this mode of pairings is conserved among several Gram-positive bacteria (*e.g.*, Durand et al., 2016; Sievers et al., 2014). This interaction hinders ribosome binding to inhibit translation initiation but in contrast to Gram-negative bacteria, does not necessarily lead to rapid depletion of the target mRNAs. This was the case of the Sigma B-dependent RsaA sRNA, which was first identified in *S. aureus* genome using bioinformatics and expression studies (Geissmann *et al*, 2009). By comparing the cytosolic protein contents of the wild type and the mutant  $\Delta$ rsaA strains by a 2D-gel approach, the synthesis of proteins belonging to the regulon of the global transcriptional regulator MgrA was altered, and the effect of RsaA was opposite to that of MgrA (Romilly et al., 2014). In agreement with these observations, RsaA was shown to repress the translation of *mgrA* mRNA by forming imperfect duplexes between its C-rich motif and RBS of *mgrA* mRNA further stabilized by a loop-loop interaction within the coding region of *mgrA*. Consequently, RsaA causes enhanced production of biofilm, decreased synthesis of capsule formation, and decreased protection of *S. aureus* against opsonophagocytic killing by polymorphonuclear leukocytes (Romilly et al., 2014).

In order to extend the repertoire of RsaA targets, we adapted for the first time the MAPS technology to *S. aureus* coupled to a complete differential proteomic analysis of the  $\Delta$ rsaA strain including cytosolic and secreted proteins. Concerning MAPS, RsaA was tagged at its 5' end with two hairpin motifs recognized by the MS2 coat



protein and expressed *in vivo*. After cell lysis, affinity chromatography purification followed by RNA-seq, we have confirmed that *mgrA* mRNA was the main mRNA target of RsaA and have revealed other mRNAs encoding cell wall hydrolases and FLIPr (also called FIR), a secreted immunomodulatory molecule that interferes with effective opsonization by the complement and/or IgG antibodies (Prat et al., 2006). The MAPS approach appears to be one powerful methodology to extend RsaA regulon of *S. aureus*.

## MATERIAL AND METHODS

### Strains, plasmids and growth conditions

*S. aureus* strains and plasmids used in this study are listed in Table S1. *E. coli* strain DC10B (Monk *et al.*, 2012) was used as a host strain for plasmid construction. Plasmids extracted from *E. coli* DC10B can be used directly for *S. aureus* electroporation. *E. coli* strain was cultivated in Luria-Bertani (LB) medium (1% peptone, 0,5% yeast extract, 1% NaCl) supplemented with ampicillin (100 µg/ml) when necessary. LB-agar plates (with or without ampicillin) were also used for growth on solid medium. *S. aureus* strains were grown in Brain-Heart Infusion (BHI) medium (Sigma-Aldrich) supplemented with erythromycin (10 µg/ml) when necessary. Blood-agar (VWR Chemicals) and BHI plates (with or without erythromycin) were also used for growth on solid medium. Plasmids were prepared from transformed *E. coli* pellets using the Nucleospin Plasmid kit (Macherey-Nagel). Transformation of both *E. coli* and *S. aureus* strains was performed by electroporation (Bio-Rad Gene Pulser).

All plasmids were prepared using pCN51 as template vector (Charpentier *et al.* 2004). Synthesis of PCR products was performed using Phusion Polymerase (Thermoscientific). To remove the cadmium inducible promoter, pCN51 was first digested by *SphI/PstI*. The P3 promoter region was then amplified by PCR and cloned into pCN51 following *SphI/PstI* digestion, forming pCN51/P3. PCR products containing either the sole MS2 tag, or MS2 tag fused to the 5' end of RsaA were cloned into pCN51/P3 following digestion of products and of the plasmid by *PstI/BamHI*. Construction of both pCN51/P3 and its derivatives (Table S1) was conducted in DC10B. Plasmids from positive clones were sequenced (GATC Biotech) before being electroporated into *S. aureus* strains.

### **Preparation of crude bacterial extracts**

For crude bacterial extract preparation, *S. aureus* cells were grown in 50 ml BHI medium to an OD<sub>600nm</sub> of 5 (approx. 6h of culture at 37°C), immediately chilled on ice, and then pelleted by centrifugation (3750 rpm, 15 min, 4°C). Supernatants were carefully removed and the pellets were resuspended in 4 mL ice-cold Buffer A (20 mM Tris-HCl pH 8,0, 150 mM KCl, 1 mM MgCl<sub>2</sub>, 1 mM DTT). Cells were then transferred into Lysis Matrix B Tubes (MP Biomedicals) and lysis was performed with FastPrep apparatus (MP Biomedicals). The tubes were then centrifuged 15 min (13000 rpm, 4°C). The supernatants (around 4 ml) were carefully withdrawn and placed in new ice-cold tubes that were maintained at 4°C until subsequent use for the affinity chromatography.

### **Affinity purification of regulatory complexes containing MS2 tagged sRNA**

The MS2 coat protein fused to maltose binding protein (MS2-MBP) was purified as previously described (Said *et al.* 2009). All steps for affinity chromatography purification were performed at 4°C. For each sample, 300 µl of amylose resin (New England BioLabs) per column was used. The columns containing the amylose resin were washed three times with 10 ml Buffer A, and 3600 pmoles of MS2-MBP (diluted in 6 ml of Buffer A) were then immobilized in each column. The crude bacterial extracts (4 ml) were directly loaded on each column, followed by three washes with 10 ml Buffer A. Finally, RNAs and proteins were eluted from the column with 1 ml of Buffer E (20 mM Tris-HCl pH 8, 150 mM KCl, 1 mM MgCl<sub>2</sub>, 1 mM DTT, 0,1% Triton X-100, 12 mM maltose). Eluted RNAs were extracted with phenol-chloroform-isoamylalcohol (25 :24 :1 (v/v), Roth) and precipitated with 3 vol of cold absolute ethanol in the presence of sodium acetate 0,3 M. The RNA samples were either used for Northern blot or treated with DNase I prior to RNA-seq analysis. Isolation of tagged sRNAs and the co-purified RNAs was performed in duplicates.

### **Northern blot**

Total RNAs were prepared from different volumes of *S. aureus* cultures taken at 2, 4 and 6 hours of culture. After centrifugation, bacterial pellets were dissolved in RNA Pro Solution (MP Biomedicals). Lysis was performed with FastPrep and the RNA

purification followed strictly the procedure described for the FastRNA Pro Blue Kit (MP Biomedicals).

Electrophoresis of either total RNA (10 µg) or MS2-eluted RNA (1 µg) was performed on 1% agarose gel containing 25 mM guanidium thiocyanate. After migration, RNAs were vacuum transferred on nitrocellulose membrane. Hybridization with specific digoxigenin (DIG)-labelled probes complementary to RsaA or the MS2-tag sequences followed by luminescent detection were carried out as described previously (Boisset *et al.* 2007).

### **RNA-seq analysis**

For both total RNA extracts and MS2-eluted RNAs, DNase I (0,1 U/µl) treatment was performed 1h at 37°C. The reactions mixtures were then purified by phenol/chloroform/isoamyl alcohol and subsequent ethanol precipitation as described above. RNA pellets were re-suspended in sterile milliQ water. RNA quality and quantity assessments were performed on Agilent Nano Chip on the bioanalyzer 2100. The RNAs for total transcriptomics were then treated to deplete abundant rRNAs, and the cDNA libraries were done using the Random Hexaprimer approach. The libraries have been sequenced by Fasteris (Switzerland). The standard protocol used is the “TruSeq Stranded mRNA” which is based on the TruSeq Illumina kit and which preserves the information on the orientation of the transcripts and produces reads of 125 nts which map on the complementary strand. The libraries have been sequenced using Illumina Hi-Seq 2500 with High-Output (HO) mode using a V4 chemistry sequencing kit (Illumina). Each RNA-seq was performed in duplicates. The reads have been then analysed using Galaxy project (Afgan *et al.*, 2016). They have been processed to remove adapter sequences and bad quality reads, converted to the FASTQ format with FASTQ Groomer (Blankenberg *et al.*, 2010), aligned on the HG001 genome (Caldelari *et al.*, 2016; GenBank accession number XXX) using BOWTIE2 (Langmead *et al.*, 2009). Finally, the number of reads mapping to each annotated feature has been counted with HTSeq (Anders *et al.*, 2015) using the interception non-empty protocol. To estimate the enrichment values for the MAPS experiment or the differential expression analysis for the transcriptomic experiment, we used DEseq2 (Varet *et al.*, 2016). The statistical analysis process includes data normalization, graphical exploration of raw and normalized data, test for differential

expression for each feature between the conditions, raw p-value adjustment and export of lists of features having a significant differential expression (threshold p-value=0.05; fold change threshold=2) between the conditions.

### **Preparation of RNAs for *in vitro* experiments**

Transcription of RsaA and RsaA-mutC was performed using linearized pUC18 vectors (Romilly *et al.* 2014). PCR fragments containing the 5'UTR and the 120 first nucleotides of selected mRNA targets were directly used as templates for *in vitro* transcription using T7 RNA polymerase (see Table S2). The RNAs were then purified using a 6% polyacrylamide-8 M urea gel electrophoresis. After elution with 0,5 M ammonium acetate pH 6.5 containing 1 mM EDTA, the RNAs were precipitated in cold absolute ethanol, washed with 85% ethanol and vacuum-dried. The labelling of the 5' end of dephosphorylated RNAs (RsaA/RsaA mutC) and DNA oligonucleotides were performed with T4 polynucleotide kinase (Fermentas ?) and [ $\gamma^{32}\text{P}$ ] ATP.

Before use, cold or labelled RNAs were renatured by incubation at 90°C for 1 min in 20 mM Tris-HCl pH 7,5, cooled 1 min on ice, and incubated 10 min at 20°C in ToeP+ buffer (20 mM Tris-HCl pH 7,5, 10 mM MgCl<sub>2</sub>, 60 mM KCl, 1 mM DTT).

### ***In vitro* translation assays**

The *in vitro* translation assays were performed using the full-length *mgrA* mRNA and the PURESYSTEMII Classic Kit (Cosmo Bio Co, Japan). The construct carries an additional sequence corresponding to the FLAG peptide, which was inserted at the 3' end of *mgrA* (Romilly *et al.* 2014). The reactions were done at 37°C for 1h in the presence of 25  $\mu\text{l}$  of the commercial solution A and 10  $\mu\text{l}$  of the commercial solution B in the presence of 10 pmoles of *mgrA* mRNA. Experiments were also carried out in the presence of increasing concentrations of RsaA or of MS2-RsaA transcript (5 to 30 pmoles). Reactions were stopped by adding an equal volume of Laemmli buffer (Bio-Rad). The proteins were detected by western blot with an antibody against the FLAG tag.

### **Gel retardation assays**

Radiolabelled purified RsaA or RsaA mutC (50000 cps/sample, concentration < 1 pM) and the cold mRNAs were renatured separately as described above. For each experiment, increasing concentrations of cold mRNAs (20-240 nM for *ssaA2\_3*, 50-900 nM for the other mRNAs) were added to the 5' end labelled RsaA or RsaA mutC in a total volume of 10 µl containing the ToeP+ buffer. Complex formation was performed at 37°C during 15 min. After incubation, 10 µl of glycerol blue was added and the samples were loaded on a 6% polyacrylamide gel under non denaturing conditions (1h, 300 V, 4°C). Quantification of the data corresponding to free RsaA or RsaA mutC and to RsaA/mRNA complex present on the autoradiography was done with ImageQuant TL software (GE Healthcare Life Sciences). Under these conditions where the concentration of the labelled RNA is negligible, the K<sub>d</sub> dissociation rate constant can be evaluated as the concentration of the cold RNA that gave 50% of binding.

### **Toeprinting assays**

The preparation of *E. coli* 30S subunits, the formation of a simplified translational initiation complex with mRNA, and the extension inhibition conditions were performed as described previously (Fechter *et al.* 2009). Increasing concentration of either RsaA or RsaA mutC concentrations were used to monitor their effects on the formation of the initiation complex with *mgrA*.

### **Mapping RsaA-mRNA interactions using footprinting**

RsaA- *flr* mRNA formation was carried out at 37°C for 15 min in ToeP+ buffer. Enzymatic hydrolysis was performed on the unlabelled mRNA (1 pmole) free or bound to RsaA in the presence of 1 µg carrier tRNA at 37°C for 5 min in 20 µl of ToeP+ buffer. The reactions were done with RNase T1 (0.0025 units), RNase V1 (0.5 units), or RNase T2 (0.05 U). The reactions were stopped by phenol extraction followed by RNA precipitation. The enzymatic cleavages in *flr* mRNA were detected by primer extension with reverse transcriptase according to Chevalier *et al.* (2009).

### **Relative quantification of RNAs by RT-qPCR**

Fresh BHI broth was inoculated with an overnight culture to an initial OD<sub>600nm</sub> of 0.05 and grown 6h at 37°C with shaking (200 rpm). The bacterial cells were harvested,

washed in 10 mM Tris buffer and diluted to a standard OD<sub>600</sub> of 1.0. One milliliter of diluted bacterial suspension was treated with 20 µg lysostaphin (1 mg/ml). The RNA isolation was performed using the RNasy Plus mini kit (QIAGEN) according to the manufacturer's instructions. Then, 1 µg of total RNA was transcribed into cDNA using Reverse Transcriptase System (Promega). The cDNA was used as a template for the real-time PCR amplification using a LightCycler instrument (Roche) and the specific primers shown in Table S2. The relative amounts of amplicons for each gene were determined using quantitative PCR relative to an internal standard (*gyrB*). The RT-qPCR experiments were processed in triplicates.

### **Differential Proteomics for cytoplasmic and secreted proteins**

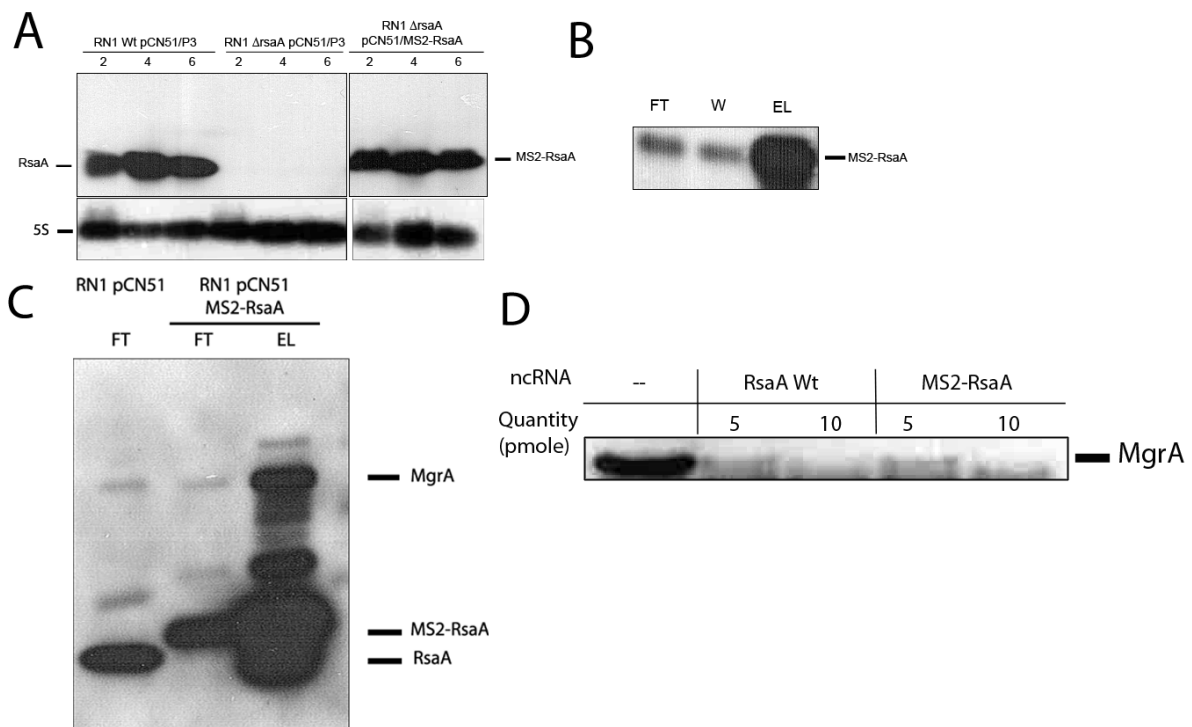
Triplicate protein extracts from supernatant or cytoplasm of HG001 (WT),  $\Delta$ *rsaA* mutant strain (LUG1630), and the same strain complemented with a plasmid expressing *RsaA*, were analyzed in separate LC/MS experiments. MS/MS spectra number were compared for each protein. Total protein extracts were prepared as follows: 1,5 mL of a *S. aureus* culture (OD<sub>600nm</sub> = 5) was centrifuged and the pellet resuspended in 150 µl of lysis buffer containing Tris 10 mM, NaCl 20 mM, EDTA 1 mM, MgCl<sub>2</sub> 5 mM, lysostaphin 50 µg/ml, 15 µl of protease inhibitors cocktail (Thermo Fischer Scientific), 2 µl DNase 10 U/µl (Roche), 2 µl RNase 500 µg/ml (Roche) and incubated for 30 min at 37 °C. Then 1 ml Trizol Reagent (Life Technologies). Upon addition of 100 µl 2-benzyl-4-chlorophenol (BCP), cells were thoroughly mixed followed by 10 min centrifugation (13000 rpm, 4°C). The upper-phase was then removed and ethanol 100% was added to the lower and pink phase. After a second centrifugation, the supernatant was placed into a new tube, and 7 volumes of cold 100% acetone were added. The proteins were then precipitated at least for 2h at -20°C. Protein pellets were centrifuged and washed with cold 70% acetone, and air-dried. Supernatants of cultures were filtered through a 0,22 µM membrane, precipitated with 5 volumes of 0.1 M ammonium acetate in 100% methanol, washed in 0.1 M ammonium acetate in 80% methanol and air-dried. Cytoplasmic or the secreted proteins (5 µg) were precipitated with 5 volumes of 0.1 M ammonium acetate in 100% methanol and then resuspended in 50 mM ammonium bicarbonate. After a reduction-alkylation step (5 mM DTT, 10 mM iodoacetamide), proteins were digested overnight with 1/25 (W/W) of trypsin and only 3 µg of proteins was vacuum

dried. The dried peptides were resuspended in 20  $\mu$ L of water containing 0.1% formic acid (solvent A). The peptide mixtures were analyzed using a NanoLC-2DPlus system coupled to a TripleTOF 5600 mass spectrometer (ABSciex) as previously described (Durand *et al*, 2015). Data were searched against the *S. aureus* N315 database from SwissProt. The first algorithm used was Mascot (version 2.2, Matrix Science, London, UK) through the ProteinScape 3.1 package (Bruker). A Spectral Counting quantitative strategy was carried on using the Mascot identification results and Proteinscape 3.1 package. To normalize the data, a correction factor was applied for each condition according to the average total spectra number for all samples.

## RESULTS

### MS2-RsaA is functional *in vivo*

The MS2 tagged version of RsaA was expressed from a plasmid under the control of the *agr*-dependent P3 promoter allowing an accumulation of the RNA at the stationary phase of growth in the mutant  $\Delta$ rsaA strain (LUG1630; Table S1). RsaA was detected by Northern blot using total RNAs extracted at 2, 4 and 6 h of growth. Using a DIG-labelled RsaA probe, we show that MS2-RsaA is efficiently transcribed from the P3 promoter (Figure 1A) in a manner highly similar to RsaA expressed from its endogenous promoter in the WT strain (Figure 1A). These data show that the steady state levels of RsaA were not affected by the presence of the MS2 tag. Affinity chromatography was then performed to verify the specific retention of MS2-RsaA by the MS2-MBP fused protein. The RNAs extracted from the different fractions (Flowthrough, Wash and Elution) were analysed using Northern blot. The data showed that only MS2-RsaA was specifically retained on the column (Figure 1B). Because we have previously shown that RsaA is a translational repressor of *mgrA* mRNA, we further have tested whether *mgrA* mRNA was co-purified together with MS2-RsaA. The Northern blot performed with a DIG-RNA probe complementary to *mgrA* revealed that *mgrA* mRNA was indeed enriched in the eluted fraction (Figure 1C).



**Figure 1 : The MS2-RsaA variant is normally expressed, correctly retained by affinity chromatography and still functional.** (A) Northern blots showing the expression of RsaA and the MS2-RsaA variant in RN1 and RN1  $\Delta$ rsaA strains. Total RNAs were prepared after 2, 4 and 6 hours of culture in BHI medium at 37°C. (B) Northern blot targeting MS2-RsaA performed on RNAs extracted after MS2 chromatography affinity; 1  $\mu$ g of total RNA was loaded on a 2% agarose gel. FT : Flow-Through ; W : Washing ; EL : Elution. (C) Northern blot targeting RsaA, MS2-RsaA and *mgrA* on RNAs extracted after MS2 chromatography affinity; 1  $\mu$ g of total RNA was loaded on a 2% agarose gel. FT : Flow-Through ; EL : Elution (D) *In vitro* translation assay performed with PURESYSTEM. The reactions were performed with 10 pmol of wild-type (WT) *mgrA* mRNA and in the presence of increasing quantities of WT RsaA or MS2-RsaA. The proteins were separated by SDS-PAGE (10%) and were revealed using a FLAG-specific antibody.



In order to verify that the MS2-tagged derivative of RsaA is functional, *in vitro* translation assays (PURE system) were performed with the whole *mgrA* mRNA whose translation is known to be repressed by RsaA (Romilly et al., 2014). A FLAG tag was added at the C-terminus part to MgrA to monitor its synthesis using Western blot analysis. Addition of increasing concentrations of RsaA or MS2-RsaA led to a strong inhibition of MgrA synthesis (Figure 1D).

All in all the addition of the MS2 tag to RsaA had no significant effect on its expression and its regulatory activities. Therefore, the MAPS approach was continued with the identification of the targetome of RsaA.

### **The MAPS approach reveals novel mRNAs as potential targets of RsaA.**

In the following experiments, we have prepared crude extracts from the  $\Delta$ rsaA mutant strain expressing MS2-RsaA or from the wild type strain expressing RsaA from the chromosome and the MS2 tag alone from a plasmid (Table S1). In this latter strain, we have shown that RsaA is expressed at a level that is comparable to MS2-RsaA (Figure 1A). Bacterial growth was done in BHI medium for 6h where the sRNA should be sufficiently abundant for the identification of the targets. After affinity chromatography, 100 ng of DNase treated and purified RNAs were used for cDNA synthesis and sequencing by a random hexaprimer approach. In parallel, we have performed with total RNA extracts of the same strains, a transcriptomic analysis to gain some indication on the expression levels of the mRNAs (Table S3). The differential expression analysis revealed no major changes in their transcriptomes and also confirmed an RsaA level only slightly more abundant when MS2-RsaA is expressed on the plasmid (1,7 fold). We could also estimate the cellular abundances of each of the transcripts of the RsaA targetome, which has been used in combination with the enrichment values of the MAPS analysis to estimate *in vivo* affinities between RsaA and its targets (supp. Materials and Methods and Table S5).

Results obtained with MS2-RsaA are shown in the Tables 1 and S4. We used Galaxy project (Afgan et al., 2016) and the Integrative Genomics Viewer (IGV) browser (Thorvaldsdóttir et al., 2013) to analyze and visualize our data, respectively. Detailed protocol for the bioinformatics analysis is provided in Material and Methods.

Target	Product	Base-pairing prediction	$\Delta G$ (Cofold) (kcal/mol)	$\Delta G$ (IntaRNA) (kcal/mol)	Affinity rank	$K_D^*$ (nM)
HG001_00625 mgrA	Transcriptional regulator MgrA	CGGAUAGUAUGUCUGGAGGAGAACUUUAUGUCUGAUC--ACAUA [51 nts] UCGCUACU GCUCAUCAGA-AGGAACCCUCUUGAAAUAUGA-UAGUGAUGUAU [30 nts] AGCGAUGA	-19	-17,8	++	≈20
HG001_02338 ssaA2_3	Staphylococcal secretory antigen SsaA precursor	CUUGGAGGACUUUAUUC AUGAAAAAUUAGUAACAGCAACUACGU GAACCCUCUUGAAAU-----UGAUGAUAG----UGAUGUA	-20,33	-7,9	+++	≈40
HG001_01977	Bacterial SH3 domain protein	AUCUCAACAACAUAGAAUAU-UACAGUCGCUAC UGGAGU-----UAUCUUUA-AU AUGUCAGCGAUG	-29,62	-15,8	++	≈200
HG001_02583 ssaA2_4	Staphylococcal secretory antigen SsaA precursor	CUUGGAGAUAAAAUGGAAUAUAAAAAGAUACUAAUUCGUUUAUUAUUGCUUUUGCAGUAC GAAACCUCU-----UGAAUUGAUGUA-GUGAUG	-18,38	-12,76	++	n.d.
HG001_02334 ssaA2_2	Staphylococcal secretory antigen SsaA precursor	UUUUUAGG-AGGAU AUUUUUAAAAUGAAGAAAAUCGCUACAGCUACUACGCG AGGAA-CCUCUUG-----AAAUUGAAUUGAUGAUGAG	-20,24	-11,6	++	n.d.
HG001_01615 cycA_1	D-serine/D- alanine/glycine transporter	GGGAGUUGUAUUUUGGCUGAAAA-AUUACAAAGGGAAC CCUCU---UGAAUUGAUGAUAGUGAUGUAUA-UUUG	-16,78	-21,2	-	n.d.
HG001_01902 rplS	50S ribosomal protein L19	GGGAAUUUAUAUGACAAUACAC CCUCUUGAA-AUUGAUGA-UAGUG	-18	-8,52	++	n.d.
HG001_01204	Hypothetical protein				++	
HG001_01150 glpF	Glycerol uptake facilitator protein	UUUUUGGAGGACAUUUAAU AGGAAACCUCUUG-AAAUUG	-15,3	-11	+	n.d.
HG001_01863	Hypothetical protein				++	
HG001_01003 flr	FPLR1 inhibitory protein precursor	GGGAGAACUU--AUU AUGAAAAAAAAUAUCACAAAAACUAUUUUCGCAUCUACA-GUAA CCUCUUGAAAUUGAUG-----AUAGUG UGAUGAUAGUG---AUGUAUAU	-19,38 (1) -15,31(2)	-13,29(1)	+++	≈100
HG001_01925 norB_4	Quinolone resistance protein norB	GGGAGG [375 nts] AACUUUGUCUAUUUUUAAGUCAUUU-AC CCUCU-----UUGAAUUGAUGAUAGUGAUGUAUAUUUG	-11,16	-12,7	+	n.d.
HG001_02606 ssaA_2	Staphylococcal secretory antigen SsaA precursor	UUUUUAUUUUGGAGGAUUUUAAUU AUGAAAAAAUCGUUACAGCUACAUCGCUACA CAGAA--GG-AAACCUCUUG-AAA-----UUGAUGAUGAUGAUGU	-25	-14,19	++	≈200

**Table S5 : *in silico* predictions of base-pairing for selected mRNAs found by MAPS with MS2-RsaA.** The red bases indicate base-pairing. In green are shown the Shine-Dalgarno sequence and the translation start codon AUG (GUG for SH3 mRNA).  $\Delta G$  for RNA hybrids were determined with CoFold and IntaRNA. Affinity rank is the ratio between the reads (normalized for the gene length and the deepness of the sequencing - “FPKM”) of RsaA\*mRNA and the MAPS. KD was experimentally and approximatively calculated by gel shift assays (see **figure 2**).

Briefly, after having aligned the reads onto our sequenced genome (Caldelari et al., 2016), counted per feature and normalized, we have estimated the enrichment of putative RNA binding partners by comparing the number of reads obtained from the tagged MS2-RsaA purification and the MS2 alone as control (also containing the MS2 tag alone). The reads were observed all along the co-purified RNA and RsaA showing that the RNAs were not degraded during the treatment (Figure S1). The data were reproduced in two independent experiments. Interestingly, the best enrichment was obtained for *mgrA* mRNA, which was previously experimentally validated (Romilly et al., 2014). We decided to fix the cut-off at around 2,5-fold using the *deseq2* method. Using these criteria of selection, we obtained a restricted list of candidates (Table 1). Interestingly, several of the mRNAs encode for proteins involved in common pathways. They include *ssaA2\_3*, *ssaA2\_4*, *ssaA\_2\_2* and *ssaA2*, all belonging to the family of staphylococcal secretory antigen A (SsaA). These proteins contain a CHAP domain corresponding to an amidase function involved in peptidoglycan metabolism (reviewed in Dubrac et al., 2008). Among the candidate targets, a mRNA encoding a protein (*orf HG001\_01977*) containing a SH3 domain, which is detected in glycyl-glycine endopeptidases like lysostaphine, was identified. We also found *flr* mRNA, which encodes a secreted immunomodulatory molecule that interferes with effective opsonization by complement and/or IgG antibodies (Stemerding et al, 2013). Other mRNAs encode the ribosomal protein L19, a drug efflux pump NorB that confers quinolone resistance, and a D-serine/D-alanine/glycine transporter. Transcriptomic analysis revealed that among the co-purified mRNA, even weakly expressed mRNAs as *ssaA\_2* or *flr* were specifically pulled out with MS2-RsaA. Therefore there is no direct correlation between mRNA abundance and their enrichment after the affinity chromatography.

We then searched for potential intermolecular basepairing interactions between RsaA and its potential mRNA targets using RNACoFold and IntaRNA (Tables 1, S5). Again, the best hit was found for *mgrA* mRNA, but stable pairings were also predicted with the four *ssaA* mRNAs and *flr* mRNA. Less stable interactions were also detected for *norB*, *cycA*, *glpF* and *rpIS*. Interestingly, most of the prediction involved the conserved UCCCA motif of RsaA and the Shine and Dalgarno (SD) sequence of the mRNAs. Only for *HG001\_01977* mRNA, stable pairings involved the 5'UTR of the mRNA with the first apical loop of RsaA (Table S5). In many cases such

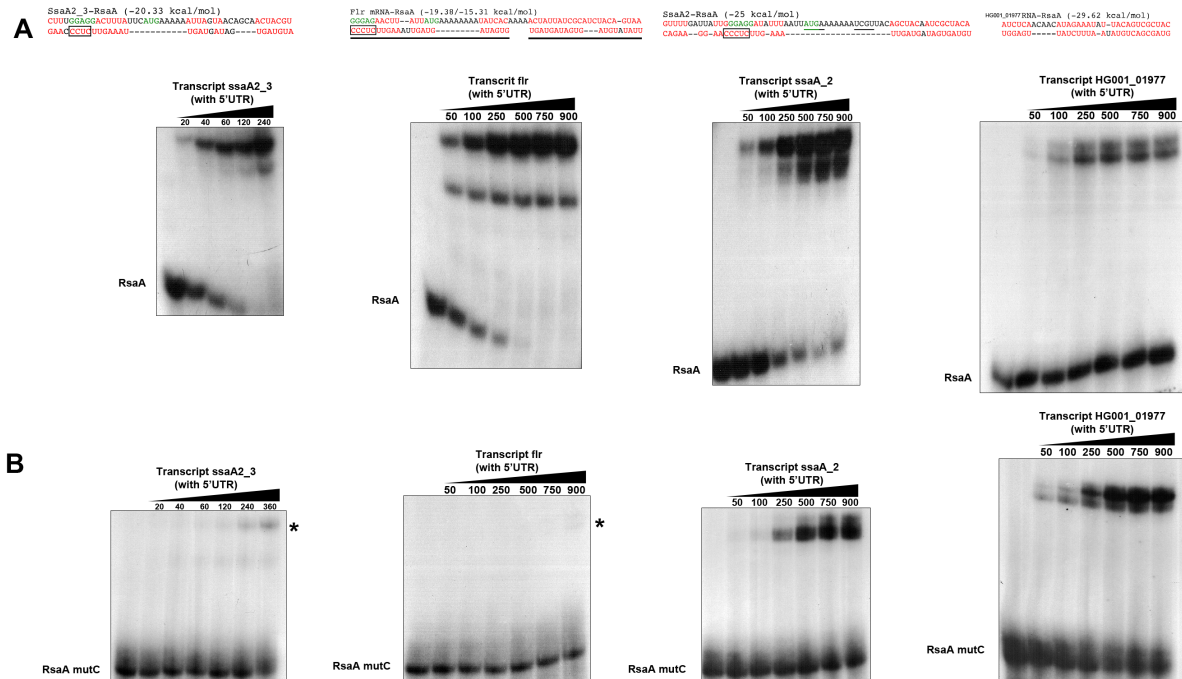
as *mgrA*, *ssaA*, *flr*, a second distant region of interactions was predicted in the coding sequences (Table S5).

These data suggested that RsaA would bind directly to many of these mRNAs, and would repress the translation through the sequestration of the SD sequence.

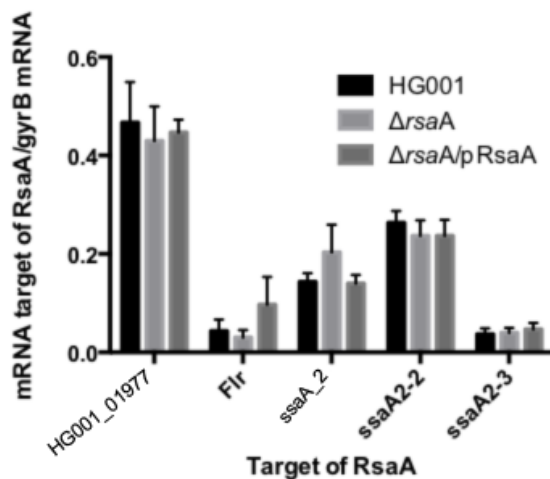
### **RsaA forms stable complexes with various mRNAs**

Based on the MAPS data, we first analyzed whether RsaA directly binds to the mRNA candidates using gel retardation assays (Figure 2). *In vitro* 5' end-labeled RsaA was incubated with increasing concentrations of the various mRNAs including the four *ssaA* species, *flr*, *norB-4*, *cycA*, *HG001\_01977*, *rplS*, and *glpF*. All the mRNA fragments contain the whole 5'UTR and at least 120 nts of the coding region (Table S2). The experiments were performed in a buffer containing magnesium and salt, without prior denaturation of the mRNA targets. The data showed that RsaA formed complexes with high affinity (see Kd value in Table S5) with *ssaA\_2*, *ssaA2\_3* and *flr*. For *flr* mRNA, two distinct complexes were observed, which can happen only if two RsaA molecules are able to bind to the mRNA. Affinity of complexes between RsaA and *HG001\_01977* or *ssaA2\_2* was weaker (> 200 nM), but the deletion of the 5'UTR of *ssaA2\_2* significantly enhanced duplex formation (Figure S2A). Finally, no gel retardation was observed when complex formation was performed with RsaA and increasing concentrations of *glpF*, *rplS*, *norB-4*, and *cycA* mRNAs (up to 1  $\mu$ M, Figure S2B).

Because the UCCCA sequence of RsaA was predicted to form basepairings with the SD of mRNA targets, we have analyzed whether mutations at these positions would affect RsaA binding as it was demonstrated for *mgrA* (Romilly et al., 2014). Complex formation was carried out under strictly identical conditions for the wild type and mutated RsaA (RsaA mutC). The data showed that RsaA mutC failed to bind to *ssaA2\_3* and *flr* mRNAs (Figure 2B). Surprisingly, the second complex with *flr* disappeared too when using RsaA mutC, which would indicate a cooperation between both sites. Hence the C-rich motif of RsaA might be the seed sequences initiating mRNA binding. In contrast, the mutations in RsaA mutC slightly altered the stability of the complex formed with *ssaA2* mRNA. It has no significant effect on SH3-like mRNA binding as expected, while RNACofold did not predict any involvement of the UCCCA motif of RsaA to interact with *HG001\_01977* mRNA (Figure 2).



**Figure 2: Gel retardation assays to monitor RsaA binding to several mRNAs.** Experiments were performed on four different possible mRNA targets of RsaA, which are *ssaA2\_3*, *flr*, *ssaA\_2*, *HG001\_01977*. The 5' end-labelled Wt RsaA (A) or the mutant RsaA mutC (B) were incubated with increasing concentrations (nM) of mRNAs. In the case of *flr* mRNA, two complexes could be shown with WT RsaA. In the gel shift assays with RsaA mutC, residual complexes are indicated with a star.



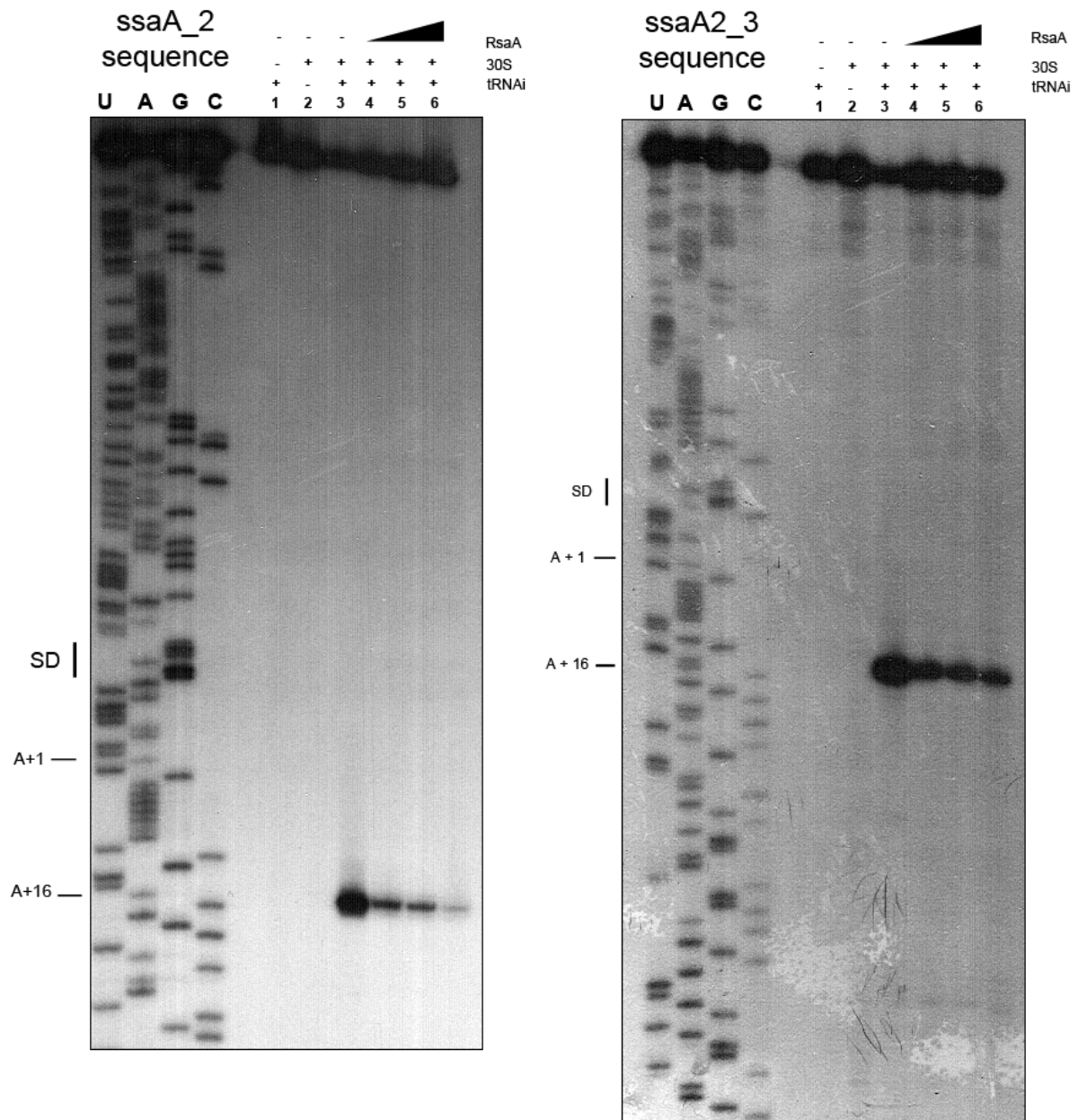
**Figure 3: The expression of target mRNAs was quantified by RT-qPCR.** The data were normalized to the level of *gyrB* mRNA expression from total RNA extracts prepared from *in vitro* culture to the late-exponential phase (6h) of HG001 (RN1 WT), the  $\Delta$ rsaA mutant strain, and the complemented strain with a plasmid expressing RsaA under its own promoter.

Taken together, the gel retardation assays correlate well with the prediction of the pairing interactions, and suggested that the C-rich motif of RsaA is one of its major regulatory determinant sites.

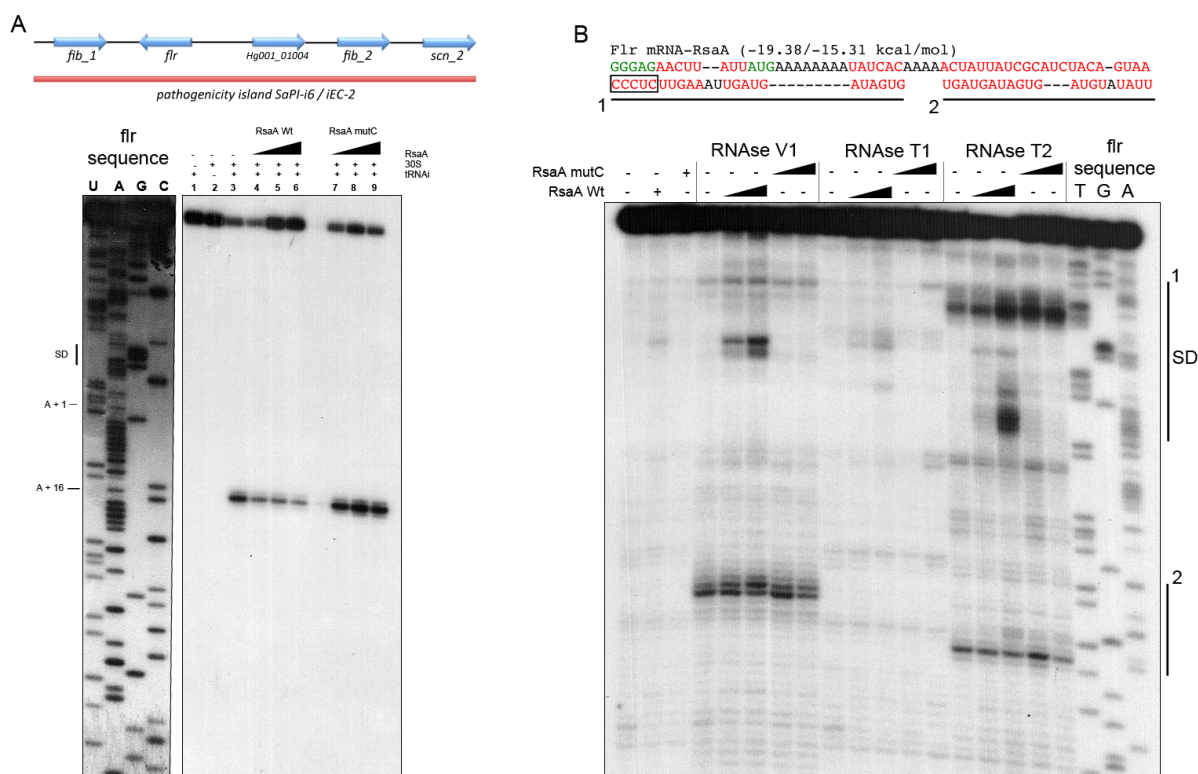
### **RsaA hinders ribosome binding site of several mRNA targets**

While sRNA-mRNA pairings can alter mRNA decay (stabilization/degradation), we have probed whether RsaA affected the steady state levels of the target mRNAs. We therefore performed RT-qPCR in triplicates on selected target mRNAs using total RNAs extracted from the wild type HG001 strain, the isogenic HG001 $\Delta$ rsaA mutant strain, and HG001 $\Delta$ rsaA complemented with pCN51/pRsaA (Table 2). The levels of each mRNA were quantified relatively to *gyrB* (Figure 3). The data showed that the deletion of *rsaA* did not significantly affect the steady state levels of *SH3*, *flr*, *ssaA2\_2* and *ssaA 2\_3* mRNAs. Only for *ssaA2* mRNA, the absence of RsaA led to a reproducibly weak increase of the mRNA yield. Such a mild RsaA-dependent effect was also observed for *mgrA* mRNA (Romilly et al., 2014).

Toe-printing assay was then used to decipher the effect of RsaA on the formation of the ternary initiation complex formed in the presence of mRNA, the initiator tRNA and the 30S subunit. The experiments performed with *flr*, *ssaA2*, *ssaA2\_2* and *ssaA2\_3* mRNAs, showed that the ribosome efficiently recognized the mRNAs as illustrated by the presence of the toe-print signal at position +16 (+1 being the initiation codon). The addition of RsaA together with the 30S subunits strongly decreased the toe-print showing that RsaA competes efficiently with the ribosome for mRNA binding (Figures 4, 5). The RsaA-dependent inhibition was albeit less efficient for *ssaA2\_2* (Figure S3). Toe-printing experiment was also performed for *HG001\_01977* mRNA (Figure S4). However in this case, we did not observe any toe-printing signal correlated with the absence of a strong SD sequence. The addition of RsaA caused a specific pause of reverse transcriptase that corresponded to the predicted basepairing interactions (Figure S5, Table 1). For *flr* mRNA, the formation of the initiation complex was completely restored when RsaA mutC was added in the reaction (Figure 5). This confirms the gel retardation assays results, which showed that the C-rich motif of RsaA is essential for stable duplex formation.



**Figure 4: Toeprinting assays to monitor the effect of RsaA on the initiation ribosomal complex formation.** Effect of RsaA on the formation of initiation ribosomal complex on *ssaA\_2* and *ssaA2\_3* mRNAs. Lane 1 : control incubation of mRNA ; lane 2 : control incubation of mRNA with 30S ribosomal subunits, without tRNAi allowing the formation of the initiation complex ; lane 3 : formation of the ribosomal initiation complex containing mRNA, the 30S subunits and tRNAi ; lane 4 to 6 : formation of the ribosomal initiation complex in presence of increasing concentrations of WT RsaA : 100, 200, 400 nM. Lanes U, A, G, C : sequencing ladders. The Shine-Dalgarno (SD) sequence, the start site of translation (A +1 of the AUG initiation codon) and the toe-printing signals (N+16) are indicated.



**Figure 5: Characterisation of the inhibitory RsaA-*flr* mRNA complex.** (A) Genetic organisation of part of the pathogenic island containing the *flr* gene and effect of RsaA on the formation of initiation ribosomal complex on *flr* mRNA. Lane 1 : control incubation of mRNA ; lane 2 : control incubation of mRNA with 30S ribosomal subunits, without tRNA<sub>i</sub> allowing the formation of the initiation complex ; lane 3 : formation of the ribosomal initiation complex containing mRNA, the 30S subunits and tRNA<sub>i</sub> ; lane 4 to 6 : formation of the ribosomal initiation complex in presence of increasing concentrations of WT RsaA : 250, 500, 1000 nM. lane 7 to 9 : formation of the ribosomal initiation complex in presence of increasing concentrations of RsaA mutC : 250, 500, 1000 nM. Lanes U, A, G, C : sequencing ladders. The Shine-Dalgarno (SD) sequence, the start site of translation (A +1 of the AUG initiation codon) and the toe-printing signals (N+16) are indicated. (B) Prediction of basepairing interactions between RsaA and *flr* together with enzymatic reactions performed on *flr* mRNA bound to increasing concentrations of RsaA (lane 4, 25 nM; lane 5, 50 nM; lane 6, 250 nM). Lanes 1, 2, 3: incubation controls on free mRNA or bound to RsaA Wt or mutC, respectively. Lanes T, G, A: dideoxy sequencing reactions. The experiments were performed with RNase V1 (V1), Rnase T1 (T1) and RNase T2 (T2). 1 and 2 represent both sites of interactions (underlined in the genome schema). S means Shine and Dalgarno sequence.



Because the gel retardation assays suggested that two RsaA molecules could bind to *flr* mRNA, we have performed footprinting experiments using the double-strand specific RNase V1 and the single-strand specific RNase T2 (Figure 5). The data revealed that the addition of RsaA induced significant changes in the region encompassing the SD sequence and part of the coding sequence from G-13 to A+43. Enhanced RNase V1 cuts were located close to the SD sequence (at positions -6 to -9) and in the coding sequence at C+30. Concomitantly, binding of RsaA reduced RNase T2 hydrolysis at positions C15, A16, C17, A33, and A42-43 while enhanced cleavages were observed at A4 to A10. This large area of changes is well correlated with the positioning of two RsaA molecules (Figure 5). In this model, the CCCU motif would interact with the SD sequence.

In conclusion, these data strongly suggested that RsaA would primarily regulate translation initiation of several of the newly identified target mRNAs by sequestering their ribosome binding sites.

### **Quantitative proteomics showed that RsaA indirectly regulates cell surface proteins**

Cytoplasm and secreted proteins extracts from the WT strain (HG001), the mutant HG001 $\Delta$ rsaA (LUG1630) strain, and the same strain complemented with a plasmid expressing RsaA under its own promoter were analysed in triplicates by LC/MSMS coupled to spectral count analysis. The synthesis of several proteins, for which the mRNAs have been demonstrated as primary targets of RsaA, were found enhanced in the mutant strains. These proteins include MgrA, SsaA\_2, SsaA2\_3 and FLIPr (Tables 2, S6). FLIPr was poorly detected in the culture supernatant, only by one MSMS spectrum in the analysis, however it was enriched by a factor of almost 2 in the  $\Delta$ rsaA strain. In these experiments, the proteins CycA and NorB\_4 were not detected while the synthesis of L19 and GlpF were found unchanged in the various strains. Interestingly, several surface proteins were found significantly activated by RsaA. These proteins include the extracellular matrix protein EbH, the immunoglobulin G binding protein A, the surface protein G SasG, the Serine-aspartate repeat-containing proteins C and D StrC/StrD, and the clumping factor ClfB. Because the corresponding mRNAs were not detected in the MAPS approach, these effects are most likely indirect.

## DISCUSSION

Many bacterial sRNAs act as antisense RNAs and form short and non-contiguous basepairing interactions with a group of mRNAs often functionally related to adapt cell growth in response to specific stress, environmental changes and various hosts. Although *S. aureus* genome encodes for numerous sRNA genes, the functions of only few of them have been identified. This is linked to the fact that pulse chase experiments followed by transcriptomic analysis has not been so well appropriate to monitor the fast regulatory events. A recent approach, the so-called MAPS, has been recently developed to detect *in vivo* the whole sets of mRNAs interacting with a defined sRNA. This powerful approach has been applied on two *E. coli* sRNAs, RyhB and RybB (Lalaouna et al., 2015). For the first time, this study provided a whole sRNA targetome and besides the previously identified mRNA targets, it has revealed the unexpected existence of sponge RNAs issued from precursor tRNAs. Here, we demonstrated that this technology could also be applied to *S. aureus* sRNAs independently of their regulatory mechanisms, and in a wild type strain context expressing all ribonucleases.

As a proof of concept, we have adapted the protocol to one small non coding sRNA, RsaA, which functionally links two master regulatory factors, Sigma B and MgrA proteins (Romilly et al., 2014). We first showed that the tagged RsaA behaved as the wild type sRNA and was able to efficiently regulate the translation of *mgrA* mRNA (Figure 1). Analysis of the enriched RNAs that co-purified with RsaA revealed that the first most abundant mRNA (15-fold) was *mgrA*. A second set of mRNAs (> 3 fold) encode several amidases that are involved in peptidoglycan metabolism and remodelling (reviewed in Dubrac et al., 2008). These enzymes play also a role in biofilm-associated infections (Resch et al, 2005). A third group of mRNAs (below 3 fold) encodes for various proteins that are associated to the membrane (transporters, efflux pump), and the secreted FLIPr protein. Search for basepairing interactions revealed that most of the mRNAs can potentially formed stable duplexes with RsaA involving either its interhelical unpaired domain encompassing the C-rich motif or the first hairpin loop motif. It is of interest that a potential correlation existed between the order of the enriched mRNAs and the strength of the predicted basepairings (Table 1). However, gel retardation assays revealed that only several mRNAs (*flr*, *mgrA*,

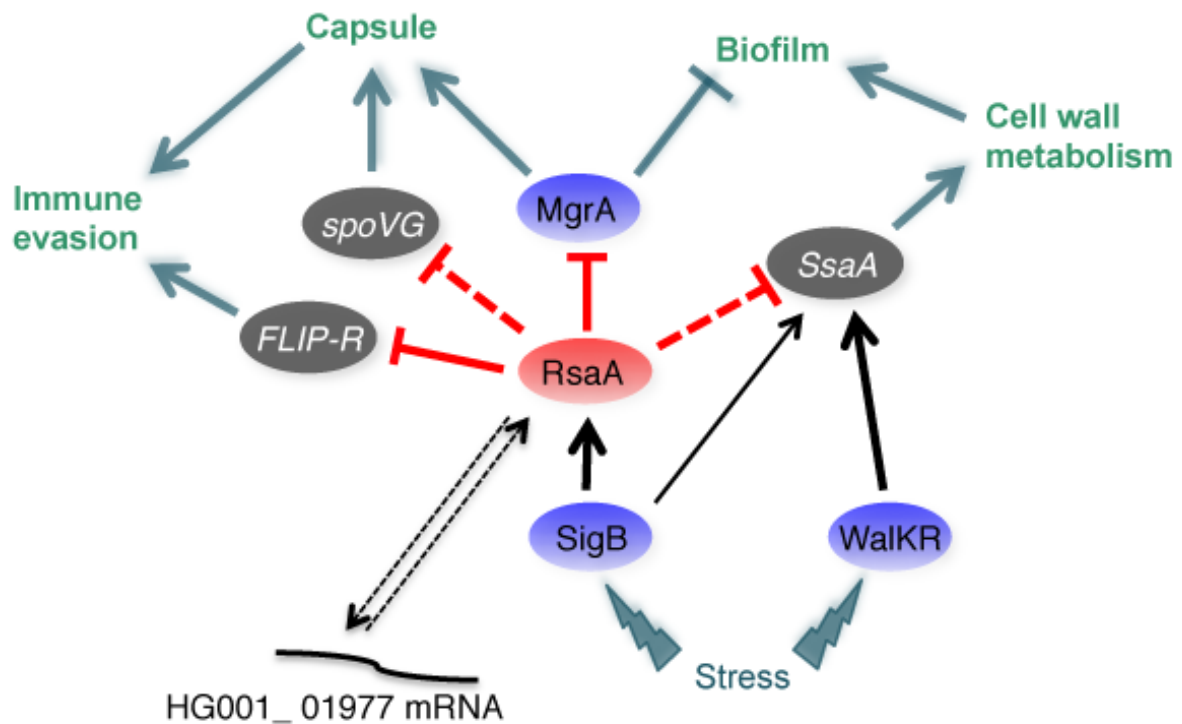
*ssaA\_2*, *ssaA2\_2*, *ssaA2\_3*, *HG001\_01977*) were competent to form stable complexes with RsaA. For the other mRNAs (*glpF*, *cycA-1*, *norB-4*, *rplS*, *ssaA2\_4*), one cannot exclude that a trans-acting RNA-binding protein would help to facilitate the formation of complexes, and/or alternatively that RsaA would bind to the mRNAs during the transcription process to avoid secondary structure elements masking the regulatory elements. However, we also cannot omit that these RNAs were retained together with another partner, through an indirect manner. For instance, in the list of enriched RNAs, we have recovered the quorum sensing induced regulatory RNAIII (enrichment just below 2,5), which has been recently shown to stabilize *mgrA* mRNA through direct binding with the 5'UTR of one of the *mgrA* transcripts (Gupta et al., 2015). Therefore, it is highly possible that RNAIII was co-purified together with *mgrA*. Toe-printing assays further demonstrated that RsaA primarily regulates translation initiation of *flr*, *ssaA2* and *ssaA2\_3* mRNAs. As established for *mgrA*, the C-rich motif of RsaA binds to the strong SD sequences of the mRNAs but in addition complementarities extend to the coding sequences of the target mRNAs. The regulatory regions present in the four *ssaA* mRNAs are highly conserved. Based on our data, we proposed that for *ssaA* mRNAs, the two interaction sites involve one molecule of RsaA while for *flr* mRNA, two RsaA molecules cover almost 40 nucleotides of the mRNA (Figures 2, 5). Interestingly enough, mutations in the C-rich motif hindered duplex formation and restored ribosome binding for *flr* mRNA, *ssaA\_2*, *ssaA2\_3* mRNAs arguing for the existence of initial pairings to occur at the SD sequence. Proteomic analysis showed that RsaA significantly repressed the synthesis of these proteins (MgrA, FLIPr, SsaA2\_3, SsaA\_2). All these data strongly supported that these mRNAs are direct targets of RsaA.

The first 5' hairpin loop of RsaA is a second regulatory site, which has been involved in the formation of a loop-loop interaction stabilizing *mgrA*-RsaA complex (Romilly et al., 2014). Here our results strongly suggested that this domain of RsaA might also bind to the large 5'UTR of the hypothetical mRNA encoding a SH3-containing protein (Figure S4). Indeed, a strong RsaA-dependent RT pause was found at the 3' boundary of the mRNA bound to RsaA, and mutation of the C-rich motif of RsaA did not alter the formation of the stable complex (Figure S4). It is intriguing that the steady state level of this RNA did not change in the WT and mutant strains (Figure 3). We also did not observe the formation of an active initiation ribosomal complex

questioning on the coding capacity of this hypothetical mRNA. The gene is located into a pathogenicity island, just downstream *sak* mRNA encoding staphylokinase (Figure S4). This exotoxin activates plasminogen to digest fibrin clots, but also inhibits phagocytosis by cleaving IgG and complement component. In the same area, several non coding RNAs have been described, one of them called SprE is expressed from the same strand between the *sak* gene and *HG001\_01977* (Pichon & Felden, 2005). Although, we need to address more specifically the functional consequences of *HG001\_01977* mRNA-RsaA complex, we cannot rule out that this hypothetical mRNA might regulate the activity of RsaA. Recent reports have indeed described novel RNA functions modulating the action of sRNAs; some of them act as RNA decoys (Figueroa-Bossi et al., 2009) or as RNA sponges (Miyakoshi et al., 2015) to regulate the stability of the target sRNA and consequently affect the expression of their regulons. More recently, Lalaouna et al. (2015) have demonstrated that the tRNA spacers can serve as RNA sponges to prevent transcriptional noise of several sRNAs in *E. coli*. In this latter case, the decay of the regulated sRNAs is not affected. It was proposed that such sRNA sponge would set up a concentration threshold that sRNAs must exceed to induce target regulation under specific conditions (Lalaouna et al., 2015). Finally, a bacteriophage encoded RNA called AgvB was able to counteract the action of the enterohemorrhagic *E. coli* GcvB sRNA through sequence mimicry (Tree et al., 2014). Its interaction does not alter the stability of GcvB although the synthesis of its targets was deregulated. Interestingly, *HG001\_01977* gene locus is located on a pathogenicity island, which arose from bacteriophage (Figure S3). Whether *HG001\_01977* regulates the RsaA action remains to be studied.

Analysis of the functions of the RsaA targets reveals several interconnected links. Four of the mRNAs encode proteins that belong to the CHAP amidase family endowed with D-alanyl-glycyl endopeptidase activity that cleave between the crossbridge and the stem peptide and one mRNA encoding a protein with a SH3 domain found in glycyl-glycine endopeptidases, that cut the pentaglycine bridge specific to staphylococcal cell wall. Cell wall metabolism is vital for bacteria and endopeptidases play multiple functions in cell shape, division, resistance to stresses, protein secretion, biofilm formation, and in host-pathogen interactions. These enzymes should also be tightly regulated to avoid perturbation of the cell wall

particularly at the stationary phase of growth. Indeed, they are under the control of several transcriptional regulatory proteins, among which the two-component system WalkR is essential (Dubrac et al., 2008; Delaune et al., 2011; Poupel et al., 2016). Inactivation of WalkR at the stationary phase of growth caused transcription arrest of these enzymes (Dubrac et al., 2008), and interestingly, the expression of SsaA and LytM were found sufficient to by-pass WalkR essentiality (Delaune et al., 2011). An additional post-transcriptional level of regulation was recently described in which *S. aureus* RNAIII contributed to repress *lytM* translation at high cell density (Lioliou et al., 2016). Hence, RsaA would be the second RNA that would coordinate the translation repression of SsaA family enzymes at the stationary phase of growth in a Sigma B-dependent manner. Among the new mRNA targets of RsaA, one of them encodes a secreted virulence factor, the FLIPr protein (Prat et al., 2006; 2009). Interestingly, the gene is located on an immune evasion cluster (pathogenicity island 4; Novick et al., 2010 and Figure 5) that includes staphylococcal complement inhibitor and several fibrinogen-binding proteins (Scn-2, Fib-1, Fib-2). FLIPr protein is a secreted staphylococcal anti-inflammatory protein that acts as an antagonist of the receptor protein FPRL1. This protein is expressed in various cells such as neutrophils, macrophages and microglia (Migeotte et al., 2006), and triggers many cellular functions including chemotaxis, exocytosis, and superoxide generation. Hence, by counteracting FPRL1, FLIPr contributes to the evasion of bacteria from the human defense system (Prat et al., 2006). This novel function of RsaA is somehow related to the translation repression of *mgrA* mRNA, a master regulator of *S. aureus* (Luong et al., 2006; Crosby et al., 2016). These studies showed that MgrA is a major activator of capsule formation and regulates autolytic activities, efflux pumps, several transporters, and virulence factors. Particularly, the two-component system ArIRS and MgrA formed a regulatory pathway that led to the repression of several cell surface proteins (SraP, SasG, FmtB, SdrD, and protein A). This repression confers to *S. aureus* the ability to form clumps in the presence of fibrinogen, which facilitates adhesion to the host tissues but also protect the bacteria. Interestingly, *mgrA* mutants are unable to form clumping due to the synthesis of the cell surface proteins Ebh, SraP and SasG. In addition, the *mgrA* mutant produced more biofilms due to the expression of SasG protein, which is known to favour intercellular contacts (Crosby et al., 2016). In agreement with this study, quantitative differential proteomic analysis



**Figure 6: Schematic drawing summarizing the regulatory networks involving RsaA and its mRNA targets in *Staphylococcus aureus* virulence.** RsaA activated by  $\sigma$ B binds to *mgrA*, *flr*, *spoVG* (Romilly et al, 2014) and *ssaA* mRNAs to inhibit their translation. The role of RsaA interaction with *HG001\_01977* is still unclear. Arrows are for activation, bars for repression. In blue are the transcriptional protein regulators, in red the regulatory RNA and in grey the virulence factors. Red lines corresponded to post-transcriptional regulation and black lines to transcriptional regulation. The regulatory events for which direct regulation is not yet demonstrated are shown by dotted lines.

revealed that the deletion of RsaA causes a strong decrease in the synthesis of the cell surface proteins EbH, protein A, SdrD, SdrC, SasG, and the clumping factor ClfB (Table S6). These effects are most likely indirect and linked to the repression of *mgrA*. Indeed, our previous study has also shown that the induction of RsaA inhibits capsule formation and enhances biofilm formation through the repression of *mgrA* (Romilly et al., 2014). The expression of RsaA was also accompanied by a decreased protection of *S. aureus* against opsonophagocytic killing by neutrophils, and acute systemic infection in mice was attenuated while chronic catheter infection was enhanced. Clearly, the RsaA-dependent translation repression of both MgrA and FLIPr supports these phenotypes while the repression of the SsaA family enzymes contributes to maintain the integrity of the cell wall, a property that is certainly essential to maintain the correct presentation of the cell surface proteins. Hence, RsaA is taking part in regulatory networks that modify the complex interactions with the host immune system (Figure 6). The MAPS approach also revealed that RsaA mediates crosstalks between the core genome and the mobile elements through the regulation of genes (*flr*, HG001\_01977) located on pathogenicity islands thus integrating horizontally acquired genes into pre-existing post-transcriptional networks (Papenfort and Vogel, 2010; Chabelskaya et al., 2010).

As it was shown for *E. coli* (Lalaouna et al., 2015), the MAPS approach can also be used to gain a better knowledge on the overall functions regulated by a specific sRNA in Gram-positive bacteria. Because these experiments have been conducted in wild type strain (in presence of any ribonucleases), we do not exclude that mRNAs that are rapidly depleted could be underestimated. In this work, we utilized an sRNA carrying a C-rich motif that has been previously shown to be an ideal motif for targeting the ribosome binding sites of mRNAs. This approach has also the advantage to look for RNA-binding proteins although crosslinking experiments would need to be introduced in order to stabilize the interactions that are often transient and dynamic. MAPS can also monitor RNA regulators of specific mRNAs providing the whole crosstalks that might exist at the post-transcriptional level (Lalaouna and Massé, 2015). More recently, novel high throughput techniques allowed to map at global scale RNA-RNA interactions *in vivo* in eukaryotes (Aw et al., 2016; Sharma et al., 2016). These methods, which combine *in vivo* crosslinks, ligation of interacting RNAs and high throughput sequencing, revealed an impressive diversity of intra- and

intermolecular RNA-RNA pairings. Such approaches applied to bacteria could certainly explore the complexity of RNA-RNA interactions taking place and to track their dynamics occurring under various stress conditions and during infection.

## ACKNOWLEDGEMENT

We are grateful to Anne-Catherine Helfer for her help in the footprinting assays, and to Lauriane Kuhn, Johana Chicher and Philippe Hammann for the mass spectrometry analysis (IBMC proteomic platform). We are thankful to Delphine Bronesky and David Lalaouna for helpful advices and discussions.

## FUNDING

This work was supported by the Centre National de la Recherche Scientifique (CNRS) to [P.R.] and has been published under the framework of the LABEX: ANR-10-LABX-0036 NETRNA to [P.R.], a funding from the state managed by the French National Research Agency as part of the investments for the future program. A Tomasini was supported by Fondation pour la Recherche Médicale.

## REFERENCES

1. Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C. *et al.* (2016) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Research*, 44, W3-W10.
2. Anders, S., Pyl, P.T. and Huber, W. (2015) HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*, 31, 166-169.
3. Aw, J.G.A., Shen, Y., Wilm, A., Sun, M., Lim, X.N., Boon, K.-L., Tapsin, S., Chan, Y.-S., Tan, C.-P., Sim, A.Y.L. *et al.* (2016) In Vivo Mapping of Eukaryotic RNA Interactomes Reveals Principles of Higher-Order Organization and Regulation. *Molecular Cell*, 62, 603-617.
4. Barquist, L. and Vogel, J. (2015) Accelerating Discovery and Functional Analysis of Small RNAs with New Technologies. *Annual Review of Genetics*, 49, 367-394.



5. Blankenberg, D., Gordon, A., Von Kuster, G., Coraor, N., Taylor, J. and Nekrutenko, A. (2010) Manipulation of FASTQ data with Galaxy. *Bioinformatics*, 26, 1783-1785.
6. Boisset, S., Geissmann, T., Huntzinger, E., Fechter, P., Bendridi, N., Possedko, M., Chevalier, C., Helfer, A.C., Benito, Y., Jacquier, A. *et al.* (2007) Staphylococcus aureus RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes & Development*, 21, 1353-1366.
7. Caldelari, I., Chao, Y., Romby, P. and Vogel, J. (2013) RNA-Mediated Regulation in Pathogenic Bacteria. *Cold Spring Harbor Perspectives in Medicine*, 3.
8. Carrier, M.-C., Lalaouna, D. and Massé, E. (2016) A game of tag: MAPS catches up on RNA interactomes. *RNA Biology*, 13, 473-476.
9. Chabelskaya, S., Gaillot, O. and Felden, B. (2010) A Staphylococcus aureus small RNA is required for bacterial virulence and regulates the expression of an immune-evasion molecule. *PLoS pathogens*, 6, e1000927.
10. Chao, Y., Papenfort, K., Reinhardt, R., Sharma, C.M. and Vogel, J. (2012) An atlas of Hfq-bound transcripts reveals 3s 3 and regulates the expression of an tory small RNAs. *The EMBO Journal*, 31, 4005-4019.
11. Charpentier, E., Anton, A.I., Barry, P., Alfonso, B., Fang, Y. and Novick, R.P. (2004) Novel Cassette-Based Shuttle Vector System for Gram-Positive Bacteria. *Applied and Environmental Microbiology*, 70, 6076-6085.
12. Chevalier, C., Geissmann, T., Helfer, A.-C. and Romby, P. (2009) Probing mRNA structure and sRNA-mRNA interactions in bacteria using enzymes and lead(II). *Methods in Molecular Biology (Clifton, N.J.)*, 540, 215-232.
13. Crosby, H.A., Schlievert, P.M., Merriman, J.A., King, J.M., Salgado-Pabón, W. and Horswill, A.R. (2016) The Staphylococcus aureus Global Regulator MgrA Modulates Clumping and Virulence by Controlling Surface Protein Expression. *PLoS pathogens*, 12, e1005604.
14. Delaune, A., Poupel, O., Mallet, A., Coic, Y.-M., Msadek, T. and Dubrac, S. (2011) Peptidoglycan Crosslinking Relaxation Plays an Important Role in Staphylococcus aureus WalkR-Dependent Cell Viability. *PLoS ONE*, 6.

15. Dubrac, S., Bisicchia, P., Devine, K.M. and Msadek, T. (2008) A matter of life and death: cell wall homeostasis and the WalkR (YycGF) essential signal transduction pathway. *Molecular Microbiology*, 70, 1307-1322.
16. Fechter, P., Chevalier, C., Yusupova, G., Yusupov, M., Romby, P. and Marzi, S. (2009) Ribosomal initiation complexes probed by toeprinting and effect of trans-acting translational regulators in bacteria. *Methods in Molecular Biology (Clifton, N.J.)*, 540, 247-263.
17. Felden, B., Vandenesch, F., Boulloc, P. and Romby, P. (2011) The Staphylococcus aureus RNome and Its Commitment to Virulence. *PLoS Pathogens*, 7.
18. Figueroa-Bossi, N., Valentini, M., Malleret, L. and Bossi, L. (2009) Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target. *Genes & Development*, 23, 2004-2015.
19. Geisinger, E., Adhikari, R.P., Jin, R., Ross, H.F. and Novick, R.P. (2006) Inhibition of rot translation by RNAIII, a key feature of agr function. *Molecular Microbiology*, 61, 1038-1048.
20. Geissmann, T., Chevalier, C., Cros, M.-J., Boisset, S., Fechter, P., Noirot, C., Schrenzel, J., François, P., Vandenesch, F., Gaspin, C. *et al.* (2009) A search for small noncoding RNAs in Staphylococcus aureus reveals a conserved sequence motif for regulation. *Nucleic Acids Research*, 37, 7239-7257.
21. Gupta, R.K., Luong, T.T. and Lee, C.Y. (2015) RNAIII of the Staphylococcus aureus agr system activates global regulator MgrA by stabilizing mRNA. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 14036-14041.
22. Huntzinger, E., Boisset, S., Saveanu, C., Benito, Y., Geissmann, T., Namane, A., Lina, G., Etienne, J., Ehresmann, B., Ehresmann, C. *et al.* (2005) Staphylococcus aureus RNAIII and the endoribonuclease III coordinately regulate spa gene expression. *The EMBO Journal*, 24, 824-835.
23. Koch, G., Yepes, A., Förstner, K.U., Wermser, C., Stengel, S.T., Modamio, J., Ohlsen, K., Foster, K.R. and Lopez, D. (2014) Evolution of resistance to a last-resort antibiotic in Staphylococcus aureus via bacterial competition. *Cell*, 158, 1060-1071.

24. Lalaouna, D. and Massé, E. (2015) Identification of sRNA interacting with a transcript of interest using MS2-affinity purification coupled with RNA sequencing (MAPS) technology. *Genomics Data*, 5, 136-138.
25. Lalaouna, D., Morissette, A., Carrier, M.-C. and Massé, E. (2015) DsrA regulatory RNA represses both hns and rbsD mRNAs through distinct mechanisms in *Escherichia coli*. *Molecular Microbiology*, 98, 357-369.
26. Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, 10, R25.
27. Lioliou, E., Fechter, P., Caldelari, I., Jester, B.C., Dubrac, S., Helfer, A.-C., Boisset, S., Vandenesch, F., Romby, P. and Geissmann, T. (2016) Various checkpoints prevent the synthesis of *Staphylococcus aureus* peptidoglycan hydrolase LytM in the stationary growth phase. *RNA biology*, 13, 427-440.
28. Luong, T.T. and Lee, C.Y. (2006) The arl locus positively regulates *Staphylococcus aureus* type 5 capsule via an mgrA-dependent pathway. *Microbiology (Reading, England)*, 152, 3123-3131.
29. Migeotte, I., Communi, D. and Parmentier, M. (2006) Formyl peptide receptors: a promiscuous subfamily of G protein-coupled receptors controlling immune responses. *Cytokine & Growth Factor Reviews*, 17, 501-519.
30. Miyakoshi, M., Chao, Y. and Vogel, J. (2015) Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. *The EMBO journal*, 34, 1478-1492.
31. Monk, I.R., Shah, I.M., Xu, M., Tan, M.-W. and Foster, T.J. (2012) Transforming the Untransformable: Application of Direct Transformation To Manipulate Genetically *Staphylococcus aureus* and *Staphylococcus epidermidis*. *mBio*, 3.
32. Novick, R.P., Christie, G.E. and Penadés, J.R. (2010) The phage-related chromosomal islands of Gram-positive bacteria. *Nature Reviews. Microbiology*, 8, 541-551.
33. Novick, R.P., Ross, H.F., Projan, S.J., Kornblum, J., Kreiswirth, B. and Moghazeh, S. (1993) Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *The EMBO Journal*, 12, 3967-3975.

34. Papenfort, K. and Vogel, J. (2010) Regulatory RNA in bacterial pathogens. *Cell Host & Microbe*, 8, 116-127.
35. Pichon, C. and Felden, B. (2005) Small RNA genes expressed from *Staphylococcus aureus* genomic and pathogenicity islands with specific expression among pathogenic strains. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 14249-14254.
36. Poupel, O., Moyat, M., Groizeleau, J., Antunes, L.C.S., Gribaldo, S., Msadek, T. and Dubrac, S. (2016) Transcriptional Analysis and Subcellular Protein Localization Reveal Specific Features of the Essential WalkR System in *Staphylococcus aureus*. *PLoS One*, 11, e0151449.
37. Prat, C., Bestebroer, J., Haas, C.J.C.d., Strijp, J.A.G.v. and Kessel, K.P.M.v. (2006) A New Staphylococcal Anti-Inflammatory Protein That Antagonizes the Formyl Peptide Receptor-Like 1. *The Journal of Immunology*, 177, 8017-8026.
38. Prat, C., Haas, P.-J., Bestebroer, J., de Haas, C.J.C., van Strijp, J.A.G. and van Kessel, K.P.M. (2009) A homolog of formyl peptide receptor-like 1 (FPRL1) inhibitor from *Staphylococcus aureus* (FPRL1 inhibitory protein) that inhibits FPRL1 and FPR. *Journal of Immunology (Baltimore, Md.: 1950)*, 183, 6569-6578.
39. Resch, A., Rosenstein, R., Nerz, C. and Götz, F. (2005) Differential gene expression profiling of *Staphylococcus aureus* cultivated under biofilm and planktonic conditions. *Applied and Environmental Microbiology*, 71, 2663-2676.
40. Romilly, C., Lays, C., Tomasini, A., Caldelari, I., Benito, Y., Hammann, P., Geissmann, T., Boisset, S., Romby, P. and Vandenesch, F. (2014) A Non-Coding RNA Promotes Bacterial Persistence and Decreases Virulence by Regulating a Regulator in *Staphylococcus aureus*. *PLoS Pathogens*, 10.
41. Said, N., Rieder, R., Hurwitz, R., Deckert, J., Urlaub, H. and Vogel, J. (2009) In vivo expression and purification of aptamer-tagged small RNA regulators. *Nucleic Acids Research*, 37, e133.
42. Sharma, E., Sterne-Weiler, T., O'Hanlon, D. and Blencowe, B.J. (2016) Global Mapping of Human RNA-RNA Interactions. *Molecular Cell*, 62, 618-626.
43. Stemerding, A.M., Köhl, J., Pandey, M.K., Kuipers, A., Leusen, J.H., Boross, P., Nederend, M., Vidarsson, G., Weersink, A.Y.L., van de Winkel, J.G.J. *et al.* (2013) *Staphylococcus aureus* formyl peptide receptor-like 1 inhibitor (FLIPr) and its

homologue FLIPr-like are potent FcγR antagonists that inhibit IgG-mediated effector functions. *Journal of Immunology (Baltimore, Md.: 1950)*, 191, 353-362.

44. Storz, G., Vogel, J. and Wassarman, K.M. (2011) Regulation by Small RNAs in Bacteria: Expanding Frontiers. *Molecular cell*, 43, 880-891.
45. Surmann, K., Michalik, S., Hildebrandt, P., Gierok, P., Depke, M., Brinkmann, L., Bernhardt, J., Salazar, M.G., Sun, Z., Shteynberg, D. *et al.* (2014) Comparative proteome analysis reveals conserved and specific adaptation patterns of *Staphylococcus aureus* after internalization by different types of human non-professional phagocytic host cells. *Frontiers in Microbiology*, 5, 392.
46. Thorvaldsdóttir, H., Robinson, J.T. and Mesirov, J.P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, 14, 178-192.
47. Tomasini, A., François, P., Howden, B.P., Fechter, P., Romby, P. and Caldelari, I. (2014) The importance of regulatory RNAs in *Staphylococcus aureus*. *Infection, Genetics and Evolution*, 21, 616-626.
48. Tree, Jai J., Granneman, S., McAteer, Sean P., Tollervey, D. and Gally, David L. (2014) Identification of Bacteriophage-Encoded Anti-sRNAs in Pathogenic *Escherichia coli*. *Molecular Cell*, 55, 199-213.
49. Udekwu, K.I., Darfeuille, F., Vogel, J., Reimegård, J., Holmqvist, E. and Wagner, E.G.H. (2005) Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. *Genes & Development*, 19, 2355-2366.
50. Updegrove, T.B., Shabalina, S.A. and Storz, G. (2015) How do base-pairing small RNAs evolve? *FEMS Microbiology Reviews*, 39, 379-391.
51. Varet, H., Brillet-Guéguen, L., Coppée, J.-Y. and Dillies, M.-A. (2016) SARTools: A DESeq2- and EdgeR-Based R Pipeline for Comprehensive Differential Analysis of RNA-Seq Data. *PLoS ONE*, 11.
52. Vogel, J. and Wagner, E.G.H. (2007) Target identification of small noncoding RNAs in bacteria. *Current Opinion in Microbiology*, 10, 262-270.
53. Wagner, E.G.H. and Romby, P. (2015) In Theodore Friedmann, J. C. D. a. S. F. G. (ed.), *Advances in Genetics*. Academic Press, Vol. 90, pp. 133-208.
54. Wang, J., Rennie, W., Liu, C., Carmack, C.S., Prévost, K., Caron, M.-P., Massé, E., Ding, Y. and Wade, J.T. (2015) Identification of bacterial sRNA regulatory targets using ribosome profiling. *Nucleic Acids Research*, 43, 10308-10320.

TABLES

Table 1: List of RNA sequenced by MAPS using RsaA as bait

Id	Gene	Product	Fold change	pvalue	padj
HG001_00625	mgrA	HTH-type transcriptional regulator MgrA	14,641	7,15E-33	1,94E-29
HG001_00516	-	RsaA	13,901	3,64E-31	4,95E-28
HG001_02338	ssaA2_3	Staphylococcal secretory antigen ssaA2 precursor	5,162	1,14E-12	1,03E-09
HG001_01977	-	Bacterial SH3 domain protein	4,024	1,63E-09	1,11E-06
HG001_02583	ssaA2_4	Staphylococcal secretory antigen ssaA2 precursor	3,868	4,16E-09	2,26E-06
HG001_02334	ssaA2_2	Staphylococcal secretory antigen ssaA2 precursor	3,455	8,21E-09	3,72E-06
HG001_01615	cycA_1	D-serine/D-alanine/glycine transporter	3,262	9,99E-08	3,88E-05
HG001_01092	rplS	50S ribosomal protein L19	2,882	1,59E-06	0,00054126
HG001_01204	-	hypothetical protein	2,889	1,93E-06	0,00058144
HG001_01150	glpF	Glycerol uptake facilitator protein	2,739	5,91E-06	0,00160586
HG001_01863	-	hypothetical protein	2,654	1,84E-05	0,00454999
HG001_01003	flr	FPRL1 inhibitory protein precursor	2,605	2,89E-05	0,00655166
HG001_01295	norB_4	Quinolone resistance protein NorB	2,581	3,96E-05	0,00827524
HG001_02606	ssaA_2	Staphylococcal secretory antigen SsaA precursor	2,413	0,0001149	0,02230708

Table 2: Differential proteomic analysis

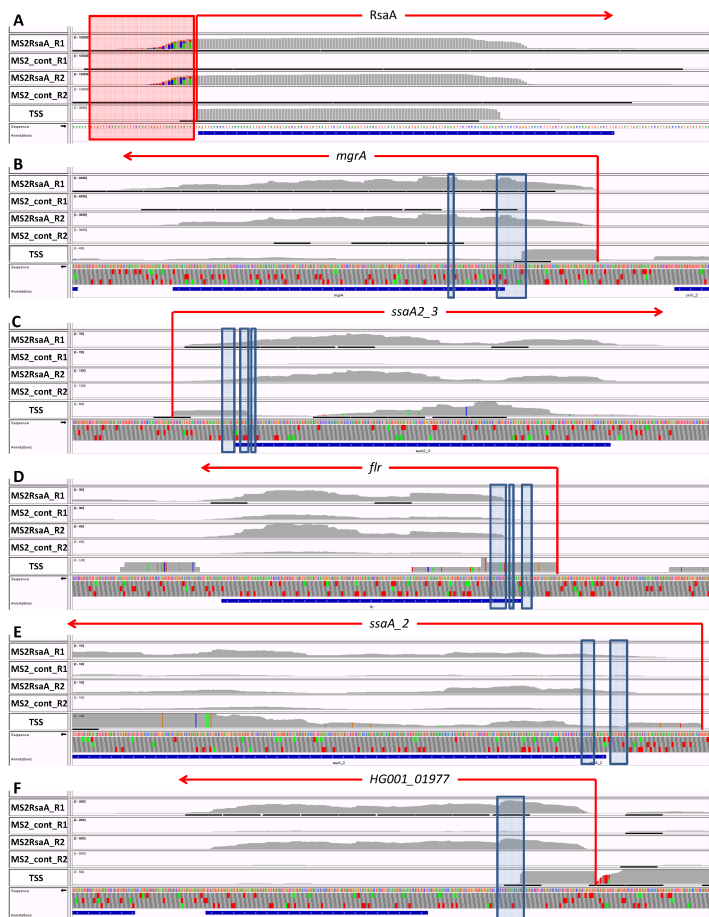
Cytoplasmic proteins

Protein Set	SPECTRAL COUNT				MS1 - LABEL FREE			
	DelRsaA/WT (1)	DelRsaA/Compl (2)	Ratio	pvalue	DelRsaA/WT (1)	DelRsaA/Compl (2)	Ratio	pvalue
mgrA	1.62	0.02295	3.70	0.00084	2.19	0.00025	12.72	0.00001
Ebh	nd	1.00000	COMPL only	0.00132	1.14	0.00924	0.07	0.06313
sasG	WT only	0.11612	COMPL only	0.00010	0.13	0.12445	0.00	0.04024
spa	WT only	0.12100	COMPL only	0.00219	0.01	0.37419	0.00	0.00039
SdrD	WT only	0.01417	COMPL only	0.00281	0.01	0.07905	0.00	0.00050
SdrC	WT only	0.00056	COMPL only	0.37390	0.01	0.11656	0.30	0.35361
ssaA2_3	1.06	0.28786	1.59	0.00069	1.09	0.73812	2.24	0.05426

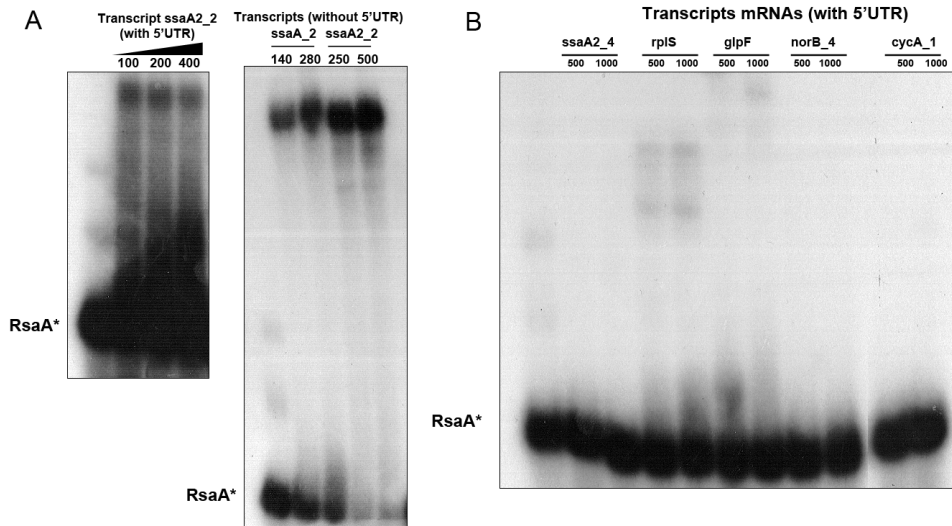
Secretome

Protein Set	SPECTRAL COUNT						MS1 LABEL-FREE						XICs MANUELS		
	DEL/WT	p-value	DEL/COMPL	p-value	COMPL/WT	p-value	DEL/WT	p-value	DEL/COMPL	p-value	COMPL/WT	p-value	DEL/WT	DEL/COMPL	COMPL/WT
ssaA2-3	1.76	0.00653	6.35	0.00058	0.28	0.00038	0.78	0.72608	13.80	0.28189	0.06	0.00053	2.14	14.60	0.15
SdrD	WT	0.00105	COMPL	0.00045	5.54	0.00107	0.44	0.00299	0.11	0.16200	4.15	0.21966	0.52	0.14	3.81
SdrC	WT	0.00335	COMPL	0.00033	1.82	0.02219	6.81	0.31236	4.42	0.34880	1.54	0.58579	0.35	0.23	1.52
ClfB	0.13	0.00246	0.05	0.00021	2.38	0.00302	1.76	0.28808	0.11	0.00256	15.44	0.00177	0.56	0.48	1.16
sasG	0.06	0.00031	0.02	0.00003	3.17	0.00011	0.63	0.52793	0.17	0.00792	3.76	0.02256	0.15	0.09	1.71
Spa			COMPL	0.03375	COMPL	0.03375	0.68	0.51458	0.06	0.14784	11.84	0.15781	0.59	0.10	5.89

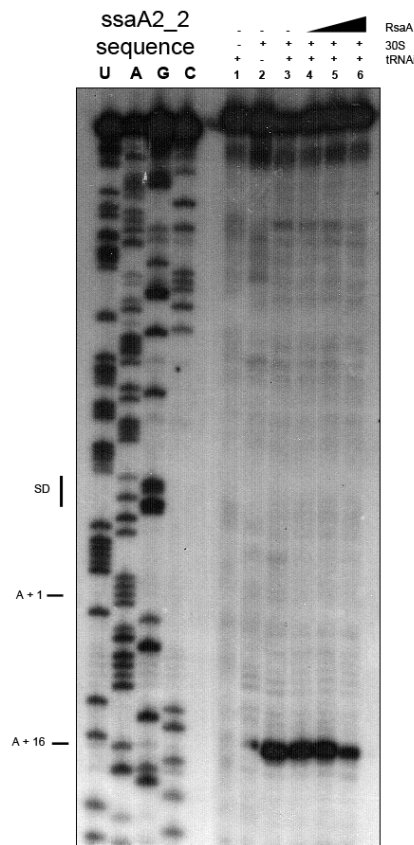
## SUPPLEMENTAL FIGURES



**Figure S1: Examples of mapped reads in RsaA and several mRNAs.** MAPS data visualized using IGV genome browser. For each gene the reads for the two MAPS experiments (MS2RsaA\_R1 or R2) and their controls (MS2\_cont\_R1 or R2) are visualized together with the Transcription Start Site (TSS) derived from Koch *et al.*, 2014. The light-red transparent box in front of RsaA is the MS2-tag. The regions of interaction with RsaA are highlighted with light-blue transparent boxes. (A) RsaA, (B) *mgrA*, (C) *ssaA2\_3*, (D) *flr* (E) *ssaA\_2* and (F) *HG001\_01977*. Note that RsaA has been described as expressed in a short and a long form (Geissmann *et al.*, 2009).



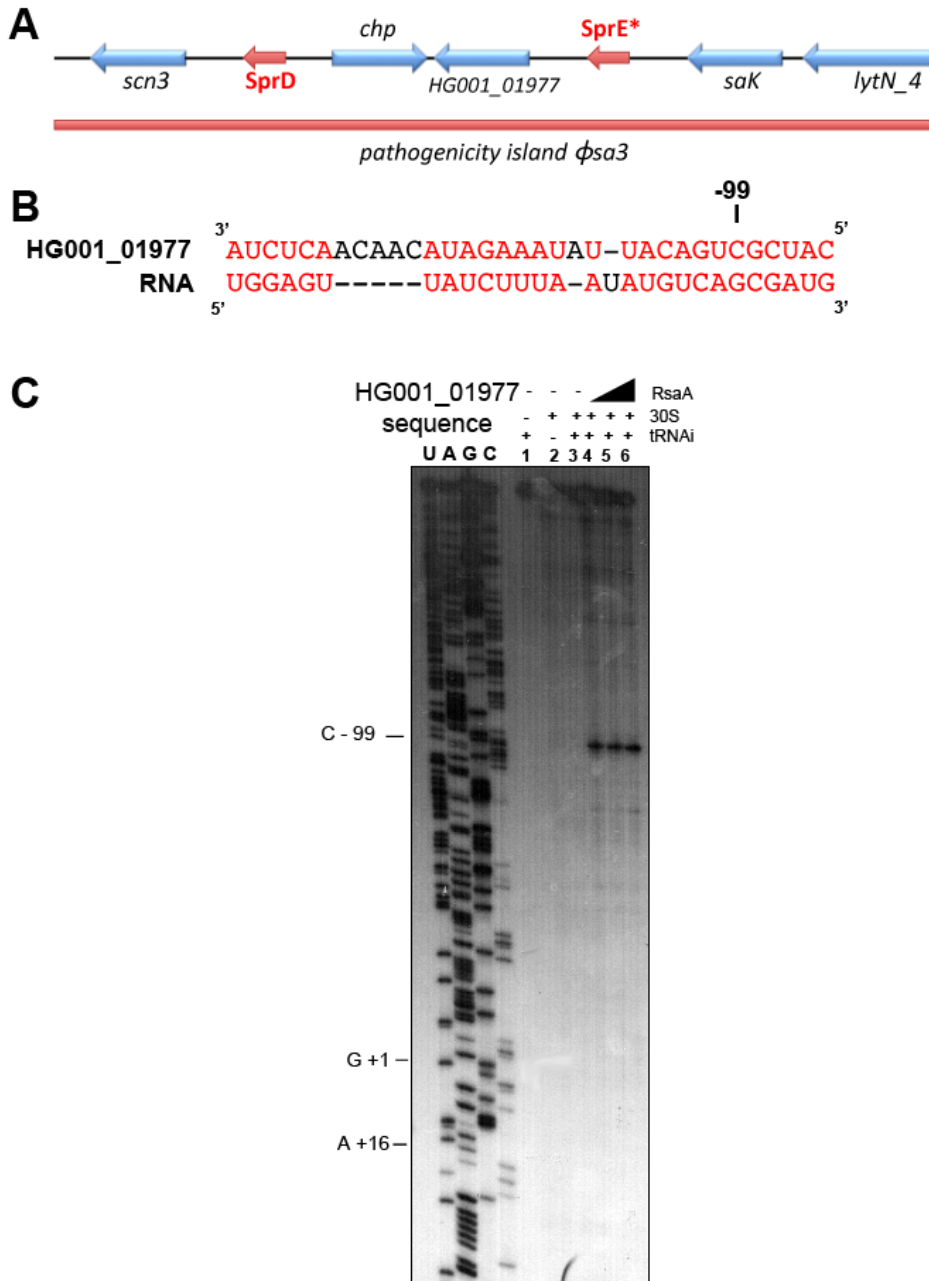
**Figure S2: Gel retardation assays on various mRNAs.** Experiments were performed on *ssaA2\_2* mRNA encompassing or not the 5' UTR (A) or *ssaA2\_4*, *rplS*, *glpF*, *norB\_4* and *cycA1* mRNA (B). The 5' end-labelled Wt RsaA was incubated with increasing concentrations (nM) of mRNAs.



**Figure S3 : Toe-print with *ssaA2\_2* mRNA.** Effect of RsaA on the formation of initiation ribosomal complex on *ssaA2\_2*. Lane 1 : control incubation of mRNA ; lane 2 : control incubation of mRNA with 30S ribosomal subunits, without tRNAi allowing



the formation of the initiation complex ; lane 3 : formation of the ribosomal initiation complex containing mRNA, the 30S subunits and tRNA<sub>i</sub> ; lane 4 to 6 : formation of the ribosomal initiation complex in presence of increasing concentrations of WT RsaA : 100, 200, 400 nM. Lanes U, A, G, C : sequencing ladders. The Shine-Dalgarno (SD) sequence, the start site of translation (A +1 of the AUG initiation codon) and the toe-printing signals (N+16) are indicated.



**Figure S4: Toe-print with with *HG001\_01977* mRNA.** (A) Genetic organization of the pathogenic island containing the *HG001\_01977* gene; sRNA are depicted in red (B) Prediction of basepairing interactions between RsaA and *HG001\_01977* mRNA.

Effect of RsaA on the formation of initiation ribosomal complex with *HG001\_01977*. Lane 1 : control incubation of mRNA ; lane 2 : control incubation of mRNA with 30S ribosomal subunits, without tRNA<sub>i</sub> allowing the formation of the initiation complex ; lane 3 : formation of the ribosomal initiation complex containing mRNA, the 30S subunits and tRNA<sub>i</sub> ; lane 4 to 6 : formation of the ribosomal initiation complex in presence of increasing concentrations of WT RsaA : 100, 200, 400 nM. Lanes U, A, G, C : sequencing ladders. The start site of translation (A +1 of the AUG initiation codon), the toe-printing signals (N+16) and the RT pausing (C-99) are indicated.

## SUPPLEMENTARY TABLES

**Table S1. Strains and plasmids used in this study**

	Relevant characteristics	Reference
<b><i>E. coli</i> strains</b>		
DC10B	dcm mutant ; used for cloning of plasmids which can be directly transformed in <i>S. aureus</i> cells	Monk <i>et al.</i> (2012)
<b><i>S. aureus</i> strains</b>		
RN HG001 (RN1)	Derivative of 8325-4, <i>rsbU</i> restored RN1, <i>agr</i> positive	Herbert <i>et al.</i> (2010)
LUG 1630	RN HG001 $\Delta$ <i>rsaA</i>	Romilly <i>et al.</i> (2014)
RN1 pCN51/P3	RN HG001 transformed with pCN51/P3	This study
RN1 $\Delta$ <i>rsaA</i> pCN51/P3	LUG 1630 transformed with pCN51/P3	This study
RN1 $\Delta$ <i>rsaA</i> pCN51/pRsaA-T2	LUG1630 transformed with pCN51/pRsaA-T2	Romilly <i>et al.</i> (2014)
RN1 pCN51/MS2	RN HG001 transformed with pCN51/MS2	This study
RN1 $\Delta$ <i>rsaA</i> pCN51/MS2-RsaA	LUG1630 transformed with pCN51/MS2-RsaA	This study
<b>Plasmids</b>		
pCN51	Shuttle vector	Charpentier <i>et al.</i> (2004)
pCN51/P3	pCN51 with P3 promoter	This study
pCN51/pRsaA-T2	pCN51 expressing RsaA (long form) under its own promoter	Romilly <i>et al.</i> (2014)
pCN51/MS2	pCN51 expressing MS2 tag under P3 promoter	This study
pCN51/MS2-RsaA	pCN51 expressing MS2-RsaA (long form) under P3 promoter	This study
pUC18-rsaA	<i>rsaA</i> (short form) under T7 promoter	Romilly <i>et al.</i> (2014)
pUC18-rsaA mutC	<i>rsaA</i> mutated in C-rich motif (short form) under T7 promoter	Romilly <i>et al.</i> (2014)

Table S2 : primers used in this study

Primer	strand	Sequence	Use
SpHI/P3	+	TTTGCATGCATACGTGGCAAAC	Plasmid construction
PstI/P3	-	CGCCTGCAGAGTTATATTAACATGC	Plasmid construction
MS2 For	+	CGCCTGCAGCGTACACCATCAGGGTACGTTTTTCGTACACCATCAGGGTACGTTT	Plasmid construction
MS2 Rev	-	CGCGGATCCAAACGTACCCTGATGGTGTACGAAAAACGTACCCTGATGGTGTACG	Plasmid construction
MS2-RsaA For	+	CGCCTGCAGCGTACACCATCAGGGTACGTTTTTCGTACACCATCAGGGTACGTTT	Plasmid construction
MS2-RsaA Rev	-	ACCATAAACAA	Plasmid construction
T7 RsaA For	+	CGCGGATCCAATAAAAAAATCCAAGCTTATCGGT	Plasmid construction
RsaA Rev	-	TAATACGACTCACTATAGGGGGTTAACATTACAAAATTGTATAGAG	<i>in vitro</i> transcription of <i>rsaA</i> RNA
T7 MS2 For	+	CGCGGATCCAAAAAAGTACACTTTGCTCATAGC	<i>in vitro</i> transcription of <i>rsaA</i> RNA
T7 MgrA For	+	TAATACGACTCACTATAGGGGGCGTACACCATCAGGGTACG	<i>in vitro</i> transcription of MS2-modified RNA
FLAG-MgrA Rev	-	AAAGGATCCTAATACGACTCACTATAGGGGCATACTTTAATTTTAGCATGGC	<i>in vitro</i> transcription of FLAG-tagged <i>mgrA</i> mRNA
T7 UTR <i>ssaA</i> _2 For	+	CCGGAATTCCTTACTTGTTCATCGTCGTCCTTGATGCTTTTTCTTTGTTTCATCAAATGC	<i>in vitro</i> transcription of FLAG-tagged <i>mgrA</i> mRNA
<i>ssaA</i> _2 CDS Rev	-	TTGTCGCAAAACTATAAAAGT	<i>in vitro</i> transcription of 5'UTR <i>ssaA</i> _2 mRNA
T7 UTR <i>ssaA</i> 2_2 For	+	TATGATTGAGCGTCATTAGA	<i>in vitro</i> transcription of 5'UTR <i>ssaA</i> _2 mRNA
<i>ssaA</i> 2_2 CDS Rev	-	TAATACGACTCACTATAGGGGGTTCGTTAATTACGAGAGAT	<i>in vitro</i> transcription of 5'UTR <i>ssaA</i> 2_2 mRNA
T7 UTR <i>ssaA</i> 2_3 For	+	TTGGGTCGTTTGGATTATAA	<i>in vitro</i> transcription of 5'UTR <i>ssaA</i> 2_2 mRNA
<i>ssaA</i> 2_3 CDS Rev	-	TAATACGACTCACTATAGGGGGTCATCAATTCGAGATGTGA	<i>in vitro</i> transcription of 5'UTR <i>ssaA</i> 2_3 mRNA
T7 UTR <i>ssaA</i> 2_4 For	+	GTTTGAGTCGTGTTGTAGTTA	<i>in vitro</i> transcription of 5'UTR <i>ssaA</i> 2_3 mRNA
<i>ssaA</i> 2_4 CDS Rev	-	TAATACGACTCACTATAGGGGGTTGAATAAAGAAAAAATTA	<i>in vitro</i> transcription of 5'UTR <i>ssaA</i> 2_4 mRNA
T7 UTR <i>flr</i> For	+	TACTATAATAATTAAGTACGCGC	<i>in vitro</i> transcription of 5'UTR <i>ssaA</i> 2_4 mRNA
<i>flr</i> CDS Rev	-	TAATACGACTCACTATAGGGGGTACTCGCATTAGTAAGTAC	<i>in vitro</i> transcription of 5'UTR <i>flr</i> mRNA
		TTGCGATTTCTAAACCTTTCCAT	<i>in vitro</i> transcription of 5'UTR <i>flr</i> mRNA

T7 UTR SH3 For	+	<u>TAATACGACTCACTATAGGG</u> ACATGCATCAACTATTTACA	<i>in vitro</i> transcription of 5'UTR <i>sh3 domain protein</i> mRNA
SH3 CDS Rev	-	AATTCTTGAATTAGTTGAAT	<i>in vitro</i> transcription of 5'UTR <i>sh3 domain protein</i> mRNA
T7 UTR <i>rpIS</i> For	+	<u>TAATACGACTCACTATAGGG</u> GTGTGGAAACACTAAAATCAC	<i>in vitro</i> transcription of 5'UTR <i>rpIS</i> mRNA
<i>rpIS</i> CDS Rev	-	GCTCACGAGTACCCTCAATGA	<i>in vitro</i> transcription of 5'UTR <i>rpIS</i> mRNA
T7 UTR <i>glpF</i> For	+	<u>TAATACGACTCACTATAGGG</u> GATTGGTATTTCTAGTTTTTC	<i>in vitro</i> transcription of 5'UTR <i>glpF</i> mRNA
<i>glpF</i> CDS Rev	-	ATCCAGCTGTGATGACAATCC	<i>in vitro</i> transcription of 5'UTR <i>glpF</i> mRNA
T7 UTR <i>cycA_1</i> For	+	<u>TAATACGACTCACTATAGGG</u> ATCTATATTTAAGGGAGAT	<i>in vitro</i> transcription of 5'UTR <i>cycA_1</i> mRNA
<i>cycA_1</i> CDS Rev	-	GAATTGAAGGGCCGGTTAATG	<i>in vitro</i> transcription of 5'UTR <i>cycA_1</i> mRNA
T7 UTR <i>norB_4</i> For	+	<u>TAATACGACTCACTATAGGG</u> GGAACAGACAAATTTCTGTT	<i>in vitro</i> transcription of 5'UTR <i>norB_4</i> mRNA
<i>norB_4</i> CDS Rev	-	CTTCAAGTATTGGTACAACA	<i>in vitro</i> transcription of 5'UTR <i>norB_4</i> mRNA

Underlined are the enzyme restriction sites. The MS2 sequence is highlighted in yellow, the T7 promoter in green, the FLAG sequence in light blue.

Note : les **tableaux S3, S4 et S6** ne sont pas présentés ici par souci de place et de confort de lecture.

## Partie 3 : Recherche et caractérisation de la fonction de RsaC

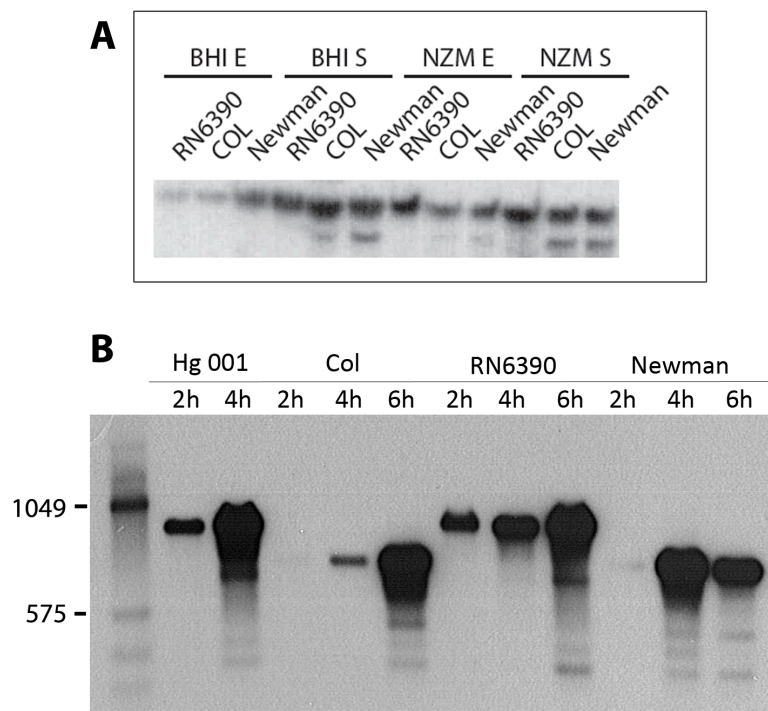
Ayant validé la preuve de concept de l'approche MAPS avec RsaA, j'ai généralisé cette méthode à un second ARNnc aux caractéristiques particulières, RsaC. Identifié en 2009 au cours par Geissmann *et al.*, la fonction de cet ARN n'avait pas encore été étudiée quand je suis arrivé au laboratoire. Tout au plus Thomas Geissmann avait remarqué une surexpression de cet ARN lorsque les cellules de *S. aureus* étaient soumises à un choc froid. Avant la mise au point du MAPS sur RsaC, j'ai caractérisé la fonction de cet ARN, par une approche plus classique impliquant l'analyse phénotypique d'une souche mutée au niveau du gène *rsaC*.

### 1. La taille de RsaC varie en fonction des souches à cause de régions répétées

RsaC a été identifié comme un ARNnc de 544 nucléotides, qui est exprimé dès la phase exponentielle de croissance (2h de culture) et qui s'accumule durant la phase stationnaire (6h), que la culture soit faite en milieu riche (BHI) ou pauvre (NZM).

Ayant plusieurs souches de *S. aureus* à disposition, j'ai vérifié à nouveau l'expression du gène *rsaC* au cours de la phase de croissance. J'ai préparé pour ce faire des ARN totaux récoltés à 2h, 4h et 6h de croissance des souches HG001, RN6390, COL et Newman et regardé l'expression de RsaC par Northern blot en utilisant une sonde DIG-ARN complémentaire de cet ARNnc. Les résultats présentés en **figure 16B** confirment l'accumulation de l'ARN en phase stationnaire de croissance, excepté pour la souche Newman où une accumulation plus importante de l'ARN à 4h est visualisée, comparée à 6h de croissance. Mais le point marquant de ces résultats est que la longueur de RsaC varie en fonction de la souche. Ainsi pour les souches Newman et COL nous observons une taille qu'on peut estimer, selon l'échelle de poids moléculaire sur la gauche, à environ 700 nucléotides. Pour les souches HG001 et RN6390 en revanche, cette taille est plus importante, de l'ordre de 1000 nucléotides (1kb). Dans tous les cas, la taille de RsaC est supérieure à 544 nucléotides, contrairement à ce qui avait été déterminé dans

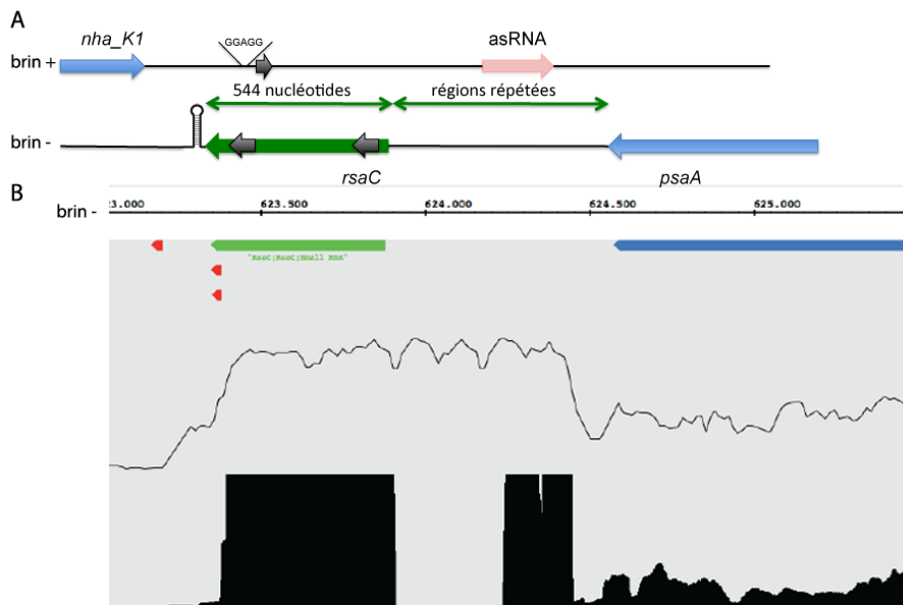
l'étude initiale de Geissmann *et al.* (2009). Par ailleurs, les expériences de « Northern blot » réalisés à l'époque (**Figure 17A**), ne montraient pas de différence de taille de l'ARN entre les souches RN6390, Newman, et COL. Seulement un produit plus court apparaissait dans les souches Newman et COL. La seule différence entre les deux expériences est que la détection de l'ARN RsaC par northern blot a été réalisée avec un oligonucléotide marqué radioactivement en 2009 par T. Geissmann alors que j'ai effectué la détection avec une sonde DIG-ARN d'une taille de 200 nucléotides, qui cible la région des 544 nucléotides déterminés par Geissmann *et al.* (2009).



**Figure 17 : Différences de taille pour RsaC en fonction des souches de *S. aureus*** - A : Northern blot pour visualiser l'expression de RsaC dans les souches RN6390, Newman et COL en phase exponentielle (E) et stationnaire (S) de croissance, en milieu riche (BHI) et pauvre (NZM). Tirée des données supplémentaires de Geissman *et al.*, 2009. B : Northern blot pour visualiser l'expression de RsaC dans les souches HG001, COL, RN6390 et Newman, à 2h, 4h et 6h de croissance.

Pour comprendre cette différence de taille observée par Northern blot (**figure 17B**), il faut regarder plus attentivement le locus de *rsaC* chez *S. aureus* (l'ARN est conservé uniquement chez cette espèce) qui présente bien des particularités. Le gène *rsaC* se situe sur le brin de polarité négative de l'ADN de *S. aureus*. Comme le montre la **figure 18A** il est situé entre les gènes *psaA* situé en amont qui code un transporteur du manganèse de type ABC et qui est en opéron avec les gènes *mntB* et *znuC\_1*, tous deux impliqués dans le transport du manganèse et du zinc, et le gène *nhaK\_1* situé sur le brin de polarité positive, qui code une pompe à protons via un mécanisme antiport. Entre le début de

RsaC tel qu'il a été identifié par Geissman *et al.* (2009) et la fin du gène *psaA* se trouvent des régions répétées. Ces régions répétées peuvent être observées expérimentalement par des approches de RNA seq et *tiling array* telles qu'exposées sur la **figure 18B**. Ce graphique résulte des données de séquençages du Dr. Alejandro Toledo à partir d'un extrait d'ARN totaux. Les *reads* obtenus après le séquençage des ARN ont été alignés sur la séquence connue du génome de RN8325-4, et on peut voir qu'il y a un *gap* de reads qui vient s'intercaler juste après le début de transcription de l'ARN RsaC. Ce *gap* est provoqué par la présence de ces régions répétées, qui empêche l'alignement correct des *reads* puisqu'à cause des répétitions, un même *read* pourrait s'aligner à plusieurs endroits de cette région. Après comparaison du locus de *rsaC* entre différentes souches de *S. aureus*, il s'avère que ce sont ces régions répétées, et surtout le nombre de répétitions, qui expliquent les différences de taille du transcrit RsaC.

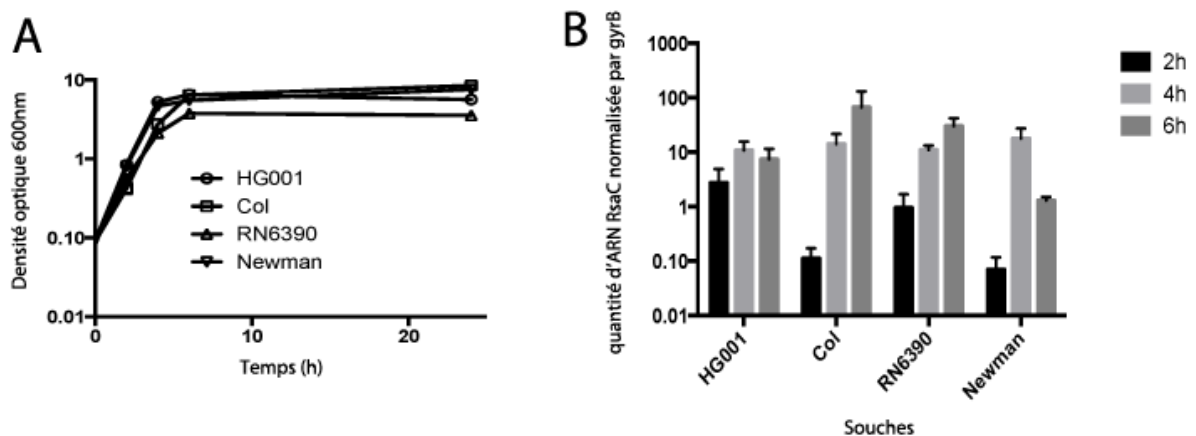


**Figure 18 : Le locus de RsaC présente plusieurs régions répétées** – A : représentation schématique du locus de *rsaC* chez *S. aureus* HG001. Le gène *rsaC* est indiqué en vert – la taille correspond à celle déterminée par Geissmann *et al.* Les régions répétées sont indiquées par la double flèche. Les petites flèches pleines noires indiquent des ORFs potentielles. B : données de *tiling array* et de RNAseq sur le locus de RsaC (les 544 nucléotides sont indiqués en vert selon les données du génome de RN 8325-4). La hauteur des zones pleines de l'ARN-seq (tout en bas) est proportionnelle au nombre de *reads* obtenu pour la séquence. Le *gap* entre les deux zones élevées indique la présence des régions répétées.

Le locus de *rsaC* présente encore deux caractéristiques qui méritent notre attention. Le premier est l'existence d'un ARN anti-sens en amont de RsaC. Une expérience d'immunoprécipitation effectuée avec une version mutée de la RNase III (qui reconnaît les ARN en double brin, et donc les duplexes formés par l'appariement d'un ARNnc avec son ARNm cible) qui est capable de lier les ARN mais qui ne peut plus réaliser le clivage a été menée par Lioliou *et al* (2012). Les ARN qui ont été co-immunoprécipités avec la RNase III ont

ensuite été séquencés et plusieurs *reads* ont été alignés à cet endroit du locus de *rsaC*, qui serait donc un potentiel ARN anti-sens dont j'ai voulu vérifier l'expression. Le second point est l'existence d'une ORF potentielle sur le brin opposé à RsaC puisqu'une séquence SD est prédite. Cette ORF putative très courte (19 codons) pourrait être celle d'un peptide de type PSM.

L'expression de RsaC a également été vérifiée par RT-qPCR par l'équipe de F. Vandenesch et comme le montre la **figure 19** ci-dessous, on observe des profils d'expression proches de ceux constatés par mes expériences de Northern, avec une accumulation de RsaC plutôt vers 4h pour les souches HG001 et Newman alors que cette accumulation est plus importante à 6h dans les souches COL et RN6390. Nous avons vérifié que les courbes de croissance des quatre souches sont identiques dans nos conditions de culture (**Figure 19A**).



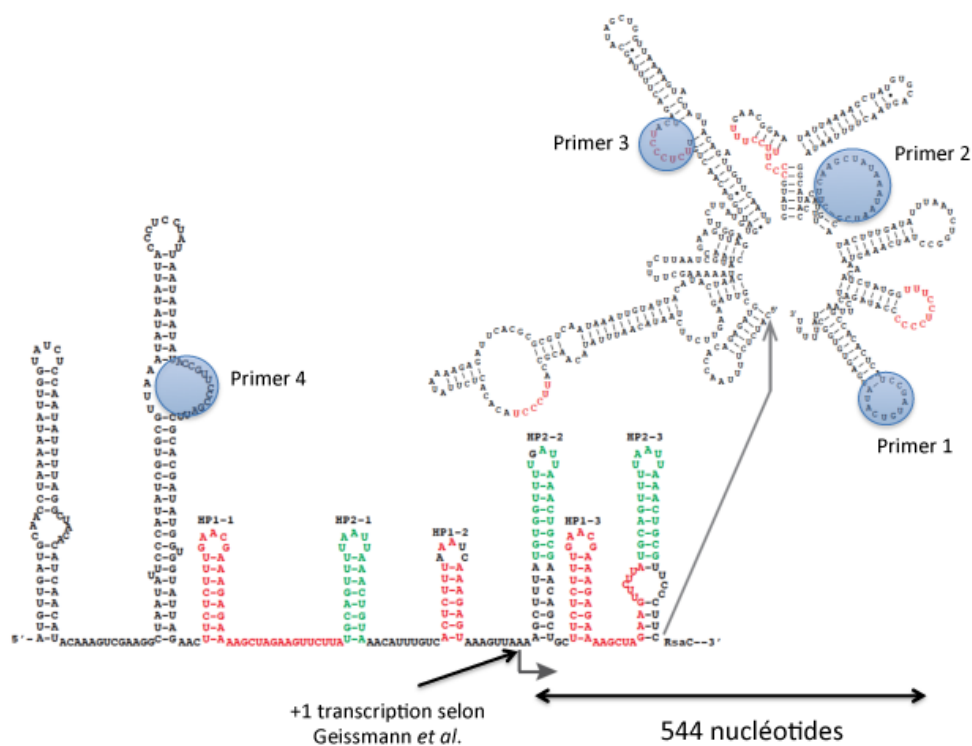
**Figure 19 : Expression de RsaC dans différentes souches de *S. aureus*** – A : courbes de croissance des différentes souches utilisées en milieu BHI à 37°C. B : Expression de RsaC chez les mêmes souches mesurée par RT-qPCR à différents temps de croissance, et normalisée par rapport à l'expression du gène de ménage *gyrB*.

## 2. Détermination du démarrage de transcription de RsaC

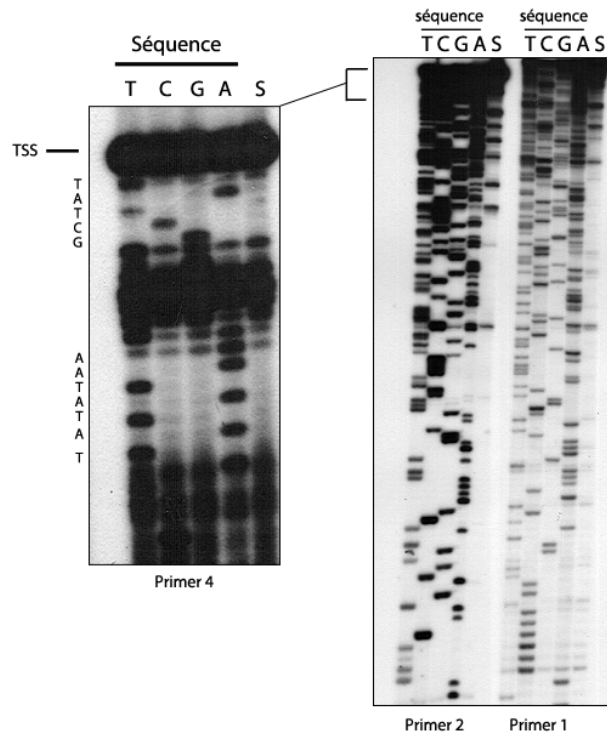
Avant de déléter le gène *rsaC* pour étudier sa fonction, il me fallait connaître les bornes du gène *rsaC*. A cause de la divergence des résultats concernant la taille de RsaC, j'ai déterminé l'extrémité 5' de l'ARNm par une approche de *primer extension* qui n'avait pas été utilisée par Thomas Geissmann. Le principe est le suivant : un oligonucléotide radio-marqué a été utilisé comme amorce pour la transcription inverse à partir d'ARN total extrait à 6h de croissance. Cet oligonucléotide est complémentaire au transcrit RsaC. La transcriptase inverse (RT) synthétise à partir de cette amorce un ADNc. L'arrêt final



correspond en général au +1 de transcription (TSS pour *transcription start site*). Il peut aussi correspondre à une coupure de l'ARN issue d'une maturation ou d'une dégradation. En parallèle un séquençage de l'ARN provenant du même extrait brut est réalisé pour déterminer à quelle base correspond l'arrêt de la RT. Ces expériences de *primer extension* ont été réalisées en collaboration avec Zong-Fu Wu, un chercheur chinois qui a passé un an au laboratoire et qui a participé au projet RsaC. Plusieurs oligonucléotides ont été nécessaires pour déterminer correctement la borne en 5' de RsaC. Ceux-ci sont listés dans la partie Matériel et Méthodes. Les trois premiers s'hybrident dans la région 3' de RsaC (celle précédemment déterminée) et nous sommes remontés de plus en plus dans la partie 5' au fur et à mesure des résultats des expériences de *primer extension*. La présence des régions répétées a posé également un problème et il a fallu sélectionner une quatrième amorce se situant dans les rares segments de séquence unique situés entre les régions répétées (**figure 20**).



**Figure 20 : Stratégie de détermination du TSS de RsaC chez *S. aureus* HG001 (RN1)** – Les 544 nucléotides déterminés par Geissmann *et al.* 2009 comprennent une partie des régions répétées qui sont ici indiquées en vert et en rouge (les structures secondaires ainsi que les régions répétées montrées ici sont uniquement prédites). Le TSS de ces mêmes travaux est indiqué. Les régions en simple brin qui ont été choisies pour la synthèses des différentes amorces pour la *primer extension* sont indiquées par les cercles bleus. Dans la partie 3', les séquences en rouge indiquent les motifs riches en C de l'ARN.



**Figure 21 : Expérience de primer extension pour déterminer le TSS de RsaC** – Les numéros des primers utilisés se rapportent à la nomenclature de la figure 18. Chacune des réactions a été réalisée sur les ARN totaux de *S. aureus* HG001. S : réaction de rétro-transcription utilisant les 4 dNTPs. T, C, G, A : réactions de séquençage spécifique pour chaque nucléotide.

Les résultats de ces expériences (montrés en **figure 21**) combinés à l'analyse du génome de *S. aureus* HG001 m'ont permis de définir correctement la borne en 5' du gène *rsaC* dans cette souche. La taille finale de RsaC obtenue est de 1117 nucléotides. Il se peut que les expériences de RACE initialement effectuées par T. Geissmann ont été perturbées par les régions répétées.

Je peux à présent avoir une vision globale du locus de *rsaC*. La borne en 5' se trouve à priori à 113 nucléotides en aval de la fin du gène *psaA*. Le promoteur du gène pourrait donc se trouver dans la région intergénique. Je n'exclue pas que RsaC soit transcrit en opéron avec *psaA* et que celui-ci soit ensuite mûri par une RNase spécifique. Sur la **figure 22** ci-après, on remarque que l'on peut distinguer deux grands domaines de RsaC : la région 3' qui contient plusieurs motifs riches en C et la région 5' qui comporte les répétitions. Le motif répété quatre fois et demi (dans la souche HG001) est constitué de deux tiges-boucles délimitées en vert et en rouge (**Figure 22**). Ce sont le nombre de ces répétitions qui varient dans les souches de *S. aureus*, ce qui explique la différence de taille de l'ARN que j'ai observé, telle que décrite dans la **figure 17B**.



Figure 22 : Séquence finale de RsaC chez *S. aureus* HG001 – Cette représentation permet très bien de voir les régions répétées, avec quatre répétitions et demi du même motif. Les séquences riches en C (en rouge) sont les mêmes qu’indiqués sur les figures 12 et 18.

### 3. Détermination de la structure secondaire de RsaC

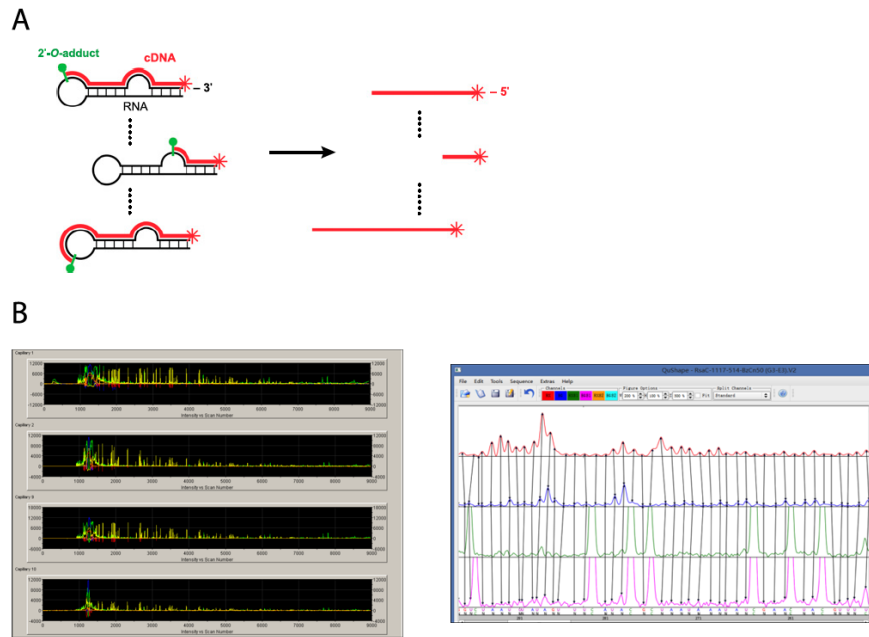
Jusqu’à présent, la structure secondaire de RsaC disponible était une structure prédite par bio-informatique. Connaissant la véritable longueur et séquence de RsaC, en collaboration avec le Dr Zong-Fu WU, la structure secondaire de RsaC a été déterminée par la technique de SHAPE (*selective 2’-hydroxyl acylation analyzed by primer extension*). Deux versions de RsaC ont été transcrites *in vitro* par la T7 RNA polymérase: la version raccourcie de RsaC de 544 nucléotides telle que déterminée par Geissmann *et al.* (2009) et la version complète de l’ARN longue de 1117 nucléotides.

Le SHAPE est basée sur l’utilisation d’un réactif chimique qui modifie chaque ribose de l’ARN. Cette réaction est couplée à la méthode de *primer extension* pour détecter les modifications dans l’ARN. La technique de cartographie en solution SHAPE se déroule en trois étapes :

- 1 , l’ARN est traité avec un réactif électrophile qui réagit avec les groupements hydroxyles (2’-OH) des riboses de l’ARN. Ce réactif (dans notre cas, le benzoyl cyanide [Sigma]) est sensible à la flexibilité du résidu et modifiera plus fortement

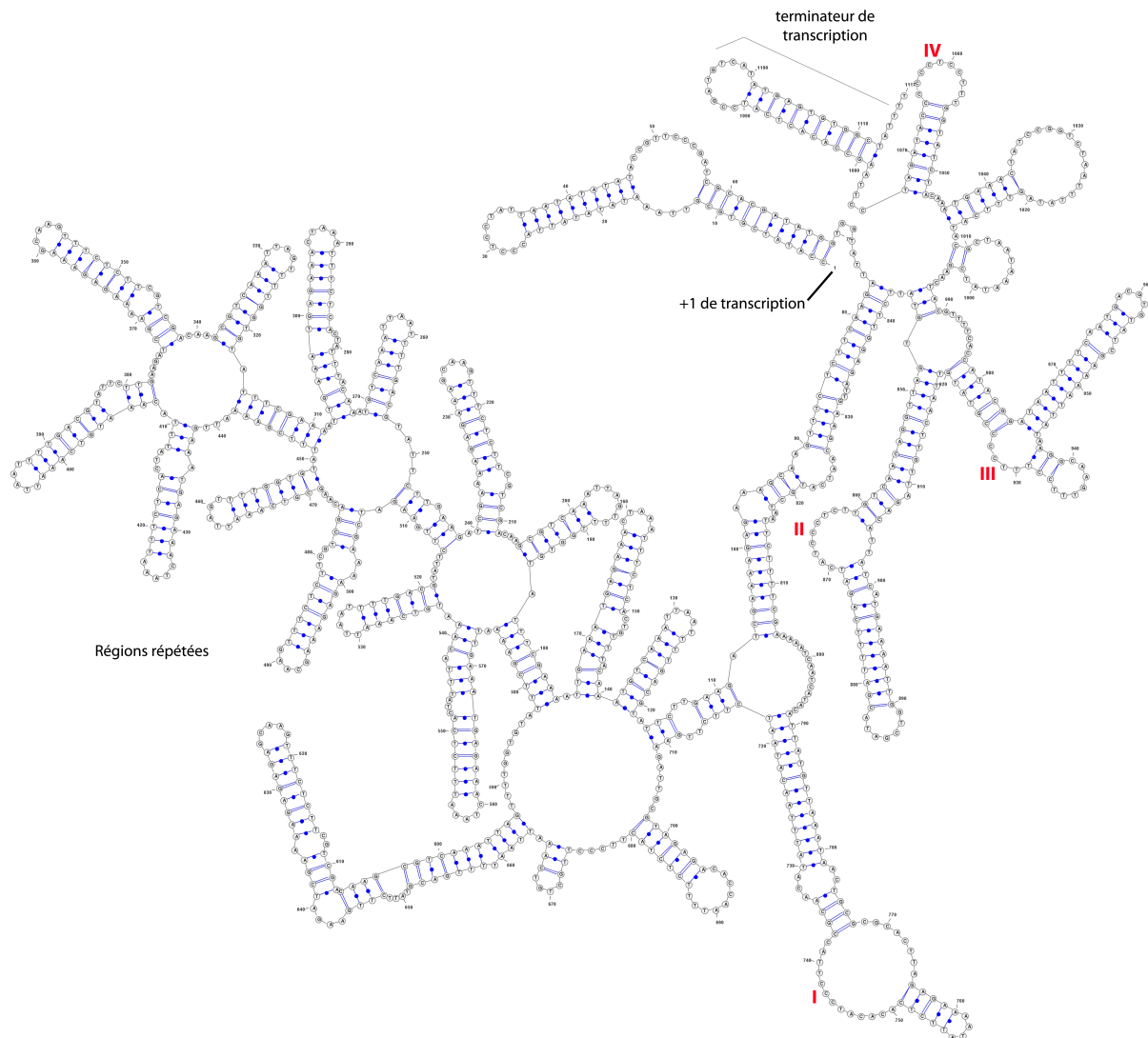
les riboses des nucléotides situés dans des régions en simple brin. La quantité du réactif utilisée est telle qu'au plus une modification aura lieu statistiquement par molécule. La réaction est arrêtée spontanément puisque la durée de demie-vie du réactif est rapide dans l'eau (dans notre cas, la réaction ne dure qu'une minute).

- 2, les sites de modification de l'ARN par ce groupement sont identifiés par RT avec une amorce ADN marquée avec un fluorophore en 5'. La modification étant encombrante suffit pour arrêter la RT. En résulte un certain nombre de cDNA de longueurs différentes, dont la taille reflète la position du nucléotide qui a été modifié. En parallèle, un contrôle d'incubation est réalisé où l'ARN est traité en absence du SHAPE. Une séquence de l'ARN doit être aussi effectuée pour déterminer la position des modifications. Les amorces utilisées pour chaque condition sont marquées avec un fluorophore différent, de façon à pouvoir discriminer les échantillons.
- 3, les produits de la rétro-transcription sont séparés par électrophorèse capillaire et la fluorescence est mesurée pour chaque position. Les résultats, présentés sous forme d'électrophorégramme sont alignés pour chaque expérience et permettent de visualiser la réactivité de chaque nucléotide à chaque position (Karabiber *et al.*, 2015). Zong-Fu Wu a combiné cette approche avec l'utilisation du logiciel d'analyse QuShape qui superpose les électrophorégrammes pour quantifier la réactivité de chaque nucléotide modifié (**figure 23**). Les quantifications sont ensuite utilisées comme contraintes pour établir une structure secondaire de l'ARN par le programme de prédiction RNA Structure. Plus le nucléotide sera modifié, plus il sera accessible au solvant et donc considéré comme étant non apparié.



**Figure 23 : Application du SHAPE pour déterminer la structure secondaire de RsaC** – A. Représentation schématique de l’approche de SHAPE. Le composé réactif (en vert) cible les nucléotides accessibles de l’ARN d’intérêt. Une amorce couplée à un fluorochrome (en rouge) est utilisée pour la réaction de rétro-transcription, qui est arrêtée par la modification. Les brins d’ADNc obtenus sont ensuite séparés en fonction de leur taille par électrophorèse capillaire. B. Exemples d’écrans d’analyses des résultats de SHAPE, soit directement par le séquenceur (à gauche), soit après acquisition des données via le logiciel QuShape (à droite).

Basé sur ces résultats, la structure secondaire de RsaC a été établie, comme le montre la **figure 24**, sans surprise, très complexe. La distinction entre la partie 5’ (qui comprend les régions répétées) et la partie 3’ (qui porte les motifs riches en C) est facilement reconnaissable. On peut également noter la présence d’une interaction étendue entre les régions 5’ et 3’ de l’ARN ce qui rapproche les deux extrémités du transcrit. Une telle organisation a été décrite pour l’ARNm *icaR* (de los Mozos et al., 2013). La région 3’ se termine par un terminateur clairement défini.

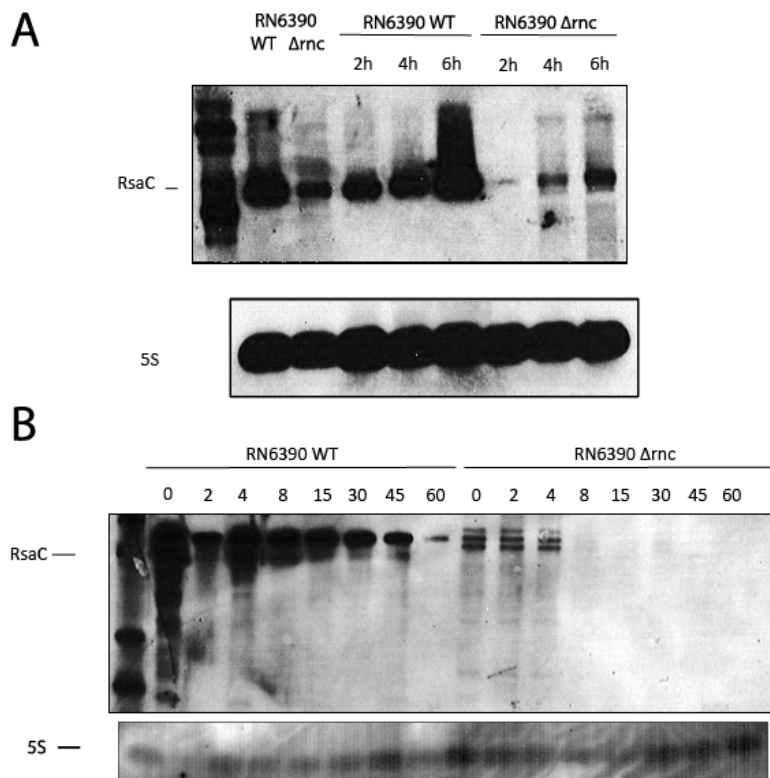


**Figure 24 : structure secondaire de RsaC** – Le TSS et le terminateur de transcription sont indiqués, tout comme les régions répétées. Les chiffres romains en rouge indiquent les quatre motifs riches en C présents sur l'ARN et déjà indiqués en figure 18.

#### 4. La stabilité de RsaC est dépendante de la RNase III

Je mentionnais précédemment la présence d'un possible ARN anti-sens à RsaC au niveau du locus du gène. L'expression de cet ARN a été recherchée dans la souche RN6390 de *S. aureus* (chez qui RsaC présente la même taille que dans la souche HG001 – voir **figure 17B**) déléetée du gène *rnc* qui code la RNase III (Lioliou *et al.*, 2012). Cette RNase est spécifique des régions ARN en double brin. Si l'ARN anti-sens à RsaC est véritablement exprimé, la RNase III devrait donc reconnaître la région d'appariement de cet ARN avec RsaC et cliver le duplex formé. En toute logique, l'expression de l'ARN anti-sens devrait être augmentée dans la souche qui n'exprime pas cette RNase. La souche  $\Delta rnc$  a également été utilisée pour vérifier la stabilité de RsaC au cours du temps. En effet, on

voit que la structure secondaire de l'ARN comporte de nombreuses régions en double brin ce qui en fait une cible potentielle pour cette RNase. Pour ce faire, j'ai visualisé par des expériences de Northern blot l'intégrité de l'ARN au cours du temps, dans la souche sauvage, et également dans la souche  $\Delta rnc$  après ajout dans le milieu de culture de la rifampicine. Cet antibiotique bloque la synthèse *de novo* d'ARN en inhibant l'activité de l'ARN polymérase bactérienne, et permet donc de suivre le temps de demi-vie des ARN et la rapidité de leur dégradation.

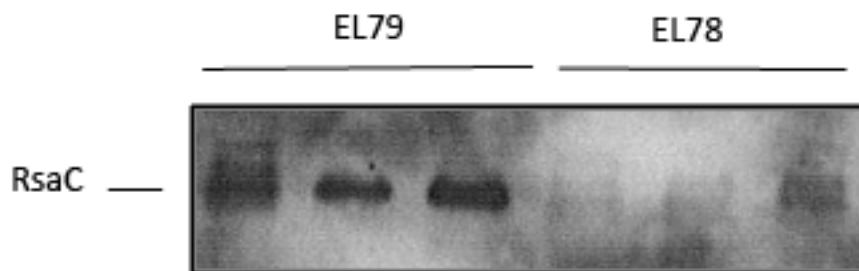


**Figure 25 : impact de la RNase III sur la stabilité de RsaC** – A. Northern blots dirigés contre RsaC et son ARN anti-sens (RsaCas) dans les souches RN6390 Wt et RN6390  $\Delta rnc$  ; les ARN totaux des deux souches ont été prélevés à 2h, 4h et 6h de culture en milieu BHI à 37°C. B. Mesure de la stabilité de RsaC par ajout de rifampicine dans le milieu de culture des souches RN6390 Wt et RN6390  $\Delta rnc$  et prélèvement des ARN totaux de 2 à 60 minutes après l'ajout de l'antibiotique.

Les résultats de ces expériences (**figure 25** ci-dessus) montrent qu'en absence de RNase III, RsaC présente une baisse très importante de sa stabilité. Le taux initial de RsaC est fortement diminué alors qu'un produit de taille plus grande est parfois visualisé (**figure 25A**) notamment chez la souche mutante à 4h et 6h de culture. Dans RN6390, RsaC s'accumule de façon importante après 6h de culture en milieu riche, et on peut voir une nette diminution de cette accumulation dans la souche  $\Delta rnc$  au même temps de culture. Après ajout de rifampicine, on voit que RsaC possède une demi-vie assez importante (estimée à 30 minutes bien que la surexposition du film m'empêche de

quantifier le signal correctement) puisqu'il est encore détecté après 60 minutes après l'ajout de l'antibiotique, alors que dans la souche  $\Delta rnc$ , la demi-vie est très fortement diminuée puisqu'au bout de 8 minutes, le pool de RsaC a intégralement disparu. On note également la visualisation de trois bandes de RsaC par northern blot, qui indiquent des ARN de taille très proches. Je n'ai pas d'indication précise sur la nature de ces ARN mais il pourrait s'agir de produits issus d'une dégradation provoquée par une autre RNase (i.e. RNase Y, RNase J1 ; Durand et al., 2016 ; annexe). Ces expériences montrent que la RNase III est impliquée dans la stabilité de RsaC. Elle pourrait participer à la maturation de l'ARN issu d'un produit de transcription plus grand incluant le gène en aval *psaA* en une forme plus stable.

Pour vérifier l'importance de l'activité enzymatique de la RNase III, j'ai également regardé l'expression de RsaC dans trois souches de RN6390  $\Delta rnc$  transformées avec le plasmide pCN51 sur lequel trois versions différentes du gène *rnc* ont été séparément incorporées : la version sauvage (souche EL79), ainsi qu'une autre version mutante (EL78) qui porte une mutation ponctuelle (l'acide aspartique en position 63 remplacé par une alanine) située dans le site catalytique de l'enzyme. Cette mutation abolit l'activité enzymatique de coupure tout en maintenant ses propriétés de fixation aux acides nucléiques (Lioliou *et al.*, 2012).



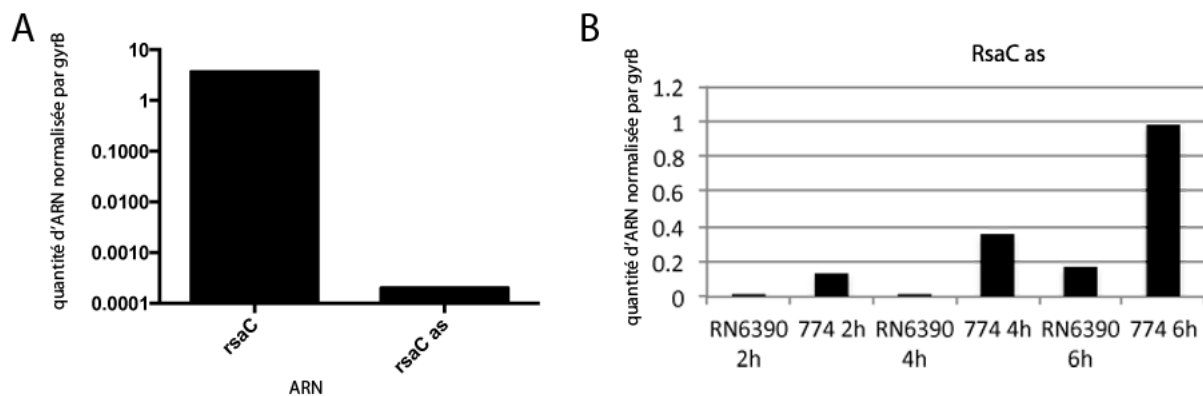
**Figure 26 : expression de RsaC chez des souches exprimant différentes versions de la RNase III** - Les ARN totaux ont été prélevés pour chacune des souches à 2h, 4h et 6h de culture en milieu BHI à 37°C. EL79 : enzyme sauvage ; EL78 : enzyme mutée D63A inactive ; EL80 : enzyme mutée E135A inactive.

Lorsqu'on compare l'expression de RsaC dans une souche qui exprime les versions mutées de la RNase III par rapport à l'enzyme sauvage, on constate une forte diminution du taux de l'ARN quelque soit le temps de culture. (**figure 26**). Je tiens néanmoins à préciser que je n'ai pas réalisé un contrôle sur l'ARNr 5S pour vérifier que les différents échantillons contenaient la même quantité d'ARN. Comme les deux RNases mutées sont



entravées dans leur action de clivage des ARN, c'est donc bien cette activité qui est requise pour stabiliser RsaC. En revanche, dans cette expérience, nous n'avons pas pu détecter de produit de transcription plus long que RsaC qui pourrait résulter d'un défaut de maturation. Le mécanisme d'action de la RNase III reste encore à élucider.

Une des possibilités serait l'intervention de l'ARN anti-sens pour cibler la maturation par la RNase III. C'est pourquoi nous avons suivi son taux par RT-qPCR et ceci de manière comparative avec RsaC, expérience réalisée en collaboration avec Karen Moreau de l'équipe de François Vandenesch (Lyon) avec qui j'ai collaboré pour ce projet de thèse.



**Figure 27 : quantification de RsaC et de RsaC anti-sens par QRT-PCR chez RN1** – A. Les ARN totaux ont été prélevés pour chacune des souches à 6h de culture en milieu BHI à 37°C. Les ARN RsaC et RsaC anti-sens ont été rétrotranscrits et ensuite amplifiés par PCR quantitative. La quantité de produits de PCR finale est proportionnelle à la quantité d'ARN présents au moment du prélèvement. Le niveau d'expression des ARN a été normalisé par rapport à celui du gène de ménage *gyrB*. B. Les mesures sont réalisées ici à plusieurs temps de culture pour la souche RN6390 sauvage et la souche RN6390  $\Delta rnc$  (774).

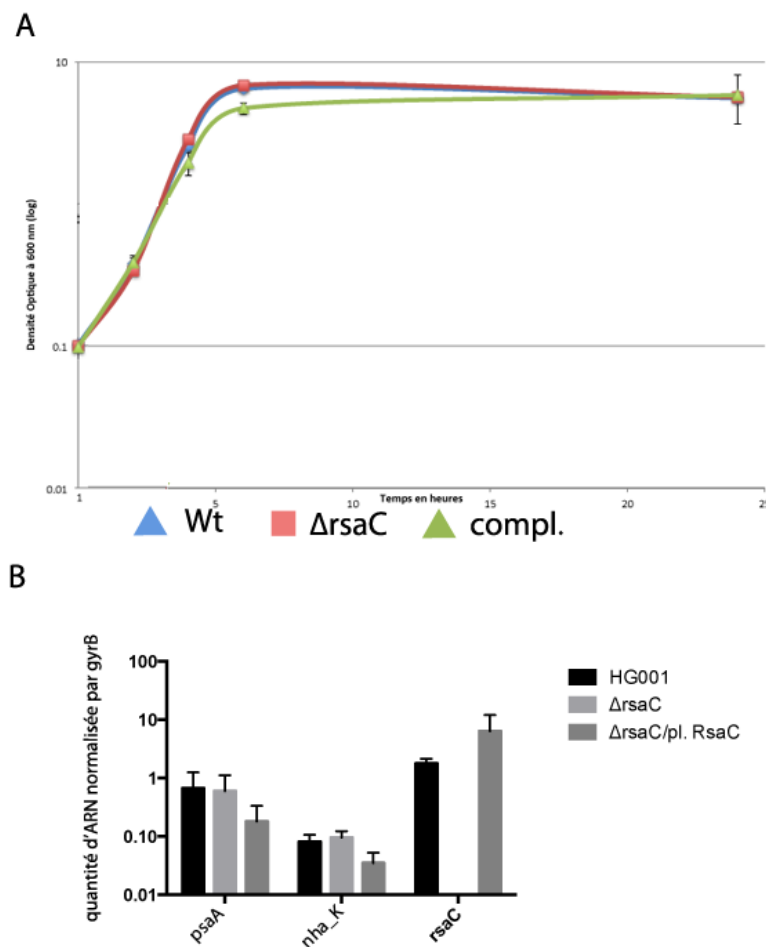
Les résultats de la **figure 27A** montrent qu'un ARN anti-sens à RsaC est effectivement exprimé mais son abondance est environ  $10^4$  fois inférieure à celui de RsaC. D'autres expériences menées dans la souche  $\Delta rnc$  ont montré que l'anti-sens est plus abondant lorsque la RNase III n'est pas présente, mais son niveau d'expression reste malgré tout très faible (**figure 27B**) suggérant que dans nos conditions de croissance, cet anti-sens ne contribue pas de manière significative à la maturation positive de RsaC.

## 5. Construction du mutant RsaC et protéomique comparative

L'identification de la véritable longueur du gène *rsaC* chez HG001 a permis de délimiter avec précision la région à déléter afin d'étudier plus précisément la fonction de cet ARN. La construction de la souche mutante a été menée par l'équipe de F. Vandenesch en

utilisant le vecteur pMAD pour effectuer une double recombinaison. Toute la région du gène *rsaC* a été délétée et remplacée par une cassette de résistance à la kanamycine. Au préalable, il a fallu vérifier si la délétion du gène entraîne une modification de l'expression des gènes qui en sont situés en amont et en aval, et notamment *psaA* puisqu'il se pourrait que RsaC soit transcrit en opéron avec *psaA*.

Dans un même temps, j'ai construit avec l'aide de Zong-Fu Wu le plasmide pCN51 qui porte le gène *rsaC* entier (1117 nucléotides) dont l'expression est contrôlée par le promoteur P3 de l'ARNIII ne connaissant pas où se situe le promoteur endogène. L'avantage majeur de ce promoteur est de pouvoir obtenir une quantité non négligeable d'ARN pour l'approche MAPS. J'ai également utilisé un plasmide pCN51 vide comme contrôle que j'ai transformé dans la souche sauvage de *S. aureus* HG001 pour m'assurer qu'il n'y aurait pas d'effet du plasmide et/ou de l'antibiotique sur l'expression de gènes.



**Figure 28 : construction du mutant  $\Delta$ rsaC chez *S. aureus* RN1** – A. Cinétique de croissance des souches sauvage (Wt),  $\Delta$ rsaC et complétée en milieu BHI à 37°C. B. Mesure de l'expression à 6h de croissance des gènes en amont (*psaA*) et en aval (*nhk*) de *rsaC* par Q-RT PCR, normalisée par l'expression du gène de ménage *gyrB*.

Les niveaux d'expression des gènes *psaA*, *nhaK* et *rsaC* ont tout d'abord été mesurés dans la souche sauvage et dans la souche  $\Delta$ *rsaC* toutes deux transformées avec le même plasmide ne contenant pas d'insert ainsi que dans la souche mutante  $\Delta$ *rsaC* complétée avec le plasmide contenant le gène *rsaC* après 6h de culture en milieu BHI à 37°C. L'expression du gène *gyrB* a également été mesurée pour normaliser les niveaux des trois autres gènes. Chacune des expériences a été réalisée en triplicats. Nous avons tout d'abord vérifié que la croissance des trois souches étaient identiques (**figure 28**). On peut constater une plus forte expression de *rsaC* dans la souche complétée, un résultat attendu puisque le promoteur P3 est sans doute plus actif que le promoteur endogène de l'ARN, et que celui ci est porté par un plasmide à plusieurs copies. L'expression des gènes *psaA* et *nhaK* n'est pas significativement modifiée par la délétion du gène *rsaC*, ni par la surexpression de RsaC. La légère diminution observée n'est pas significative.

J'ai mené également une étude protéomique comparative avec la souche HG001 et  $\Delta$ *rsaC* pour déterminer les protéines dont la synthèse est influencée par la présence ou l'absence de RsaC. Pour ce faire, j'ai récupéré des extraits protéiques totaux de trois cultures différentes des deux souches concernées qui ont ensuite été séquencés par l'approche protéomique déjà décrite dans la Partie 1 des Résultats. Chronologiquement parlant, la plateforme avait entre temps affiné ses protocoles pour des analyses quantitatives plus précises. Enfin, deux méthodes d'analyse des résultats sont utilisées par la plateforme: l'approche *spectral count* et l'approche MS1 *label-free* (Durand et al., 2015). Les deux approches sont suivies de méthode d'identification par des algorithmes différents (Mascot et Paragon, respectivement). Les résultats sont ensuite croisés puis validés manuellement. La variation de synthèse d'une protéine d'un facteur supérieur à 1,5 et inférieur à 0,66 peut être considérée comme significative, et l'utilisation de triplicats biologiques permet d'avoir une pertinence statistique par l'obtention d'une p-valeur (qui doit être inférieure à 0,05). L'expérience de protéomique comparative n'a pas donné un très grand nombre de protéines dont la synthèse était perturbée par RsaC : 7 protéines différentiellement exprimées dans les 2 souches, dont 6 pour laquelle la synthèse est activée en absence de RsaC et une pour laquelle la synthèse est inhibée en son absence. Les résultats après validation manuelle par l'équipe de protéomique sont présentés dans le tableau ci-dessous. Il sera important de refaire ces expériences en

utilisant la souche complémentée afin d'être sûr que les effets observés dépendent de RsaC.

ID	Gène	Description	Nombre de peptides obtenus	Fold change $\Delta$ /Wt
Hg001_01477	sodA	Superoxide Dismutase [Mn/Fe]	8	2.23 - 4.61
Hg001_00640	nagA	N-acetylglucosamine-6-phosphate deacetylase	1	3.27
Hg001_01840	yraA	Putative cysteine protease	4	2.05 – 2.25
Hg001_02120	thiD	Hydroxymethylpyrimidine/phosphom	1	2.39
Hg001_01689	ribH	et-hylpyrimidine kinase	1	2.41
Hg001_02278	rpsQ	6,7-dimethyl-8-ribityllumazine synthase	1	4.83
Hg001_00570	psaA	30S ribosomal protein S17	2	0.04
		Manganese ABC transporter substrate-binding lipoprotein precursor		

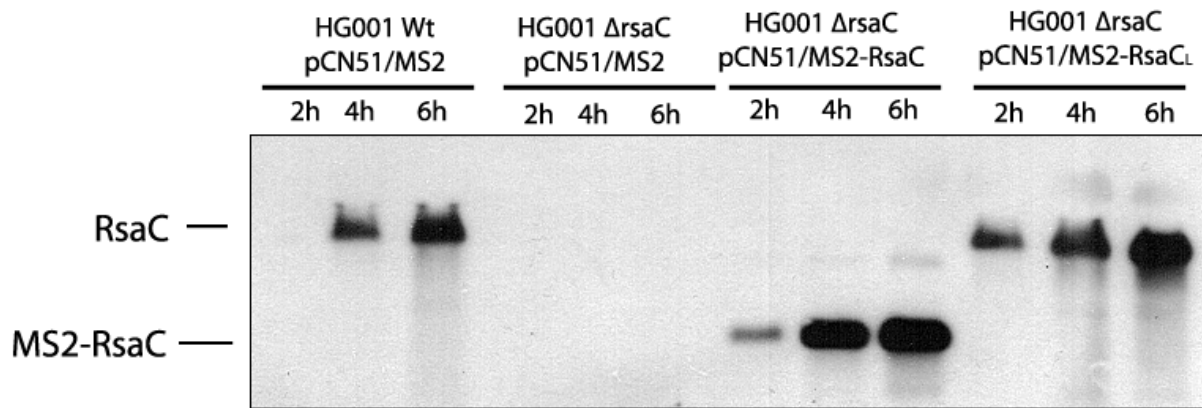
**Tableau 3 : analyse par spectrométrie de masse du protéome de RsaC** - Le *fold change* est la différence de synthèse de la protéine observée entre la souche  $\Delta$ rsaC et la souche sauvage (*S. aureus* HG001). La validation manuelle implique que la p-valeur pour ce fold change est inférieure à 0,05.

Comme le montre le **tableau 3**, c'est la superoxide dismutase codée par le gène *sodA* qui est retrouvée la plus abondamment en absence de *rsaC*, avec un maximum de 8 peptides obtenus et un taux augmenté jusqu'à 4,61 dans la souche mutée. SodA est une enzyme dimérique dont la fonction est de réduire les ions superoxyde en peroxyde d'hydrogène H<sub>2</sub>O<sub>2</sub>, qui est ensuite pris en charge par la catalase pour être transformée en H<sub>2</sub>O et O<sub>2</sub>. Elle assure donc une fonction importante dans la protection contre le stress oxydatif qui peut être causé par l'accumulation des espèces réactives de l'oxygène (ROS) lors de la phase stationnaire de croissance mais également par la production de ROS produits par les macrophages lorsque les bactéries sont internalisées (voir introduction). Et si au départ le rôle de SodA a été démontré pour la protection contre le stress oxydatif d'origine intracellulaire (Clements *et al.*, 1999), elle a également une importance dans la protection du stress oxydatif d'origine extra-cellulaire, et a donc potentiellement un lien avec la virulence (Karavolos *et al.*, 2003). Concernant les autres enzymes trouvées par l'approche du protéome, NagA est une enzyme impliquée dans la synthèse de l'un des composés du peptidoglycane qui constitue l'enveloppe cellulaire. Le gène de la protéine YraA est une potentielle cystéine protéase et RpsQ est une protéine ribosomale de la

petite sous-unité 30S. ThiD et RibH sont, quant à elles, impliquées dans les voies de biosynthèse des acides nucléiques. En revanche, il est très intéressant de constater que l'unique protéine retrouvée moins abondamment dans la souche mutante  $\Delta$ rsaC soit PsaA, une protéine impliquée dans le transport du manganèse – qui est l'ion dont a besoin SodA pour fonctionner – et dont le gène est celui situé juste en amont de *rsaC*. Or j'ai vérifié que la délétion de *rsaC* n'a pas d'influence sur le taux de l'ARNm *psaA* (**figure 28**). Il est donc possible que la diminution de la quantité de PsaA en absence de RsaC soit liée à un contrôle, direct ou indirect, de son expression par cet ARN au niveau traductionnel.

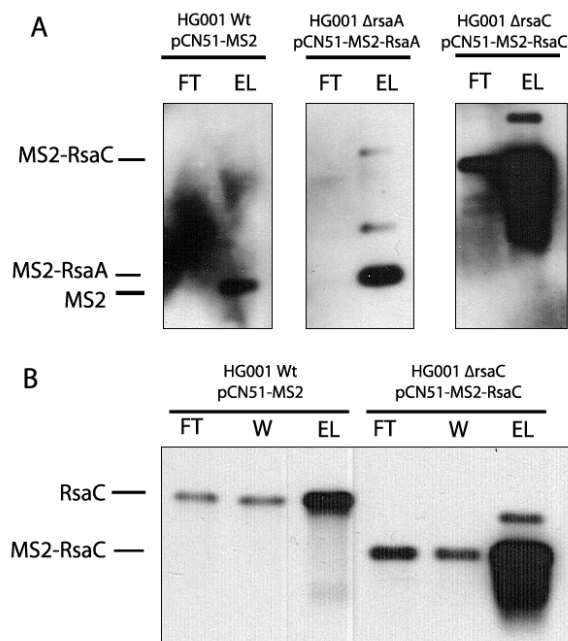
## 6. Application du MAPS à RsaC

Ayant obtenu la souche mutante HG001  $\Delta$ rsaC, j'ai mis en place l'approche MAPS pour RsaC. J'ai d'abord construit deux plasmides pCN51 portant les constructions MS2-RsaC (qui comporte les 544 nucléotides identifiés par Geissman *et al.*) et MS2-RsaCL (pour lesquels les 1117 nucléotides sont inclus). Comme pour la construction MS2-RsaA, ces deux ARN étiquetés sont placés sous le contrôle du promoteur P3. Les souches HG001 Wt et  $\Delta$ rsaC ont également été transformées avec un plasmide pCN51-P3-MS2 qui exprime que l'étiquette MS2, en guise de contrôle. Une fois les souches obtenues (qui sont, pour rappel, listées dans le Matériel et Méthodes), j'ai dans un premier temps vérifié que les ARN sont correctement exprimés, par Northern blot avec une sonde DIG-ARN spécifique de RsaC. Sur la **figure 30** ci-après, on peut voir que les deux ARN étiquetés sont correctement exprimés et s'accumulent en phase stationnaire de croissance, ce qui est dû au contrôle de leur expression par le promoteur P3. La taille de la construction MS2-RsaCL est à la même longueur que la version sauvage de l'ARN (le gel ne permet pas de discriminer ici les 46 nucléotides de différence liés à la présence de l'étiquette MS2) mais les ARN sont homogènes et ne présentent pas de produits de dégradation (**figure 29**).



**Figure 29 : Expression des variants MS2-RsaC chez *S. aureus*** – Les ARN totaux des quatre différentes souches de *S. aureus* ont été prélevés à 2h, 4h et 6h de culture en milieu BHI à 37°C.

J'ai ensuite vérifié que MS2-RsaC est correctement retenu sur la colonne de chromatographie d'affinité. J'ai procédé à l'expérience en suivant le protocole déjà appliqué pour MS2-RsaA. J'ai recherché la présence de MS2-RsaC sur les ARN purifiés depuis l'éluat par Northern blot, en déposant une même quantité d'ARN (1 µg) de chaque échantillon.



**Figure 30 : MS2-RsaC est correctement retenu sur la colonne de chromatographie d'affinité** – A. Visualisation des ARNs MS2, MS2-RsaA et MS2-RsaC dans les *flow-through* (FT) et éluats des extraits bruts des souches utilisées. Des sondes DIG-ARN spécifiques de MS2, RsaA et RsaC ont été utilisées pour chaque expérience. B. Visualisation par Northern blot de RsaC et MS2-RsaC dans le FT, *washing* (W) et EL des souches *S. aureus* wt et exprimant MS2-RsaC.

La présence de l'étiquette MS2 sur un ARN de courte taille entraîne sa rétention spécifique sur la colonne de chromatographie d'affinité comme le montre la **figure 30A**. On voit clairement l'enrichissement important de l'étiquette MS2 et de l'ARN MS2-RsaA

dans l'éluat (EL) par rapport à la fraction de *flow-through* (FT). En revanche avec la construction MS2-RsaC, si on observe également un enrichissement important dans l'éluat, on remarque que l'ARN est également présent avec un signal assez fort dans la fraction FT. J'ai donc voulu vérifier si l'ARN RsaC peut être retenu sur la colonne de chromatographie d'affinité sans l'ajout de l'étiquette MS2. C'est ce qu'on peut observer en effet en **figure 30B** : la version sauvage de RsaC (longue de 1117 nucléotides), est enrichie dans l'éluat, ce qui peut être dû aux nombreuses tiges boucles présentes qui pourraient être reconnues par la protéine de fusion MS2 CP-MBP.

Ceci est à prendre en compte lors de l'analyse des résultats du MAPS. En effet, si RsaC est retenu sur la colonne et enrichi dans l'éluat, certains de ses ARN cibles peuvent se retrouver parmi ceux élués dans la situation contrôle, et ne pas être pris en compte à cause d'un enrichissement trop faible. Les résultats montrent malgré tout que MS2-RsaC est correctement exprimé et est fortement retenu sur la chromatographie d'affinité, ce qui m'a permis de poursuivre les expériences avec cet ARN. En revanche, pour MS2-RsaC<sub>L</sub>, aucun enrichissement de l'ARN n'a été observé dans l'éluat (résultats non montrés). Il se peut que ce soit lié à un défaut de conformation de l'étiquette MS2 – ou de l'ARN entier - lorsque l'étiquette est présente et qui l'empêche d'être retenu sur la colonne. Une autre hypothèse est qu'au vu de la structure secondaire complexe de RsaC<sub>L</sub> (**figure 24**), l'étiquette MS2 ne soit plus accessible à la protéine de fusion MBP-CP. Pour la suite des expériences, j'ai donc utilisé la version courte (544 nucléotides) de RsaC, avec l'hypothèse sous-entendue que ce sont les motifs riches en C qui sont les plus importants pour la fonction de régulateur. Ces motifs sont fortement conservés dans les souches de *S. aureus* alors que les régions répétées sont présentes en nombre variable – même si la séquence elle-même est aussi conservée (**figure 22**).

Comme pour les travaux présentés en seconde partie des Résultats, les ARN purifiés à partir de deux éluats de la souche HG001  $\Delta$ rsaC pCN51/MS2-RsaC ont été séquencés par la société Fasteris (Suisse) qui utilise le séquençage à haut débit sur le principe Illumina. Les données de séquençage (appelées *reads*) ont ensuite été placées par un *workflow* sur la plateforme informatique Galaxy. Le *workflow* est une suite de modules informatiques, de simple d'utilisation, qui permet de vérifier l'intégrité des séquences, leur alignement,

le pourcentage du génome couvert et de procéder ensuite à une analyse différentielle des séquences obtenues en fonction des conditions des expériences.

Dans mon cas, le *workflow* utilisé sur Galaxy comprend les étapes suivantes :

- FastQC : un module de contrôle de qualité des échantillons
- FastQ Groomer : un module qui permet de convertir les données des *reads* dans un format utilisable par le module BOWTIE2
- BOWTIE2 : ce module va permettre d'aligner les *reads* sur le génome de référence (ici HG001).
- Flagstat : le module permet d'obtenir des statistiques en fonction des résultats donnés par le précédent module (comme le pourcentage de *reads* alignés sur le génome, par exemple)
- HTseq : ce module va permettre de compter les *reads* alignés sur le génome.
- DEseq (ou DEseq2) : ces modules finaux vont eux permettre l'analyse différentielle entre les différentes conditions à partir des résultats du précédent module. Pour cette étape, nous avons appliqué deux paramètres pour prendre en compte une différence d'expression significative : un *fold change* entre les conditions qui soit supérieur ou égal à 2,5, et une p-valeur inférieure ou égale à 0,05.

Gene ID	Gène	Description	foldChange RsaC/contrôle	p valeur
HG001_00569	<b>RsaC</b>		<b>69.24</b>	3.26E-22
HG001_01477	<b>gene=sodA</b>	<b>Superoxide dismutase [Mn/Fe] 1</b>	<b>47.94</b>	1.2532E-19
HG001_00671	<b>gene=fecD_1</b>	<b>Fe(3+) dicitrate transport system permease protein FecD</b>	<b>14.66</b>	3.49098E-06
HG001_01627	gene=gloB_1	Hydroxyacylglutathione hydrolase	13.6	3.37123E-11
HG001_02152	gene=fba	Fructose-bisphosphate aldolase	9.32	6.29566E-09
HG001_00934	gene=ptsI	Phosphoenolpyruvate-protein phosphotransferase	7.2	1.76119E-07
HG001_02633		DNA-binding transcriptional regulator EnvR	6.28	0.000114846
HG001_02589		Virus attachment protein p12 family protein	5.82	0.009996967
HG001_00967		Cysteine-rich secretory protein family protein	5.8	3.35627E-06
HG001_01721		product=2-succinylbenzoate--CoA ligase	5.77	4.25473E-06
HG001_02376		hypothetical protein	5.74	2.05146E-05
HG001_00933	gene=ptsH	Phosphocarrier protein HPr	5.73	3.25008E-06
HG001_02411		hypothetical protein	5.49	1.24727E-05
HG001_01479	<b>znuB</b>	<b>High-affinity zinc uptake system membrane protein ZnuB</b>	5.43	2.62239E-05
HG001_01478	<b>zur</b>	<b>Zinc-specific metallo-regulatory protein</b>	5.39	0.000147351
HG001_00252		FtsX-like permease family protein	4.91	2.35389E-05
HG001_01657	ccpA	Catabolite control protein A	4.88	2.66109E-05

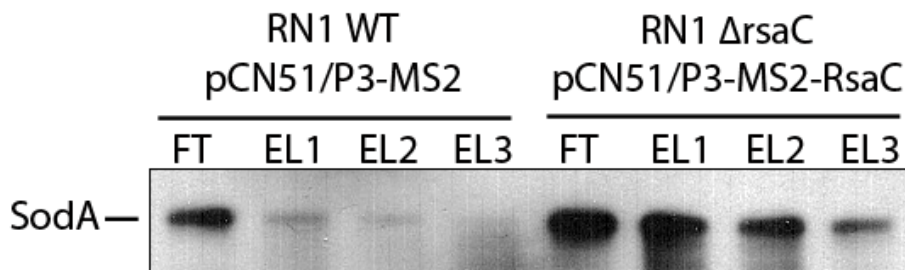


HG001_00968		hypothetical protein	4.43	8.65208E-05
HG001_02659		hypothetical protein (note: Glutathione peroxidase homolog BsaA)	4.34	0.01809111
HG001_01718	ytkD	Putative 8-oxo-dGTP diphosphatase YtkD	4.15	0.000335422
HG001_01220	MsrR	Regulatory protein MsrR	4.09	0.00057282
HG001_00556	sarA	Transcriptional regulator SarA	4.04	0.000215301
HG001_01925	cobQ	Cobyric acid synthase	4.03	0.000821702
HG001_01480	znuC	High-affinity zinc uptake system ATP-binding protein ZnuC	4	0.001378957
HG001_01720	menC	o-succinylbenzoate synthase	3.97	0.000209762
HG001_00712	cggR	Central glycolytic genes regulator	3.89	0.000253853
HG001_02590	feoB	Ferrous iron transport protein B	3.87	0.000527373
HG001_01416	ssrB	Sensor protein SrrB	3.79	0.000360116
HG001_00203	ypdA	Sensor histidine kinase YpdA	3.62	0.001144169
HG001_02591		FeoA domain protein	3.61	0.030171166
HG001_00540	ywgN	Putative NAD(P)H-dependent FMN-containing oxidoreductase YwqN	3.55	0.000677896
HG001_00544	adh	Alcohol dehydrogenase	3.48	0.000654743
HG001_02421		hypothetical protein	3.43	0.025868788
HG001_01626	nhaX	Stress response protein NhaX	3.37	0.001071101
HG001_01301		PepSY-associated TM helix	3.35	0.001727898
HG001_01298	ald1	Alanine dehydrogenase 1	3.34	0.001002313
HG001_01150	glpF	Glycerol uptake facilitator protein	3.33	0.003730805
HG001_01845	lytN	putative cell wall hydrolase LytN precursor	3.27	0.001400277
HG001_01832		hypothetical protein	3.26	0.001662192
HG001_00028	rlmH	Ribosomal RNA large subunit methyltransferase H	3.25	0.001692295
HG001_02747		hypothetical protein	3.25	0.001699863
HG001_01262		hypothetical protein	3.25	0.028686553
HG001_01710	yvgN	Glyoxal reductase	3.23	0.002126565
HG001_00932		hypothetical protein	3.21	0.001611189
HG001_02440	nirC	Nitrite transporter NirC	3.2	0.001543335
HG001_00637	gclR	HTH-type transcriptional repressor GlcR	3.18	0.048553574
HG001_00204	lytR	Sensory transduction protein LytR	3.17	0.016696224
HG001_02519		hypothetical protein	3.13	0.003792509
HG001_01730		hypothetical protein	3.13	0.010742031
HG001_01731		transposase	3.1	0.034927694
HG001_00202	scdA	Iron-sulfur cluster repair protein ScdA	3.07	0.00206853
HG001_02748		hypothetical protein	3.07	0.003413345
HG001_01118		hypothetical protein	3.07	0.003464677
HG001_01343		Ferredoxin	3.07	0.006466669
HG001_00832	oppB	Oligopeptide transport system permease protein OppB	3.07	0.008214483
HG001_00386	treB	PTS system trehalose-specific EIIBC component	3.04	0.002182803
HG001_01119	rplGA	putative ribosomal protein YlxQ	3.04	0.003317533
HG001_01926	murE	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2%2C6-diaminopimelate ligase	3.03	0.004841301
HG001_02416		hypothetical protein	3.02	0.006427676
HG001_02375	mro	Aldose 1-epimerase precursor	2.99	0.003957899
HG001_01717		Prolyl oligopeptidase family protein	2.92	0.006025427
HG001_01550	cymR	HTH-type transcriptional regulator CymR	2.85	0.006978541
HG001_00961		hypothetical protein	2.82	0.005815034
HG001_01116	rimP	Ribosome maturation factor RimP	2.81	0.005677231

HG001_01929	dinB	DNA polymerase IV	2.8	0.008331867
HG001_02069	rex	Redox-sensing transcriptional repressor Rex	2.79	0.006013999
HG001_00568	nhak	Sodium%2C potassium%2C lithium and rubidium/H(+) antiporter	2.77	0.00613882
HG001_02746	immR	HTH-type transcriptional regulator ImmR	2.73	0.006803647
HG001_01117	nusA	Transcription termination/antitermination protein NusA	2.69	0.006734444
HG001_02544	ywaC	GTP pyrophosphokinase YwaC	2.63	0.00753323
HG001_02331		hypothetical protein	2.6	0.011583311
HG001_02304	moaA	Cyclic pyranopterin monophosphate synthase	2.57	0.009325689
HG001_00249	focA	putative formate transporter 1	2.55	0.010782534
HG001_00251	hel	Lipoprotein E precursor	2.53	0.032249308

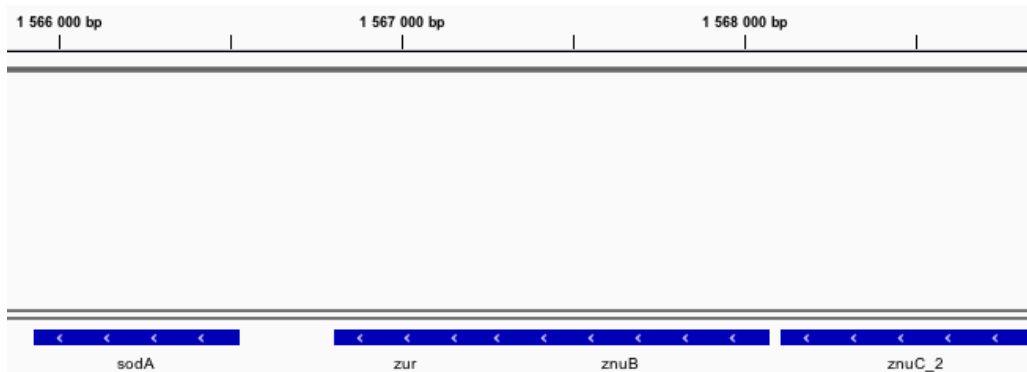
**Tableau 4 : résultats du MAPS réalisé avec MS2-RsaC** – L’analyse a suivi les différents modules listés précédemment avec le module DESeq2 pour l’analyse finale. Les gènes surlignés dans une même couleur indiquent qu’ils sont impliqués dans des fonctions similaires. En jaune : métabolisme du fer et stress oxydatif ; en orangé : métabolisme et transport du zinc ; en rose : ARNm retrouvé avec RsaA ; rose saumon : transport des phosphates. En gras sont indiqués les gènes qui nous ont semblé les plus dignes d’intérêt.

Comme le montre le tableau ci-dessus, nous avons retrouvé 104 ARNs enrichis avec MS2-RsaC après notre expérience de MAPS. Il est logique de retrouver, comme il s’agit de notre ARN « appât », RsaC tout en haut du tableau, et fortement enrichi par rapport à la condition contrôle où l’ARN est exprimé mais non retenu par l’étiquette MS2. En revanche, l’enrichissement pourrait certainement être plus important puisque j’ai observé que RsaC est capable d’être retenu sur la colonne par lui-même (**figure 30B**). Le premier *hit* après RsaC est également le plus intéressant : il s’agit de l’ARNm *sodA*, l’ARN de la superoxide dismutase, protéine qui a été retrouvée en premier *hit* également lors des expériences de protéomique comparative (**tableau 4**), ce qui fait de l’ARNm *sodA* un très bon candidat cible de RsaC. Cet enrichissement dans l’éluat de HG001 qui exprime MS2-RsaC par rapport à la souche contrôle a également été vérifié par Northern blot avec un sonde DIG-ARN spécifique de l’ARNm *sodA* (**figure 31** ci-dessous).



**Figure 31 : L’ARNm *sodA* est enrichi dans l’éluat provenant de l’extrait brut contenant MS2-RsaC** – Northern blot réalisé contre *sodA* avec une sonde DIG-ARN spécifique sur les extraits totaux (FT) et éluats (EL1-3) des souches HG001 exprimant MS2-RsaC et la souche contrôle après passage sur la colonne de chromatographie d’affinité. Pour les FT, 5 µg d’ARN ont été déposés, pour les éluats, 0,5 µg d’ARN a été utilisé.

Parmi les autres ARNm retrouvés, certains ont plus retenu mon attention que d'autres, notamment *fecD*, *feoB*, *scdA*, *zur*, *znuB* et *znuC\_2* qui sont impliqués dans le transport du fer et du zinc. De plus, les gènes *zur*, *znuB* et *znuC* sont les gènes situés juste en aval de *sodA* (**figure 32**). Une régulation de tout l'opéron par RsaC pourrait donc être envisageable.



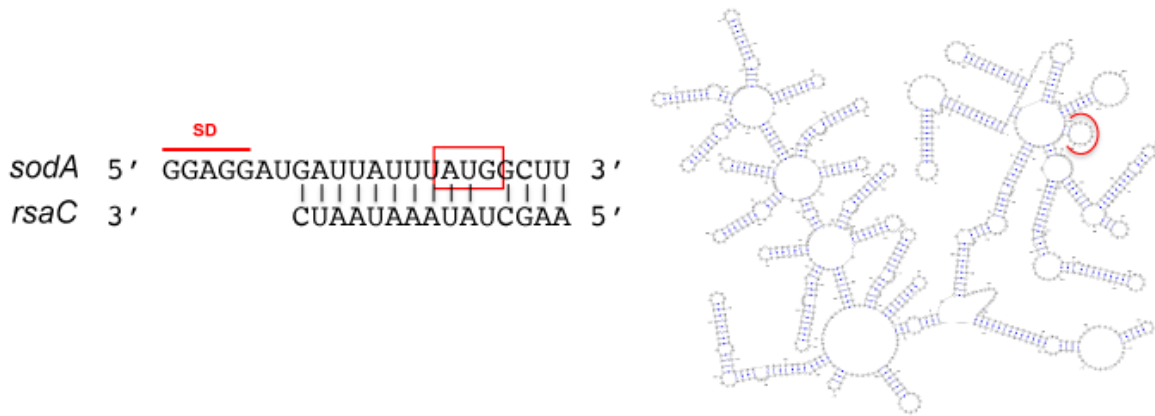
**Figure 32 : Localisation de *sodA* sur le génome de *S. aureus*** – Cette représentation est issue directement du logiciel IGV (*Integrative Genomics Viewer*). La localisation exacte à partir de l'origine de réplication est indiquée tout en haut.

L'ARNm *cymR*, malgré sa position très basse dans le tableau (avec une variation d'enrichissement de 2,85) n'est pas à négliger car c'est un facteur de transcription impliqué dans la régulation de gènes impliqués dans la résistance au stress oxydatif, qui est lui-même capable de sentir la présence de ROS dans la cellule (Ji *et al.*, 2012). Les ARNm *lytN* et *lytR* constituent également des cibles potentielles intéressantes pour leur importance dans la modulation de la synthèse du peptidoglycane, qui peut également avoir un impact sur la virulence. Enfin, je note l'ARNm *rimP* qui est un facteur impliqué dans la maturation du ribosome. Cet ARNm est intéressant car le produit de sa traduction fait partie des protéines qui sont plus exprimées (mais dans ce cas, avec une différence relativement faible) en absence de *rsaC*, comme pour SodA.

J'ai croisé les données du MAPS avec celles obtenues lors des expériences de protéomique pour voir si les protéines et ARNm correspondent entre les différentes approches. Avec les résultats que j'ai obtenus, seules deux protéines ont leurs ARNm retrouvés dans l'expérience MAPS : il s'agit de RimP et de SodA, qui est le premier *hit* dans chacune des approches. Pour RimP, son enrichissement dans la souche  $\Delta$ *rsaC* par rapport à la souche sauvage était de 3,5 fois – mais la p-valeur n'étant que de 0,4, sa validation manuelle n'a pas été effectuée (ce qui explique son absence du **tableau 3**).

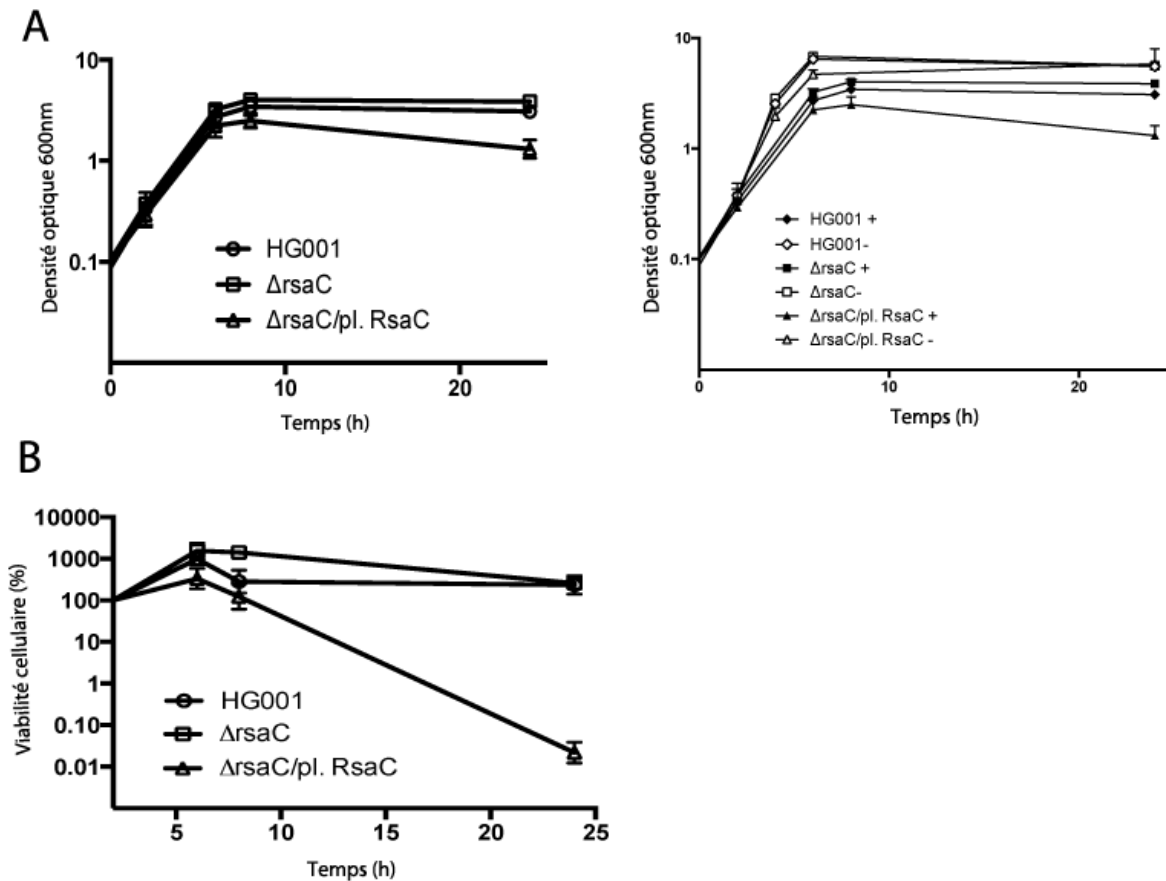
## 7. Implication de RsaC dans la résistance au stress oxydatif

L'approche de MAPS appliquée à RsaC tout comme l'analyse de protéomique ont fait ressortir la même protéine et son ARNm correspondant : SodA. Une analyse de bio-informatique prédit un appariement entre les deux ARN qui se trouve très proche du site SD de l'ARm *sodA* (**figure 33**).



**Figure 33 :** *prédiction d'appariement de séquences entre RsaC et l'ARNm sodA* – A droite est indiquée en rouge le motif de RsaC impliqué dans l'appariement.

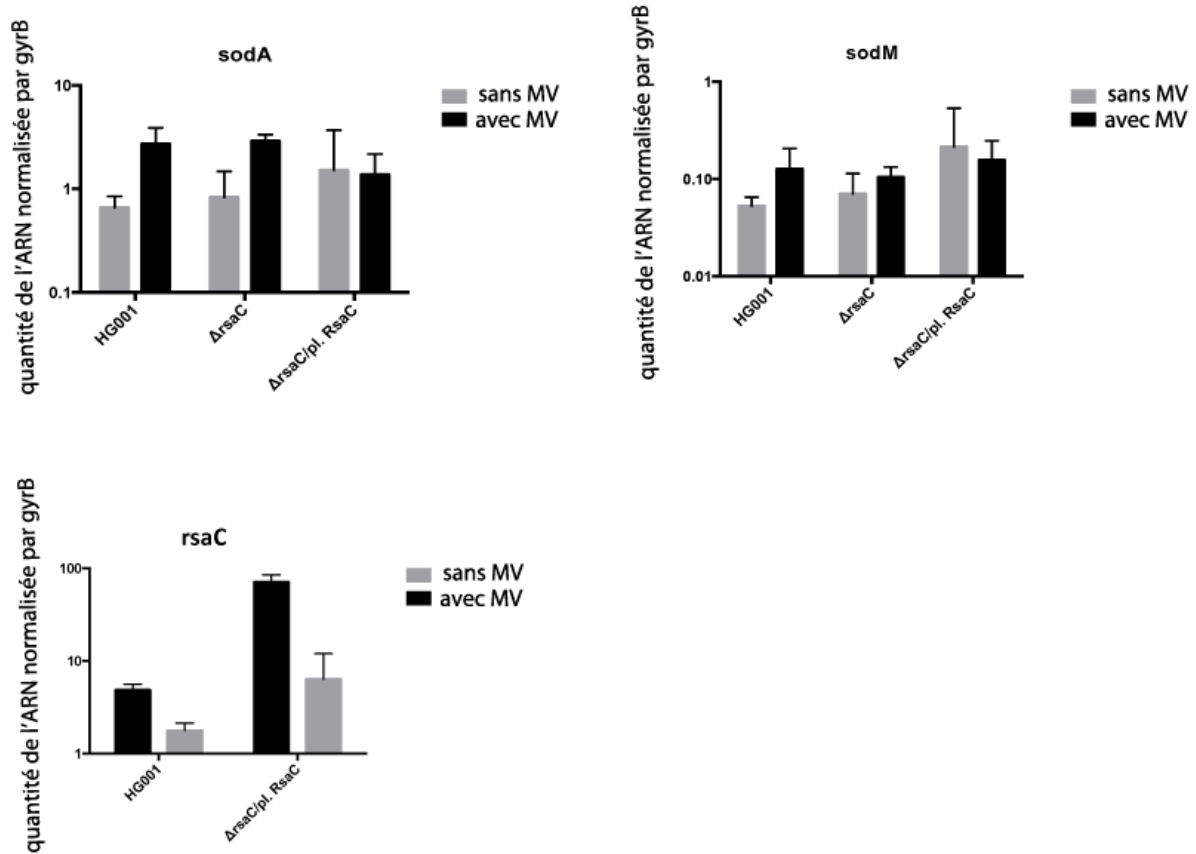
La protéine SodA est impliquée dans la résistance au stress oxydatif et permet notamment de lutter contre l'accumulation de ROS interne. Pour vérifier l'implication et l'importance de RsaC dans la résistance au stress oxydatif, nous avons décidé de tester la résistance au méthyl-viologène (MV) des souches de *S. aureus* HG001 sauvage,  $\Delta$ *rsaC* et complémentée. Ce composé est capable de pénétrer dans les cellules de *S. aureus* qui génère l'accumulation d'ions superoxydes. Les expériences présentées ci-après ont été réalisées à Lyon et à chaque fois en plusieurs réplicats biologiques. Dans un premier temps, les courbes de croissance des trois souches ont été mesurées en milieu BHI à 37°C avec ou sans ajout de MV après 2h de croissance. De même, la viabilité des cellules sans MV ou après ajout de MV à 2h de croissance a été mesurée par comptage cellulaire.



**Figure 34 : Impact du MV sur la croissance et la survie cellulaire** – A. Gauche : courbes de croissance des souches HG001 (O),  $\Delta$ rsaC (□), et  $\Delta$ rsaC/pl. RsaC (Δ) en BHI à 37°C après le stress oxydatif interne par ajout de MV à 2h de croissance. Les données représentent la moyenne de trois expériences différentes. Droite : courbes de croissance des souches HG001 (◇, ◆),  $\Delta$ rsaC (□, ■),  $\Delta$ rsaC/pl. RsaC (Δ, ▲) en absence (symboles vides) et présence (symboles pleins) de MV (ajouté à 2h de croissance). B. Viabilité cellulaire des souches HG001 (O),  $\Delta$ rsaC (□), et  $\Delta$ rsaC/pl. RsaC (Δ).

On peut observer un effet de RsaC sur la croissance cellulaire en présence de MV dans la **figure 34A**. En effet, après 20h de culture, la densité optique des cultures de la souche complémentée HG001  $\Delta$ rsaC/pl. RsaC est plus faible que celle des souches sauvage et mutante. Le MV induit un stress oxydatif interne. La protéine SodA permet de résister à ce stress. Or comme RsaC est surexprimé dans la souche complémentée, la quantité de SodA est moindre, et la souche devient plus sensible au stress oxydatif. Cette sensibilité est due à la présence de MV car en l'absence de ce composé dans le milieu de culture, les trois souches montre une courbe de croissance identique. Cette sensibilité accrue au stress oxydatif lorsque RsaC est surexprimé est corroboré dans la **figure 34B** : les cellules qui surexpriment RsaC voient leur viabilité diminuée drastiquement au cours du temps comparé aux deux autres souches.

Dans un second temps, nous avons regardé l'impact de la présence de MV sur le taux de RsaC, de sa potentielle cible *sodA* mais aussi sur le gène *sodM* qui code la seconde superoxyde dismutase présente chez *S. aureus*.



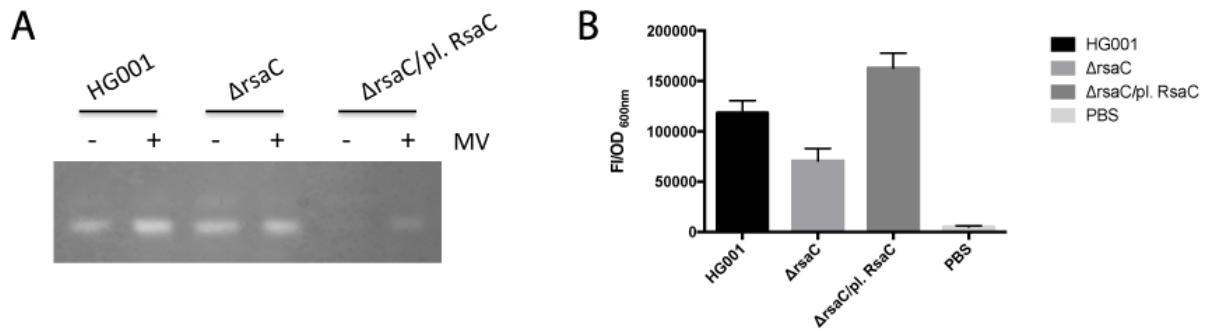
**Figure 35 : Impact du MV sur l'expression de RsaC, *sodA* et *sodM*** – L'expression des trois gènes a été quantifiée et normalisée par la quantité de *gyrB* par RT-qPCR sur des ARN totaux prélevés après 6h de culture des souches en milieu BHI à 37°C, avec ou sans ajout de MV à 2h de culture. Les données présentent les résultats de trois expériences séparées.

Comme ni la protéine SodM ni son ARNm n'ont été retrouvés dans nos expériences de protéomique comparative et de MS2-Seq, cela constitue un bon témoin négatif pour ces expériences. Comme on peut le voir dans la **figure 35**, l'expression de *sodM* n'est pas influencée de façon significative par la présence de MV, que ce soit en absence ou en présence de RsaC. Un résultat qui n'est de toute façon pas surprenant car SodM est une protéine plus impliquée dans la réponse au stress oxydatif d'origine extra-cellulaire (Karavolos *et al.*, 2003). Pour *sodA*, les quantités de l'ARNm entre la souche sauvage et la souche mutante dans chaque condition, ne subissent pas de variation significative. La présence de MV induit légèrement la transcription de *sodA* dans les deux souches. En revanche, dans la souche complétementée, lorsque RsaC est surexprimé, on remarque que la quantité de l'ARNm *sodA* ne varie pas, lorsque le MV est ajouté en milieu de culture.

En revanche, ce taux diminue légèrement dans la souche complémentée par rapport aux souches mutantes et sauvage en présence de MV suggérant une inhibition de l'expression de *sodA* par RsaC.

Concernant RsaC, on observe qu'il est, en absence ou en présence de MV, plus exprimé dans la souche complémentée que dans la souche sauvage, ce qui s'explique par la présence du gène sur le plasmide (plusieurs copies du gènes) et par le fait que nous ayons utilisé le promoteur P3 qui est induit fortement à 6h. On constate également un taux plus abondant de *rsaC* en présence de MV. Ceci suggère que le MV peut avoir une action sur la transcription (que ce soit sur le promoteur endogène ou le promoteur P3) ou affecté de manière indirecte la stabilité de RsaC. Ces résultats posent une contradiction avec ce qu'on observe concernant l'expression de SodA. En effet, dans la souche sauvage, si RsaC est plus exprimé en présence de MV, alors nous devrions voir une diminution de la quantité de l'ARNm *sodA* dans la même condition, ce qui n'est pas ce que nous observons. De plus, en visualisant l'ARNm *sodA* par Northern blot dans les souches sauvages, mutantes et complémentées pour *rsaC*, je n'ai observé aucune différence (résultats non montrés). L'ensemble des résultats sera traité plus en détails dans la partie Discussion mais il semble au vu de ces expériences que l'effet de RsaC sur SodA ne se fasse pas au niveau transcriptionnel. Au vu de la prédiction d'appariement obtenue par bio-informatique, il me semble donc que la régulation de SodA se ferait au niveau post-transcriptionnel, possiblement par une inhibition de la traduction de l'ARNm *sodA*.

Nous avons également vérifié l'impact de RsaC sur la résistance au stress oxydatif par trois autres expériences, inspirées notamment des travaux de Clements *et al.*, 1999. Le premier test est une mesure de l'activité de SodA depuis des extraits protéiques totaux issus des souches après ajout de MV dans le milieu de culture. Les protéines totales sont séparées par électrophorèse sur gel natif et l'activité de SodA est mesurée par coloration négative du gel avec du NBT (*nitro blue tetrazolium*). La seconde approche est une mesure par fluorescence des quantités de ROS dans chaque souche après 6h de culture, qui correspond à la phase stationnaire de croissance pendant laquelle les ROS, générés par le métabolisme, s'accumulent.



**Figure 36 : RsaC joue un rôle dans la résistance au stress oxydatif** – A. Mesure de l'activité SodA chez les souches sauvage, mutante  $\Delta$ rsaC et complétée sur des extraits protéiques prélevés après 8h de culture à 37°C en BHI. L'ajout de MV a été réalisé après 2h de culture. B. Mesure par fluorescence des ROS présents chez les mêmes souches après 6h de culture à 37°C en milieu TSB. Le PBS sert de contrôle négatif.

Dans la figure **36A** on observe en premier lieu que l'ajout de MV provoque une légère augmentation de l'activité SodA pour les 3 souches. En absence de RsaC, les différences avec la souche sauvage ne sont pas frappantes, même si l'activité SodA dans la condition sans MV semble plus importante que chez la souche sauvage, ce qui pourrait s'expliquer par une levée de l'inhibition de la traduction de l'ARNm *sodA* en absence de l'ARN. On pourrait également s'attendre à ce que l'activité de SodA soit augmentée par ajout de MV. C'est ce que l'on observe chez la souche mutante lorsque l'on compare l'activité de SodA en absence et en présence de MV. En revanche, on ne voit pas de différence notable en présence de MV entre la souche mutante et la souche sauvage. Dans la souche complétée, la surexpression de RsaC a un effet sur l'activité de SodA. En absence de MV, l'activité de SodA est nulle ou extrêmement faible et n'est pas détectée par notre méthode de coloration du gel. En présence de MV, on observe une petite activité de SodA, mais qui est bien inférieure à celle observée dans les deux précédentes conditions. La surexpression de RsaC diminue donc l'activité SodA, ce qui va dans le sens de notre hypothèse d'une inhibition de la traduction de l'ARNm *sodA* par RsaC, mais ceci ne semble se produire que lorsque RsaC est surexprimé (puisque nous n'observons pas d'effet sur l'activité de l'enzyme en comparant seulement les souches sauvages et mutantes). On ne peut donc pas écarter l'hypothèse que l'effet de RsaC sur SodA soit uniquement dû à la surexpression de l'ARN et n'ait pas de réelle signification biologique.

Nous observons dans les résultats montrés en **figure 36B** une diminution de la quantité de ROS chez la souche mutante par rapport à la souche sauvage, et cette quantité ré-augmente, et devient plus importante dans la souche complétée, toujours par rapport à la souche sauvage. La surexpression de RsaC dans cette dernière explique le

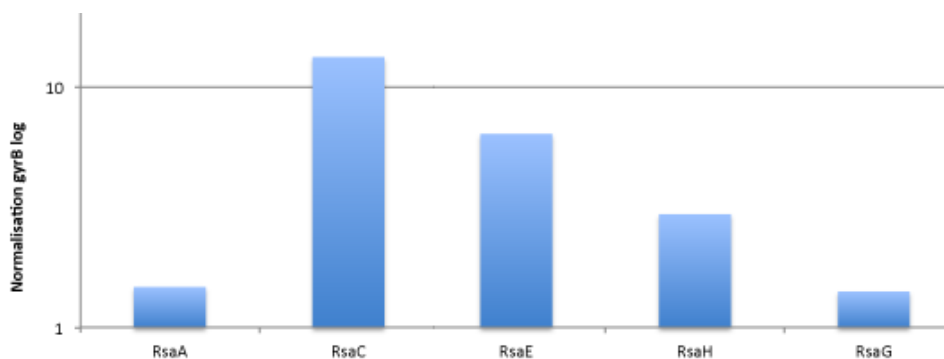


résultat : comme il y a plus de RsaC, il y a moins de protéine SodA qui est traduite, et donc plus de ROS accumulées après 6h de croissance. Mis bout à bout, ces résultats montrent clairement une implication de RsaC dans la résistance au stress oxydatif, et les résultats seront discutés plus en détails dans la partie Discussion de ce manuscrit.

## 8. Un lien possible entre RsaC et la persistance bactérienne

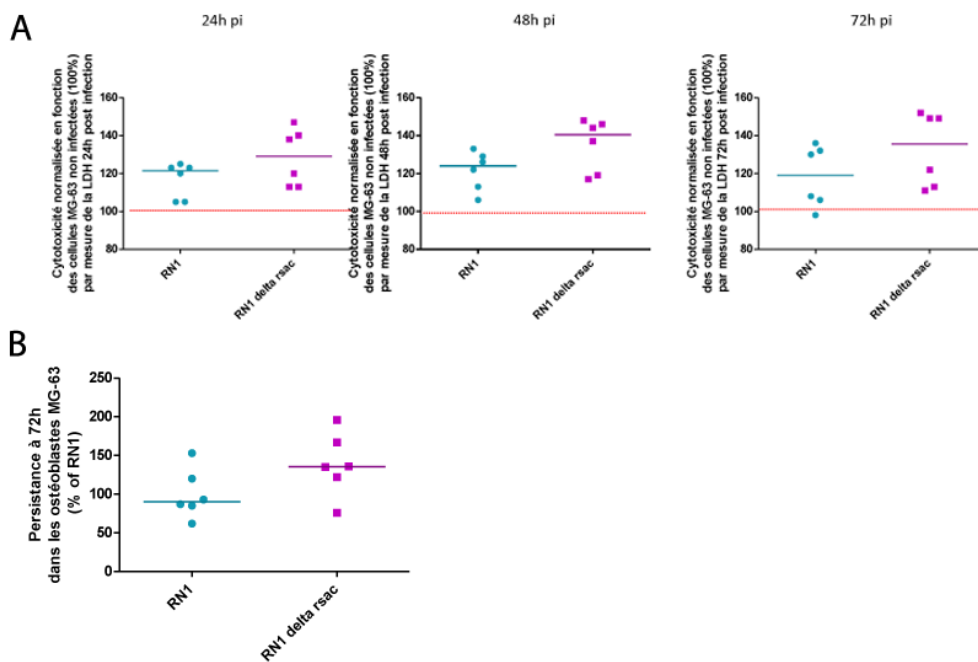
Une étude récente (Szafranska *et al.*, 2014) a montré que dans un modèle d'ostéomyélite (infections d'ostéoblastes par *S. aureus*) l'ARN RsaC est surexprimé par rapport aux conditions de culture *in vitro*. Dans ce modèle infectieux en particulier, dans lequel les cellules bactériennes sont internalisées par les ostéoblastes, l'ARN pourrait donc jouer un rôle pour la survie cellulaire, et donc la persistance bactérienne au sein de ces cellules.

Toujours en collaboration avec Karen Moreau de l'équipe de F. Vandenesch à Lyon, nous avons dans un premier temps confirmé les résultats de Szafranska *et al.* en mesurant par RT-qPCR les niveaux d'expressions de différents ARNnc de *S. aureus* au cours de l'infection d'ostéoblastes. Les cellules de *S. aureus* ont été incubées avec les cellules MG63 et récoltées 24h après l'infection suivant un traitement au lysozyme et à la gentamycine pour lyser les bactéries extracellulaires. Sur les cellules bactériennes intracellulaires, les ARN ont été extraits, purifiés de toute trace d'ADN, et des réactions de rétro-transcription spécifiques pour chaque ARNnc, suivies d'amplification par PCR, ont été conduites. Les résultats de cette expérience sont montrés dans la **figure 37**.



**Figure 37 : Expression intracellulaire (cellules MG63) de différents ARNnc de *S. aureus*** – Les mesures de RTq-PCR sont réalisées 24h post-infection (p.i.) et normalisées par la quantité de l'ARN du gène de ménage *gyrB*. Les mesures de trois autres ARNnc de *S. aureus* (RsaE, RsaH et RsaG) sont montrées à titre de comparaison.

On peut voir d’après les résultats que RsaC est l’ARNnc le plus exprimé au cours de l’infection d’ostéoblastes, ce qui vient confirmer les mesures déjà effectuées par Szafranska *et al.* Il est intéressant que RsaE qui est impliqué dans le maintien d’une homéostasie redox intracellulaire (Durand *et al.*, 2015). Nous avons donc voulu voir en utilisant nos souches mutantes si RsaC tient un rôle important dans la survie de *S. aureus* dans les cellules MG63 et s’il peut également jouer un rôle dans la persistance cellulaire, c’est-à-dire la capacité à rester viable après une longue durée dans le milieu intracellulaire – et une capacité qu’on peut mettre directement en relation avec les infections chroniques à *S. aureus* en milieu hospitalier.



**Figure 38 : Tests de cytotoxicité et de persistance cellulaire dans les cellules MG63 (ostéoblastes)** – A. Test de cytotoxicité par mesure de la LDH chez les cellules infectées avec *S. aureus* WT et  $\Delta$ rsaC. Les mesures sont effectuées pour 6 lots de cellules infectées par souche, 24h, 48h et 72h après l’infection. Les barres latérales représentent la médiane de chaque ensemble de mesure. B. Mesure de la persistance cellulaire par comptage cellulaire des *S. aureus* encore viables 72h après infection des cellules MG63.

Pour ce faire, nous avons dans un premier temps mesuré la cytotoxicité des souches *S. aureus* sauvage et  $\Delta$ rsaC à 24h, 48h et 72h p.i. des cellules MG63 par dosage de la lactate déshydrogénase (LDH), une enzyme importante dans le métabolisme des sucres et qui est relarguée dans le milieu extracellulaire lors de dommages cellulaires. Les résultats présentés dans la **figure 38A** ci-dessous montrent que la souche mutante provoque plus de dommages dans les cellules que la souche sauvage, et ce de façon plus importante après 48h p.i. RsaC pourrait donc avoir un effet atténuateur de la virulence. Dans un second temps, c’est la persistance bactérienne des deux mêmes souches de *S. aureus* qui

a été mesurée 72h après infection des cellules MG63. Les résultats peuvent être observés en figure **38B** et montrent que la souche  $\Delta$ rsaC est retrouvée en plus grand nombre que la souche sauvage, montrant une meilleure persistance. RsaC semble donc ici avoir un effet inhibiteur sur la persistance bactérienne, qui pourrait être relié à la résistance au stress oxydatif. Les résultats seront vus plus en détails dans la partie Discussion & Perspectives.

**L'ARN RsaC de *S. aureus* n'est pas un ARN simple à étudier. L'obtention du mutant  $\Delta$ rsaC a été hasardeuse, et l'étude de sa fonction a été entravée par la nécessité de caractériser correctement la longueur de l'ARN, qui varie entre les souches à cause de la présence de régions répétées, dont l'importance est encore à vérifier. Néanmoins j'ai pu, dans notre souche modèle HG001, déterminer la taille et la structure secondaire du transcrit. Les analyses combinées de protéomique comparative et de MAPS ont permis de dégager une cible potentielle majeure : Soda. Les tests physiologiques réalisés, comme d'autres ARN retrouvés dans mes expériences de MAPS, pointent toutes vers un lien important de RsaC dans la réponse au stress oxydatif. Il reste encore des expériences à réaliser pour m'assurer que Soda est une cible primaire de RsaC mais l'élucidation de la fonction de RsaC est à présent à portée de main.**

## Partie 4 : Mise au point de la technique Grad-Seq

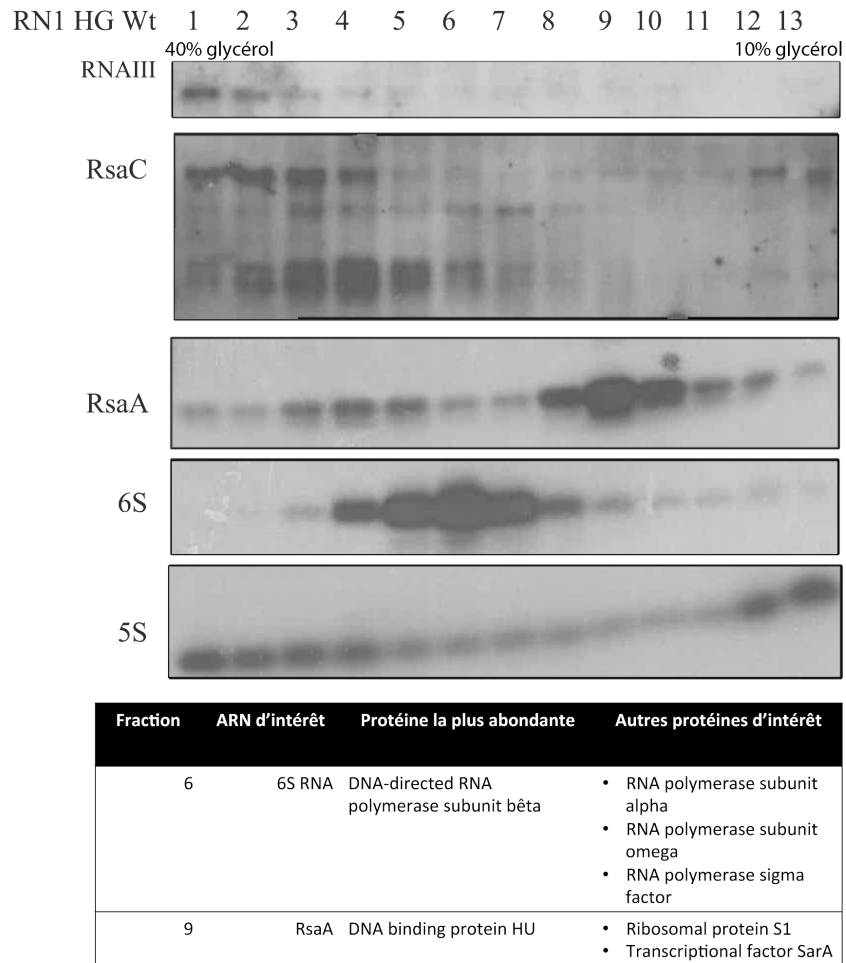
Comme je l'ai expliqué dans les objectifs de mon projet de thèse, l'approche MAPS n'est pas la seule que j'ai employée pour aller rechercher les cibles primaires des ARNnc de *S. aureus* et les éventuelles protéines associées. En effet, la technique du MS2 n'est spécifique que d'un seul ARN donné à la fois. Il nous a semblé intéressant d'adopter une approche plus globale, qui permette d'analyser un ensemble d'ARNnc de la bactérie ainsi que leurs partenaires ribonucléiques et protéiques. En nous basant sur les travaux de Wassarman & Storz (2000), nous avons décidé d'appliquer la méthode de séparation des complexes ribonucléoprotéiques formés *in vivo* par centrifugation sur gradient de glycérol. Après centrifugation, les gradients sont séparés en différentes fractions (les plus lourdes à 40% de glycérol, aux plus légères qui ne contiennent que 10% de glycérol) sur chacune desquelles les ARN et les protéines peuvent être séquencés.

Avant d'appliquer la méthode pour nos ARN d'intérêt, nous avons utilisé l'ARN 6S comme preuve de concept. En effet, l'approche des gradients de glycérol est celle qui a permis de mettre en évidence l'interaction de l'ARN 6S avec l'ARN polymérase, par co-localisation de l'ARN avec les sous-unités  $\beta$  de l'enzyme dans les mêmes fractions (Wassarman & Storz, 2000). J'ai donc recherché sur mon premier gradient de glycérol la localisation dans les fractions de l'ARN 6S par northern blot, puis regardé dans un second temps quelles sont les protéines les plus abondantes dans les fractions dans lesquelles l'ARN 6S est le plus enrichi. Le séquençage des protéines a été réalisé par spectrométrie de masse comme décrit précédemment dans ce manuscrit. Sur les mêmes fractions, j'ai également recherché par Northern blot la présence de plusieurs ARN d'intérêt : RsaA et RsaC bien entendu, mais également l'ARNIII, ainsi que l'ARN 5S.

Sur les résultats présentés en **figure 39**, on observe dans un premier temps que chaque ARN présente un profil de migration différent, avec des enrichissements spécifiques. Ainsi, l'ARN 6S présente un enrichissement très fort au niveau de la fraction 6, située au milieu du gradient de glycérol, avec un profil de migration qui ressemble fortement à

celui observé chez *E. coli* (Wassarman & Storz, 2000). En regardant par spectrométrie de masse les protéines les plus abondantes présentes dans la fraction 6, nous avons obtenu comme premiers *hits* les sous-unités  $\beta$ , mais aussi alpha, de l'ARN polymérase. L'interaction de l'ARN 6S avec ce complexe protéique étant déjà connue, et ces résultats venant corroborer ceux de Wassarman & Storz (2000), j'ai ainsi pu valider la pertinence de l'application de cette approche chez *S. aureus*.

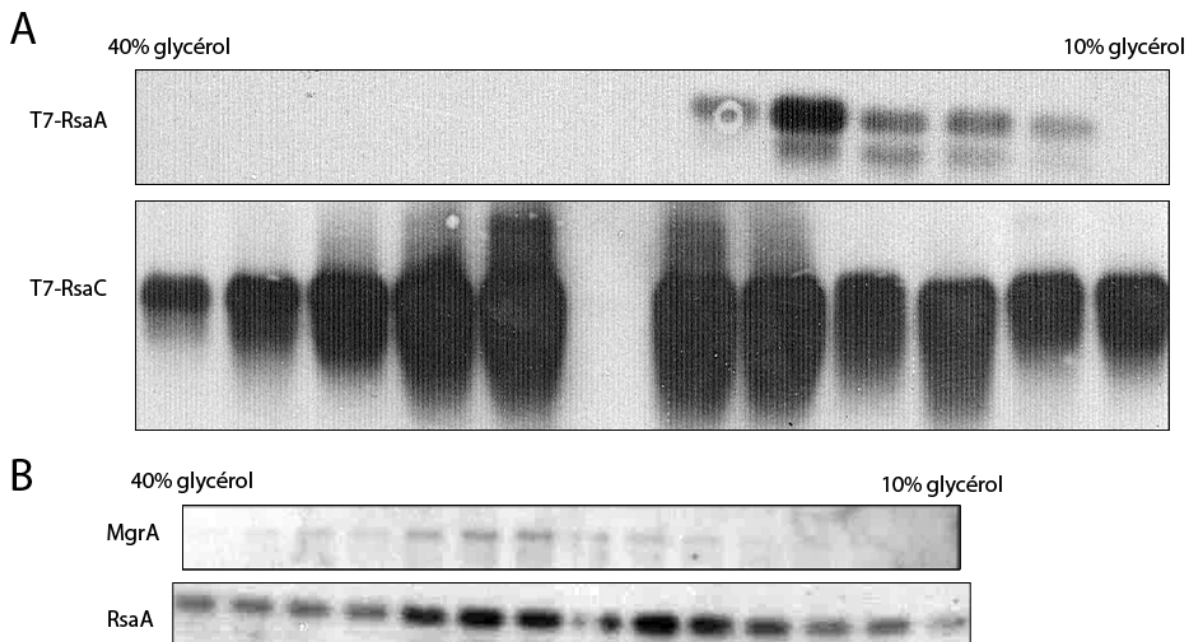
On peut à présent commenter la localisation des autres ARN étudiés. L'ARNr 5S semble diffuser dans diverses fractions du gradient de glycérol. Une grande partie néanmoins a sédimenté avec les ribosomes au fond du gradient. Cependant il est clair qu'une certaine fraction de l'ARNr 5S se retrouve dans la bactérie sous forme libre ou associée à des complexes protéiques intermédiaires du ribosome. Il se peut que cette fraction corresponde à un ARNr 5S en cours de maturation. L'ARNIII est retrouvé dans le fond du gradient de glycérol, certainement en association avec les ribosomes puisqu'il contient dans sa séquence une petite ORF qui code l'hémolysine delta. Pour RsaC, la présence dans le bas du gradient de glycérol laisse aussi penser à une interaction avec des complexes protéiques de grande taille, et il n'est pas exclu que cet ARN possède également une ORF qui code un petit peptide de type PSM, comme chez l'ARNIII, même si nous n'avons pas réussi pour le moment à le détecter. On remarque également pour RsaC la présence d'ARN de tailles différentes, mais il n'est pas certain que ces fragments relèvent d'une maturation de RsaC (car nous avons vu qu'il possède une importante stabilité), mais plutôt à une dégradation au cours de la centrifugation du gradient.



**Figure 39 : Validation de l'approche du gradient de glycérol chez *S. aureus*** – Les différents ARN ont été localisés sur les fractions du gradient après purification des ARN et dépôt d'une même quantité d'ARN totaux pour chaque fraction. La détection est réalisée à l'aide de sondes DIG-ARN spécifiques pour chaque ARNnc. En bas, le tableau indique quelles sont les protéines les plus abondantes retrouvées dans les fractions d'enrichissement de l'ARN 6S (n°6) et de RsaA (n°9).

RsaA présente deux pics d'enrichissements dans son profil de migration, un dans la fraction 9 et un second au niveau des fractions 3 et 4, qui se rapprochent des fractions les plus lourdes (avec le plus fort pourcentage de glycérol) du gradient. Ce résultat montre que RsaA peut interagir avec des protéines qui l'entraînent dans le fond du gradient, car une expérience de gradient de glycérol sur lequel je n'ai déposé que l'ARN RsaA transcrit *in vitro* montre que cet ARN migre très peu dans le gradient (**figure 40A**). Il y a donc forcément des facteurs protéiques ou autres ARNm qui sont impliqués dans cette localisation observée *in vivo*. J'ai testé cette propriété également pour RsaC (544 nucléotides) que j'ai transcrit *in vitro* et déposé sur gradient, mais cet ARN réussit à pénétrer par lui-même dans les fractions lourdes du gradient, ce qui amène sa

visualisation dans ces mêmes fractions *in vivo* à re-considérer d'un point de vue biologique.



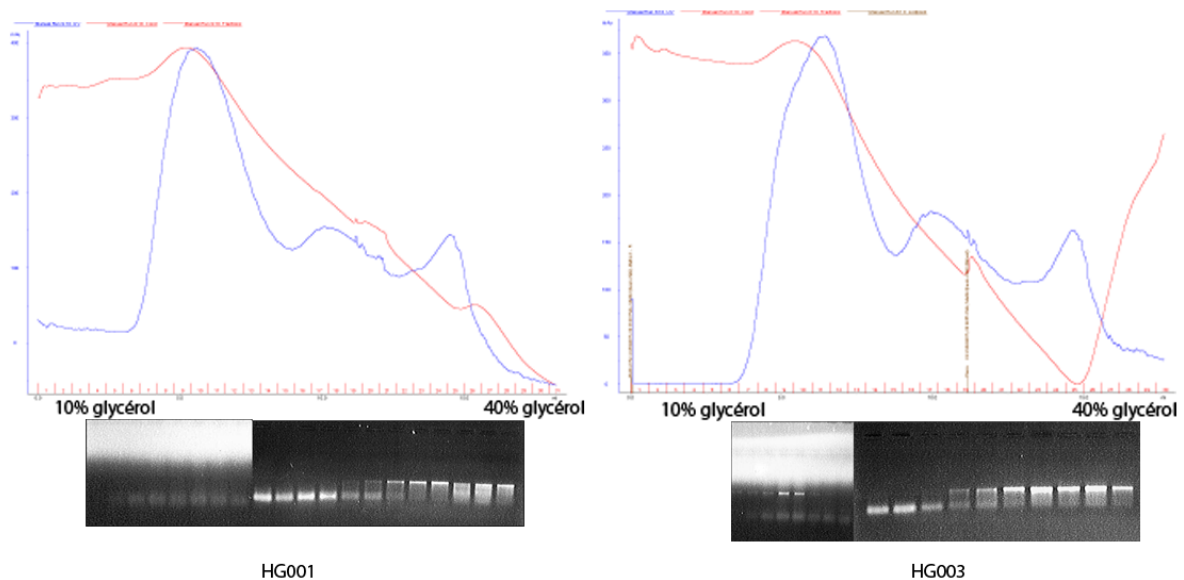
**Figure 40 : Le gradient de glycérol permet de suivre les interactions ARN/ARN et ARN/protéines *in vivo*** – A. Northern blot réalisé sur un gradient de glycérol sur lequel seul le transcrit RsaA ou RsaC (transcrit *in vitro*) a été déposé. Les résultats montrent deux gradients différents, un réalisé pour chaque ARN, qui a été détecté à l'aide d'une sonde DIG-ARN spécifique. La fraction vide pour RsaC correspond à une perte des ARN au cours de la purification. B. Visualisation par Northern blot des ARN RsaA et MgrA sur un gradient de glycérol réalisé avec les extraits totaux de la souche HG001.

Comme pour l'expérience de MAPS, nous avons localisé l'ARNm *mgrA* dans le gradient de glycérol et comparer sa répartition à celle de RsaA. Comme le montre la **figure 40B** on observe un enrichissement de MgrA dans l'un des pics où se trouve RsaA. Cette co-localisation est importante puisqu'elle nous montre que les duplexes d'ARN chez *S. aureus* sont stables et peuvent être suivis après l'étape de centrifugation.

Afin de séquencer tous les ARNs présents dans chaque fraction d'un gradient de glycérol, nous avons commencé une collaboration avec l'équipe de J. Vogel localisée à Würzburg, qui a réalisé le séquençage sur toutes les fractions du gradient. J'ai donc réalisé deux nouveaux gradients en prenant des extraits bruts de la souche HG001 mais également d'une seconde souche très proche, appelée HG003. Entre temps, notre méthode de collecte des fractions a évolué (en utilisant une machine AKTA®) et il a donc fallu s'assurer que la validité de notre approche n'était pas entachée par ce nouvel outil technologique. De même, j'ai voulu comparer les gradients entre eux pour

m'assurer qu'il n'y a pas de différence significative dans la séparation des molécules dans les extraits bruts des souches HG001 et HG003.

La première indication de la reproductibilité des expériences se fait lors de la collecte des fractions par le système AKTA®. En effet, la densité optique à 280 nm (par laquelle on peut mesurer la quantité de protéines) est mesurée au fur et à mesure que les fractions passent dans l'appareil de collecte. En comparant les deux profils obtenus par cet outil de mesure, on peut voir en **figure 41** que les gradients de HG001 et HG003 sont extrêmement similaires, et présentent trois grands pics d'absorbance à 280 nm.



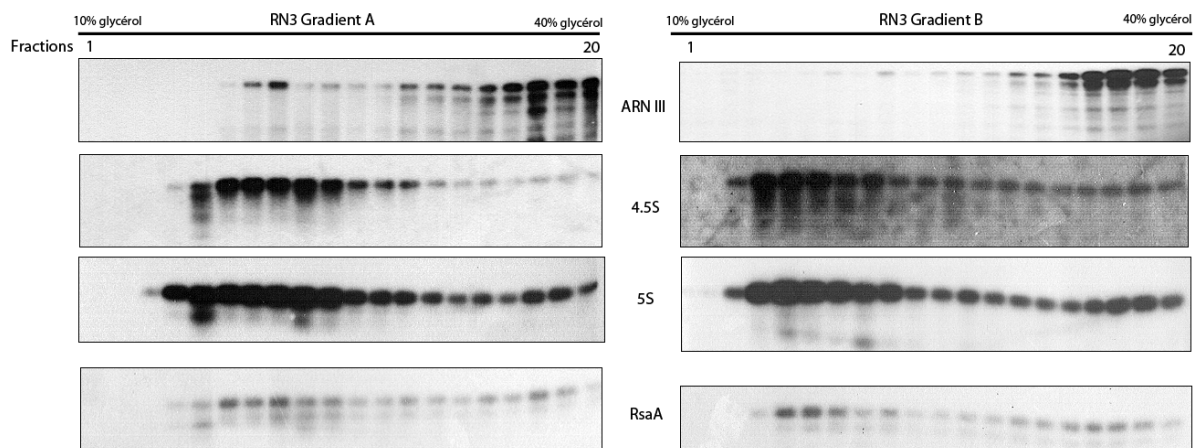
**Figure 41 : Vérifications de la reproductibilité de la collecte des gradients de glycérol par AKTA® et de l'intégrité des ARN après centrifugation.** – Profils d'absorbance à 280 nm des gradients réalisés sur les extraits bruts des souches HG001 et HG003 après collecte par le système AKTA. L'absorbance est indiquée en bleu, la conductivité en rouge. En dessous de chaque profil, les ARN correspondant issus de chaque fraction pour chaque gradient sont déposés (1 µg) sur gel d'agarose et l'intégrité est vérifiée par coloration du gel au BET.

Le pic le plus à gauche, situé dans les fractions légères, correspond certainement aux protéines libres ou associée en complexes de petite taille. Le pic du milieu correspond au complexe de l'ARN polymérase ; quand au pic tout à droite, il correspond certainement aux sous-unités 30S du ribosome. Les ribosomes entiers et polysomes ne sont pas récoltés mais sont sédimentés dans les culots de mes gradients, à cause de la durée de la centrifugation. J'ai également vérifié l'intégrité des ARN en déposant 1 µg d'ARN totaux de chaque fraction sur un gel d'agarose, coloré ensuite au bromure d'éthidium (BET). Comme le montre la **figure 41**, les ARN sont dans un bon état : on ne voit pas d'échelles de dégradation (aussi appelée « smear ») et deux bandes donnent un signal important : il s'agit des ARNr 23S et 16S.



Les expériences présentées montrent donc que l'approche du gradient de glycérol est reproductible et applicable aux deux souches analysées. Pour la suite des manipulations, j'ai continué de travailler avec la souche HG003 uniquement, en utilisant deux gradients réalisés avec deux extraits bruts issus de réplicats biologiques.

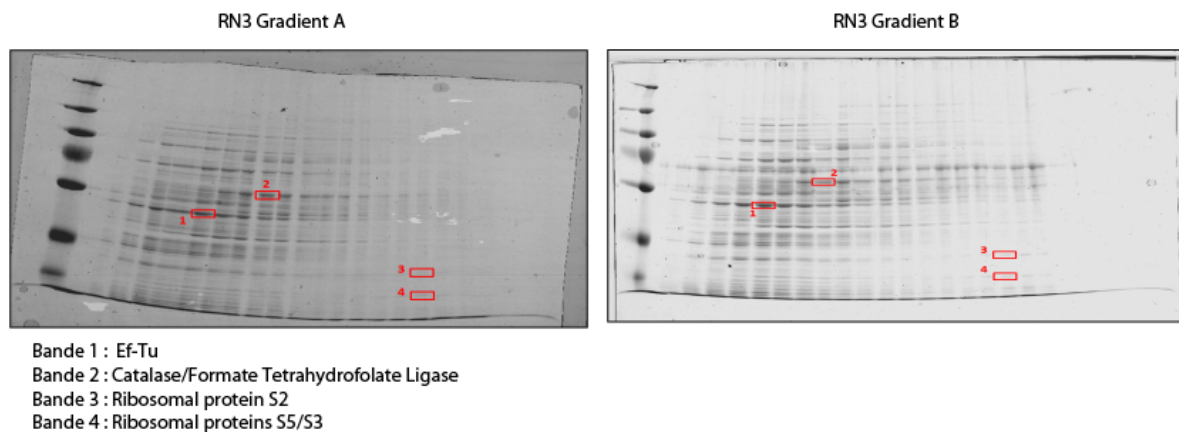
J'ai procédé à une vérification de la localisation de certains ARN d'intérêt dans ces gradients par Northern blot, de la même façon que décrit précédemment (§). Les ARN choisis cette fois sont l'ARNIII, RsaA, l'ARNr 5S et aussi l'ARN 4,5S qui est connu pour interagir avec la RNase P lors de sa maturation (Spitzfaden *et al.*, 2000). J'ai de manière parallèle déposé une même quantité de protéines de chaque fraction sur un gel dénaturant de polyacrylamide 12% que j'ai ensuite coloré au bleu de coomassie. En comparant les gels en fonction des deux souches, j'ai choisi 5 bandes situées dans les mêmes fractions qui ont été analysées par spectrométrie de masse pour vérifier qu'elles correspondent bien aux mêmes protéines.



**Figure 42 : Vérifications de la reproductibilité des gradients de glycérol par Northern blot** - Les fractions de chaque gradient sont numérotées de 1 (fractions légères) à 20 (fractions lourdes). Chaque ARN a été détecté par une sonde DIG-ARN spécifique. La même membrane a été utilisée pour la détection de tous les ARN.

La comparaison des expériences northern blots effectuées sur les deux différents gradients nous montre que l'approche que j'ai utilisée, malgré le changement de moyen de récolte des fractions, est toujours très reproductible, car les ARN présentent les mêmes profils de migration dans les deux cas (**figure 42** ci-dessus). On remarque à nouveau une répartition assez homogène pour l'ARNr 5S, avec une plus grande présence dans les fractions légères. Un profil similaire est observé pour l'ARN 4.5S. De même, on retrouve l'ARNIII fortement enrichi dans les fractions les plus lourdes du gradient, ce qui peut être dû soit à sa taille importante (513 nucléotides, ce qui n'est pas éloigné du RsaC

que j'ai transcrit *in vitro* est qui est capable d'aller dans le fond du gradient juste par sa taille) mais aussi par son interaction avec les ribosomes de *S. aureus* puisqu'il a une fonction d'ARNm. Enfin, RsaA présente à nouveau deux pics d'enrichissements, un présent dans les fractions légères, un second présent dans les fractions lourdes. Comparé au premier northern blot que je montrais en **figure 39**, on voit une distance plus importante entre les deux pics d'enrichissement, ce qui est probablement dû au plus grand nombre de fractions (20 ici contre 13 auparavant) que j'obtiens avec la collecte du gradient par l'AKTA®.

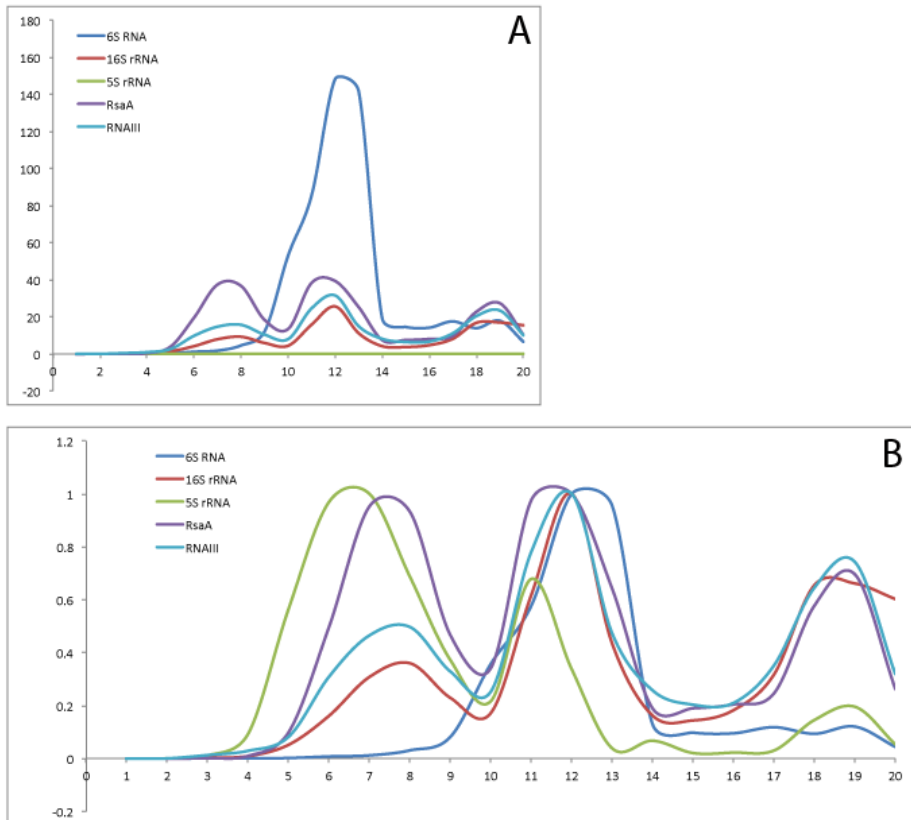


**Figure 43 : Vérifications de la reproductibilité des gradients de glycérol SDS-PAGE et spectrométrie de masse** – Une même quantité de protéines pour chaque fraction du gradient a été déposée sur gel dénaturant de polyacrylamide 12%-SDS. La coloration des gels est réalisée au bleu de coomassie, les bandes indiquées en rouge ont été découpées et séquencées par MS comme indiqué précédemment dans le manuscrit. L'identification des bandes est donnée en dessous des gels.

Concernant les protéines analysées par spectrométrie de masse, les résultats ont là aussi été concluants sur la reproductibilité de mes gradients (**figure 43**). Pour chacune des bandes découpées et séquencées par spectrométrie de masse, j'ai retrouvé les mêmes protéines en *first hit* ce qui, combiné à l'aspect extrêmement semblable de la répartition des protéines que l'on peut voir sur les gels colorés au bleu de coomassie, nous indique que les gradients sont reproductibles ainsi que la répartition des ARN et des protéines dans chacune des fractions. J'ai donc envoyé les ARN totaux de chaque fraction du gradient A à l'équipe de J. Vogel qui ont procédé à leur séquençage par RNA-seq avec la société Vertis.

Pour pouvoir vérifier l'enrichissement des ARN en fonction des fractions, les *reads* obtenus ont d'abord été normalisés par les niveaux d'ARN *spike-in* (ce sont des ARN dont la séquence et l'expression sont parfaitement connus dans les conditions utilisées, rajoutés à l'expérience comme contrôle interne) (Wang *et al.*, 2009). Pour obtenir une

mesure de l'abondance de chaque ARN sur une même échelle, il est également possible de diviser l'abondance de l'ARN dans chaque fraction par son maximum d'abondance présent sur l'ensemble des fractions. De cette façon, nous obtenons ensuite des valeurs comprises entre 0 et 1 pour chaque ARN, ce qui permet de les comparer entre eux. Les opérations de normalisation ont été réalisées par A. Smirnov.



**Figure 44 : Visualisation de l'abondance d'ARNnc de *S. aureus* par quantification et normalisation des données de RNA-seq** – A. Les données ont d'abord été normalisées avec l'ajout dans le set de données des ARN spike in. B. Les mêmes données sont ici normalisées par le maximum d'abondance de chaque ARN.

La visualisation d'ARNnc connus de *S. aureus* permet également de vérifier la « bonne conduite » du gradient effectué. En **figure 44A**, l'enrichissement de l'ARN 6S dans les fractions 10 à 14 du gradient est signe de la présence de l'ARN polymérase et corrèle avec le pic d'absorbance à 280 nm observé sur les profils de la **figure 41A**. On peut observer que chaque ARN présente une légère accumulation dans ces mêmes fractions, ce qui est attendu puisque chaque ARN est naturellement transcrit au cours de la croissance. Il vaut mieux par contre se référer à la **figure 44B** pour avoir une meilleure idée des enrichissements de chaque ARN puisqu'ici les valeurs sont normalisées par le maximum d'abondance pour chaque ARN donné. On constate que le pic de l'ARN 6S censé être avec l'ARN polymérase est décalé par rapport aux autres ARN, et ceci indique

un problème technique, très certainement lié à une dégradation des ARN. De même, les profils de migration dans la partie lourde du gradient ne sont pas aussi sinusoïdes que dans la partie légère. En ajoutant le fait que la durée de la centrifugation utilisée empêche de voir les ribosomes 70S ainsi que les polysomes, le protocole que j’ai utilisé est mieux applicable pour étudier les complexes de faible poids moléculaire. Il devra être ajusté pour visualiser les polysomes et autres complexes de fort poids moléculaire.

Une autre façon de regarder ces données, très visuelle également, est simplement d’appliquer un code de couleurs en fonction des abondances des ARN, ce qui permet de voir rapidement dans quelle fraction un ARN est enrichi comme le montre la **figure 45** ci-dessous. Je préfère préciser malgré tout que les résultats de RNA-seq constituent à l’heure actuelle des résultats préliminaires ; l’analyse est toujours en cours et se heurte notamment à des problèmes d’annotations de génome.

	Fractions																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
ARN 6S	7.8847E-05	0.00035466	0.02996682	0.05209375	0.26307917	1.12111806	1.73928414	4.54693515	12.1018675	53.047942	85.6809567	148.427197	141.723901	19.111751	14.4999959	14.1323085	17.5348658	1.38757919	17.8712278	6.52910833
rsaA	0.00033427	0.00166694	0.04704002	0.37374375	3.88416042	19.5342049	37.4377476	36.6299045	18.3039367	13.4222265	38.34745	39.293225	25.205581	7.41970333	7.45610146	8.027125	9.66654444	2.2692629	27.3391	10.3173875
rsaF	9.7864E-05	0.00021094	0.12241811	0.30205625	1.4038	2.09910417	2.66391548	2.59320403	1.76843061	1.19336835	3.7692	4.725275	1.31605485	1.14854	0.78495386	0.85925625	1.38948333	0.29652668	2.9667	0.9828625
rsaG	1.3072E-05	9.0233E-05	0.00282033	0.0060375	0.03386667	0.14918667	0.30771982	0.4937139	0.45193585	0.23670387	0.78306667	1.12073333	0.3123466	0.17175	0.09165507	0.06983646	0.09515556	0.0442221	0.16076667	0.07005833
rsaA	0.00033427	0.00166694	0.04704002	0.37374375	3.88416042	19.5342049	37.4377476	36.6299045	18.3039367	13.4222265	38.34745	39.293225	25.205581	7.41970333	7.45610146	8.027125	9.66654444	2.2692629	27.3391	10.3173875
rsaH	0.00042455	0.00181819	0.07152246	0.30105625	1.03196667	3.66097917	5.49323209	5.01567219	2.72031274	1.74965346	3.9034	3.7277	1.39104685	0.72571	0.54163122	0.5034	0.738775	0.14036429	1.53	0.6383375
Sau-5971	0.00254528	0.15296267	10.1520599	21.0266208	39.6494842	119.508437	118.254264	103.830483	47.0047375	26.28207	90.5107333	66.32605	17.5582384	10.42018	5.29736557	4.4617042	5.27806111	1.07303611	17.6434667	6.88730833
ssrS	7.6847E-05	0.00035466	0.02996682	0.05209375	0.26307917	1.12111806	1.73928414	4.54693515	12.1018675	53.047942	85.6809567	148.427197	141.723901	19.111751	14.4999959	14.1323085	17.5348658	1.38757919	17.8712278	6.52910833
RNAIII	0.00083924	0.0094946	0.40213742	0.9133375	2.57828239	9.74019444	14.8585647	15.6566479	10.3518001	7.96487196	24.058333	31.51535	14.8967026	8.10974667	6.40650616	6.61740417	11.077375	2.0331053	23.4048	10.6537793
rsaK	3.9175E-05	0.00071654	0.02814	0.06725	0.24926667	0.78783333	1.02161472	0.87192775	0.49923816	0.2441818	0.63	0.7286	0.20978618	0.1104	0.06859308	0.05055	0.06956667	0.00951684	0.1364	0.060215
rsaA	0.00033427	0.00166694	0.04704002	0.37374375	3.88416042	19.5342049	37.4377476	36.6299045	18.3039367	13.4222265	38.34745	39.293225	25.205581	7.41970333	7.45610146	8.027125	9.66654444	2.2692629	27.3391	10.3173875
rsaH	0.00042455	0.00181819	0.07152246	0.30105625	1.03196667	3.66097917	5.49323209	5.01567219	2.72031274	1.74965346	3.9034	3.7277	1.39104685	0.72571	0.54163122	0.5034	0.738775	0.14036429	1.53	0.6383375
rsaE	9.7864E-05	0.00021094	0.12241811	0.30205625	1.4038	2.09910417	2.66391548	2.59320403	1.76843061	1.19336835	3.7692	4.725275	1.31605485	1.14854	0.78495386	0.85925625	1.38948333	0.29652668	2.9667	0.9828625

Figure 45 : Visualisation de l’abondance d’ARNnc de *S. aureus* sur chaque fraction du gradient de glycérol – Chaque chiffre représente l’abondance du même ARN pour chaque fraction. Le code couleur indique une abondance très forte (rouge) à très faible (vert foncé).

Dans ce tableau on peut clairement retrouver l’enrichissement dans les fractions 10 à 13 de l’ARN 6S déjà observé auparavant, et également constater que la majorité des ARNnc choisis dans cette figure ont également une abondance plus forte dans ces fractions. Certains sont également présents dans les fractions lourdes du gradient de glycérol, ce qui sous-entend leur interaction avec des complexes de taille plus importante. D’autres se cantonnent aux fractions légères, ce qui démontre encore une fois la spécificité de la localisation de chaque ARNnc dans cette approche. J’aimerais en dernier lieu attirer l’attention sur deux détails : la première est l’abondance très forte d’un petit ARN qui a un enrichissement très important dans les fractions légères (6 à 8) en plus de présenter une abondance dans toutes les fractions bien plus fortes que pour les autres ARNnc. De façon curieuse, cet ARN annoté Sau5971 a été identifié auparavant : son expression est notamment constitutive chez une souche de *S. aureus* présentant un phénotype de *Small*

*Colony Variant* (Abu-Qatouseh *et al.*, 2010) – un phénotype généralement associé avec la persistance, la résistance aux antibiotiques et les infections chroniques. En regardant sa localisation sur le génome de HG001 avec le programme Integrative Genomics Viewer (Robinson *et al.*, 2011), j'ai retrouvé la séquence correspondante de Sau-5971 dans une région intergénique, située en 3' de la partie codante du gène HG001\_00312 (dont le produit est de fonction encore inconnue). Il pourrait donc, dans cette souche, s'agir d'un ARNnc situé dans la région 3'UTR de ce gène, ce qui soulève une certaine curiosité puisque nous étudions dans l'équipe un ARNnc, RsaG, qui est lui aussi situé dans la région 3'UTR d'un gène (*uhpT*) – et pourrait donc ne pas être le seul dans ce cas chez *S. aureus*, puisqu'une étude récente a montré que près de 35% des ARNm de *S. aureus* possèdent des régions 3'UTR longues de plus de 100 nucléotides, qui peuvent être impliquées dans des processus de régulation à plusieurs niveaux, comme c'est le cas pour l'ARNm *icaR* dont la partie 3'UTR contient une séquence anti-SD qui va s'apparier avec la séquence SD du messenger pour inhiber sa traduction dans un premier temps, et entraîner la dégradation de l'ARN par la RNase III (De Los Mozos *et al.*, 2013).

L'autre point que je souhaite soulever par la **figure 45** est la présence en double de certains ARN (RsaA, l'ARN 6S, ou encore RsaH) – et ceci provient directement de problèmes d'annotation des gènes dans les différentes bases de données utilisées pour aligner les *reads* (nous disposons du génome séquencé de HG001, pas de HG003, et cette première analyse a été faite en utilisant le génome d'une souche voisine, N315 – dont l'annotation laisse vraisemblablement à désirer). Ce qui soulève l'une des difficultés dans l'analyse des données de RNA-seq qui reposent sur les annotations de génomes qui peuvent être incomplètes ou erronées. L'analyse de ces données de RNA-seq pourra être ré-effectuée en utilisant notre génome séquencé de HG001 ce qui permettra d'éviter ces doublons.

**En me basant sur les travaux de Wassarman & Storz, j'ai réussi à appliquer l'approche de séparation des complexes ribonucléoprotéiques par centrifugation sur gradient de glycérol à *S. aureus*. J'ai vérifié la reproductibilité de cette approche ainsi que son efficacité en utilisant l'ARN 6S comme preuve de concept. J'ai également montré que cette méthode permet de suivre les duplexes ARNnc-ARNm stables formés *in vivo* chez la bactérie, ce qui est l'un de ses avantages. La**

**combinaison du gradient aux techniques de séquençage à haut débit telles que la LC MS/MS et le RNA-seq devrait permettre de donner une idée des interactions ARN/ARN et ARN/protéines qui se forment *in vivo*, à grande échelle, chez *S. aureus*.**

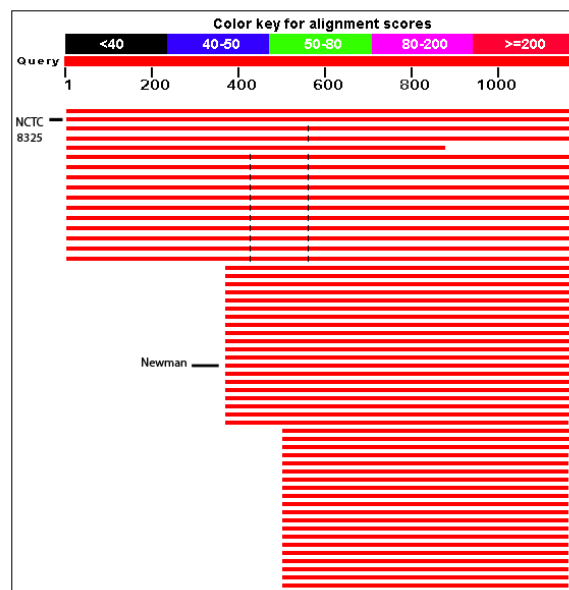
# DISCUSSION ET PERSPECTIVES

# Partie 1 : RsaC, un ARN aux propriétés intrigantes

Au cours de ces quatre années de thèse je me suis attelé à résoudre la fonction de l'ARNnc RsaC, découvert en 2009 dans l'équipe par Geissmann *et al.* Le peu de données présentes à mon arrivée au laboratoire m'ont demandé de partir à zéro avec cet ARN, et la première surprise est venue de la découverte des régions répétées présentes dans la partie 5' de cet ARN, dont le nombre diffère en fonction des souches.

## 1. Quelle est la fonction des régions répétées ?

RsaC est un ARN conservé uniquement chez *S. aureus* et pas chez les autres espèces du genre *Staphylococcus*. Le principe de conservation des séquences chez les espèces vivantes veut qu'une séquence conservée soit synonyme de fonction biologique à préserver pour l'organisme. On constate en revanche que, si les régions répétées sont à chaque fois présentes chez RsaC, leur nombre est souche-dépendant. Une façon très visuelle de voir cette différence est de simplement « aligner » la séquence complète de RsaC de HG001 sur les génomes des autres souches référencées de *S. aureus* via l'outil informatique BLAST (NCBI).



**Figure 46 :** Résultat de BLAST sur la séquence complète de RsaC – Le score d'alignement de séquence est indiqué par un code couleur en haut de la figure. La longueur des alignements est alignée sur la séquence donnée.



La différence de longueurs observée se rapporte directement aux résultats de la **figure 17**. On constate qu'il existe trois grandes populations de RsaC, la plus longue étant présente chez HG001, avec deux populations de taille inférieure qui varie par le nombre de régions répétées dans leur partie 3'. Pour expliquer l'origine de ces répétitions, il est possible qu'au cours de la réplication du génome, l'ADN polymérase ait butée sur la région répétée et ait polymérisé la séquence plusieurs fois. A l'inverse il pourrait également s'agir de régions qui ont été perdues chez certaines souches. Dans les deux cas il s'agirait d'évènements séparés qui se sont produits au cours de l'évolution et qui ont conduit à ces différences observées aujourd'hui entre les souches de *S. aureus*. Ces régions qui forment des tige-boucles au niveau de l'ARN font penser à certains éléments répétés présents chez d'autres bactéries comme les séquences REP-2 (pour *Repetitive Extragenic Palindromic*) identifiées chez plusieurs espèces du genre *Neisseria* (Delihias, 2011). Ce type de répétition fait entre 120 et 150 nts et se retrouve dans des régions intergéniques. Chez *Neisseria*, elles permettent notamment de contrôler l'expression de gènes impliqués dans l'interaction de la bactérie avec les cellules de l'hôte infecté. Mais leur nombre de répétitions (de 23 à 26) est considérablement plus important que chez *S. aureus* pour RsaC. Chez d'autres bactéries, les régions répétées ont également un rôle dans la plasticité du génome, en servant de site d'intégration pour les IS comme c'est le cas chez *E. coli* ou *S. enterica* (Delihias, 2011). Je n'ai pas observé d'intégration de telles séquences chez les différentes souches de *S. aureus* mais les séquences pourraient avoir un rôle à jouer dans le contrôle de l'expression de l'ARN mais aussi dans l'évolution du génome de *S. aureus*.

Ceci pose malgré tout légitimement la question de leur utilité. En supposant que la fonction de RsaC soit la même chez toutes les souches de *S. aureus*, on peut se demander pourquoi l'évolution a permis la conservation de ces différences dans le nombre de régions répétées. Ont-elles réellement une fonction pour l'ARN ? Servent-elles de motif purement structurel pour stabiliser ou entraîner la maturation de l'ARN ? Sont-elles reconnues par d'autres facteurs, acides nucléiques ou protéines, cellulaires ?

Outre les conséquences évolutives, les motifs répétés peuvent avoir différentes conséquences fonctionnelles chez les bactéries. Il existe notamment chez *S. aureus* de petites séquences riches en G-C répétées appelées séquences STAR (pour *S. aureus*

*repeats*) (Cramton *et al.*, 2000). Longues de 14 nts, ces répétitions se trouvent dans des régions intergéniques à des *loci* conservés entre les différentes espèces de *S. aureus* ; elles pourraient notamment jouer un rôle dans la pathogenèse car on les retrouve en plus forte abondance chez les deux souches de *S. aureus* les plus virulentes connues (Purves *et al.*, 2012). D'autres ARN possèdent des répétitions de motifs reconnus spécifiquement par des protéines, mimant ainsi les cibles de ces dernières. C'est notamment le cas pour les ARNnc CsrB et CsrC qui possèdent chacun 22 et 13 motifs GGA respectivement qui sont reconnus par la protéine CsrA. La répétition de ces motifs permet de piéger la protéine CsrA et de l'empêcher d'assurer son rôle de régulation sur ses cibles ARN qui présentent le même motif (Romeo *et al.*, 2013). Cependant ces séquences répétées de RsaC sont plus importantes qu'un motif de 3 nucléotides et forment des tiges-boucles successives (**figure 24**), formant notamment trois structures « étoilées » à quatre branches dans la partie 3'. On pourrait s'imaginer que ces structures soient reconnues par des protéines mais l'expérience de MAPS avec la forme longue de RsaC n'ayant pas fonctionné il n'est pas encore possible de conclure sur cette hypothèse. L'expérience devra être reconduite et réalisée en même temps qu'avec la forme courte (544 nts) de RsaC pour comparer si des protéines sont spécifiquement enrichies en présence des régions répétées (voir §5 discussion).

## 2. L'ARNm SodA est-il une cible directe de RsaC ?

RsaC présente dans sa partie 5' des motifs riches en C qui sont également répétés et accessibles puisqu'ils se situent dans des régions en simple brin. Ces motifs sont présents chez d'autres ARNnc de *S. aureus* comme l'ARNIII, RsaA, RsaE, mais aussi RsaG. Pour les trois premiers, ces motifs riches en C s'apparient avec la séquence SD d'ARNm cibles pour réguler leur traduction et induire ou non une dégradation des ARNm réprimés. De façon surprenante, il a été montré dans l'équipe que RsaG s'apparie avec deux motifs riches en G présents sur un autre ARNnc, RsaI, conduisant certainement à un contrôle d'un ARN sur l'action de l'autre (Desgranges E. et Bronesky, D., résultats non montrés).

Ainsi, les motifs riches en C des ARNnc de *S. aureus* pourraient constituer une signature d'un motif d'appariement avec un ARNm. Comme j'ai retrouvé l'ARNm *sodA* dans mes

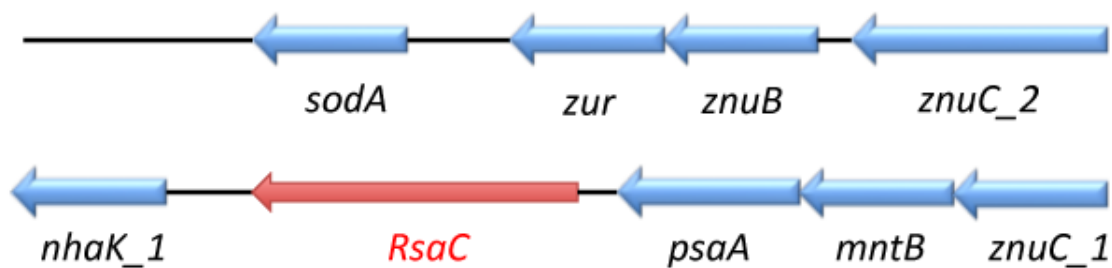
expériences de MAPS (**tableau 3**) et que le niveau de la protéine est également perturbé chez le mutant  $\Delta$ *rsaC* comme vu en protéomique comparative (**tableau 2**) alors que le niveau de l'ARNm n'est pas affecté (**figure 35**), il est probable que l'un des motifs riche en C puisse reconnaître la séquence SD de l'ARNm *sodA* pour réguler sa traduction (dans notre hypothèse, il s'agirait d'une inhibition). La prédiction d'appariement présentée en **figure 33** se trouve en amont de la séquence SD mais nous n'avons pour le moment aucune vérification expérimentale. Pour vérifier que l'interaction est directe la technique de retard sur gel pourrait être utilisée en utilisant *sodB* marqué et des concentrations croissantes de RsaC (version longue et courte) comme décrit pour RsaA (manuscrit 2, A. Tomasini et al.). Cependant, la taille de RsaC étant relativement grande, la cartographie en solution serait sans doute plus appropriée pour définir la ou les régions de RsaC dont la réactivité des nucléotides changerait en présence de l'ARNm.

Un autre résultat interpellant est la présence dans les données de protéomiques de la protéine PsaA, dont le gène se trouve juste en amont de *rsaC*. Nous avons vérifié par RT-qPCR que la délétion du gène *rsaC* n'a aucun effet sur la transcription de *psaA* (**figure 27B**); l'effet observé lors de l'étude de protéomique est donc post-transcriptionnel. Mais à l'inverse de SodA et des autres protéines retrouvées enrichies en absence de RsaC, PsaA est moins abondante dans ces conditions, ce qui sous-entendrait que RsaC pourrait avoir un effet activateur sur PsaA. J'ai aussi montré que RsaC est stabilisée par l'action de la RNase III. En d'autre terme aurions nous besoin d'une maturation spécifique de RsaC pour induire cet effet stabilisateur sur *psaA* ?

La protéine PsaA est un transporteur ABC des ions  $Mn^{2+}$  et  $Zn^{2+}$  - qui utilise donc l'énergie dégagée par l'hydrolyse d'un ATP pour faire rentrer ces ions dans la cellule. Ceci est intéressant car SodA est une superoxide dismutase qui a besoin d'ions  $Mn^{2+}$  pour fonctionner correctement; la présence de ces ions augmente de plus son activité, mais il a également été démontré que ces ions par eux-mêmes ont des propriétés de protection contre le stress oxydatif (Karavolos *et al.*, 2003). Les données de protéomique ont été obtenues sur des cultures de *S. aureus* récupérées à 6h, en phase stationnaire de croissance. Il a été montré qu'en conditions de laboratoire, SodA est la plus exprimée en milieu de phase exponentielle à 4h de croissance alors qu'à 6h son expression est diminuée (Karavolos *et al.*, 2003). Il n'y a donc pas de contradiction à

supposer qu'en phase stationnaire, RsaC puisse avoir un rôle inhibiteur sur la traduction de l'ARNm *sodA* et en même temps active l'expression de *PsaA* pour moduler la réponse de la bactérie à résister à l'accumulation des ROS durant la phase stationnaire de croissance, en mêlant l'action de protéines et l'équilibre homéostatique. En revanche, l'ARNm *psaA* n'a pas été retrouvé enrichi dans notre expérience de MAPS, ce qui peut supposer deux choses : soit la régulation post-transcriptionnelle observée en protéomique n'est pas due à une interaction directe entre les deux ARN, ou bien si cette interaction est directe, elle pourrait impliquer les régions répétées – ou la partie 5' de *RsaC* – qui n'était pas présente sur notre construction utilisée en MAPS.

Il est intéressant de mentionner ici que j'ai aussi retrouvé l'ARNm *ZnuC\_2* dans cette expérience MAPS. Ce gène fait partie du système *ZnuABC* où *znuC\_1* est transcrit en opéron avec *psaA*. La protéine *ZnuC\_1* est impliquée dans l'utilisation des ions  $Zn^{2+}$ , qui sont les autres ions également capables d'être transportés par *PsaA*. La régulation observée pour *PsaA* peut être donc due directement à un effet de *RsaC* sur cet opéron.



**Figure 47 :** Représentation schématique des loci de *sodA* et *psaA* chez *S. aureus* HG001 – Les ORFs sont indiquées en bleu, les ARNnc en rouge.

J'ai constaté de plus que de nombreux ARNm codant pour des protéines dont la fonction sont le transport et l'utilisation des ions  $Zn^{2+}$  ont été retrouvés enrichis par MAPS. Il faut évidemment valider ces cibles, qui ne sont pas toutes les plus enrichies de l'expérience, mais cela indique clairement une tendance de *RsaC* dans le métabolisme des ions métalliques (en l'occurrence, zinc, fer et manganèse) qui sont importants dans de nombreux processus biologiques, et notamment la résistance au stress oxydatif. Ce qui m'amène au point suivant.

### 3. RsaC est-il vraiment impliqué dans la résistance au stress oxydatif ?

#### a. Effet de RsaC sur l'expression de SodA et lien avec le stress oxydatif

Si les résultats de MAPS et de la protéomique, concernant d'une part l'enrichissement de l'ARNm avec RsaC *in vivo* et la surabondance de la protéine en l'absence de l'ARN, ont été validés avec une p-valeur qui les rendent significatifs, les expériences pour valider l'action de RsaC sur SodA sont bien plus discutables. Comme on le voit premièrement dans la **figure 34** lorsqu'on mesure les croissances des souches sauvages, mutantes et complémentées pour RsaC, ainsi que leur viabilité cellulaire, on n'observe de différence que chez la souche complémentée, lorsque RsaC est surexprimé. La souche sauvage et la souche  $\Delta$ rsaC semblent se comporter de la même façon. Il se pourrait que dans les conditions de nos études, la concentration intracellulaire de RsaC ne serait pas suffisante pour obtenir un phénotype clair. On retrouve ce même constat en **figure 35**. En absence de MV, le niveau d'expression de SodA est le même chez la souche sauvage et la souche mutante, et n'a de variation que chez la souche complémentée. En présence de MV, le niveau d'expression de SodA est plus important qu'en son absence, mais on n'observe pas à nouveau de différence significative d'expression de la protéine entre la souche sauvage et la souche mutante. Seule la souche complémentée présente une expression moins forte de SodA, allant dans le sens que RsaC, lorsqu'il est surexprimé et en présence de MV, entraîne une inhibition de l'expression de SodA.

De même, lorsqu'on regarde l'expression de RsaC on constate que celle-ci est fortement induite dans la souche complémentée en présence de MV (plus de dix fois comparé à la souche sauvage dans la même condition). Pourtant l'effet observé sur l'expression de SodA n'est pas proportionnel à cette surexpression de RsaC. On peut donc se demander si l'effet que nous voyons n'est pas accentué par une trop forte concentration de RsaC, qui ne correspondrait pas à une concentration biologiquement possible dans l'environnement naturel de la bactérie. A l'inverse, l'expression de RsaC n'est peut-être pas suffisante dans nos conditions de croissance en laboratoire pour observer l'effet sur SodA et, par extension, un véritable phénotype lié à cet ARN. Même si mes expériences

de protéomiques réaliées dans ces conditions m'ont malgré tout déjà montré un effet (**tableau 2**).

Lorsqu'on teste directement l'activité de SodA comme montré dans la **figure 36A**, on n'observe pas de différence entre l'activité de la protéine chez la souche sauvage et la souche mutante, en l'absence et en présence de MV, alors que nous pouvions nous attendre à avoir une activité plus importante de l'enzyme dans la souche  $\Delta$ rsaC. A nouveau, le seul effet observé se fait chez la souche complémentée, lorsque RsaC est surexprimé, ce qui renvoie à la pertinence de cette surexpression ; une façon de pallier à ce problème sera de déterminer la séquence promotrice de RsaC et de cloner à nouveau l'ARN avec sa séquence promotrice endogène dans le pCN51. Mais malgré cela, on ne pourra pas exclure une surexpression due au nombre de copies de plasmide dans la cellule, et il faudrait soit passer par un plasmide à faible nombre de copies, ou simplement réaliser la complémentation directement dans le chromosome bactérien. Le choix d'avoir utilisé des plasmides vient également de la difficulté à réaliser ce genre de manipulations génétiques chez *S. aureus* (j'ai passé moi-même plusieurs mois à essayer d'obtenir un mutant de *rsaC* par différentes approches – double recombinaison par utilisation du 5FO (Monk *et al.*, 2012) ou du système Cas9/CRISPR - avant de laisser l'équipe de F. Vandenesch réaliser ce mutant). Les régions répétées dans RsaC ont été effectivement gênantes pour la recombinaison génétique nécessaire pour déléter le gène dans le chromosome bactérien. Enfin, je voudrais ajouter que les résultats présentés en **figure 35** sont issus d'expériences de RT-qPCR qui sont très sensibles et sujettes à des variations qui sont souvent importantes (voir les barres d'erreur pour SodA et SodM). De même, seul le gène *gyrB* a été utilisé pour la normalisation des quantifications, il ne serait peut-être pas inutile d'utiliser un second gène de ménage pour cette normalisation.

Au contraire des précédentes expériences, celles présentées en **figure 36B** vont, elles, dans le sens de notre hypothèse pour le rôle de RsaC. Lorsqu'on mesure la concentration de ROS dans la cellule on constate que celle-ci diminue chez la souche  $\Delta$ rsaC et qu'elle est plus importante que dans la souche sauvage et la souche complémentée. On sait que l'activité de SodA est de réduire les ROS, et notamment les ions superoxydes en O<sub>2</sub> et H<sub>2</sub>O, ce qui entraîne donc une diminution de leur concentration intracellulaire. Si RsaC

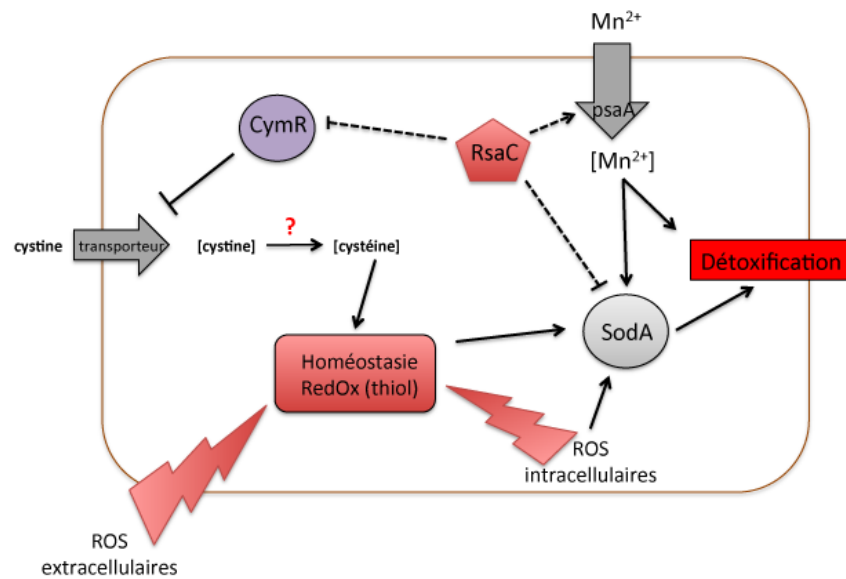
entraîne bien une inhibition de l'expression de SodA, alors il n'est pas surprenant de voir la concentration de ROS diminuer en absence de RsaC (puisque l'inhibition de SodA sera donc levée). De même, l'augmentation de leur concentration devrait avoir lieu dans la souche complémentée, lorsque RsaC, surexprimé, inhibe plus fortement l'expression de SodA. En définitive, si le rôle de RsaC dans la réponse au stress oxydatif n'est pas encore clairement défini, les résultats que j'ai obtenus convergent vers une implication de RsaC dans ce processus et une régulation de l'expression de SodA, probablement au niveau post-transcriptionnel.

#### *b. Autres cibles potentielles de RsaC*

Même si je n'ai pas eu le temps de faire toutes les validations nécessaires, les résultats de l'expérience de MAPS avec RsaC (**tableau 3**) montrent, parmi tous les ARNm qui ont été retrouvés et séquencés, des interconnexions plus ou moins fortes entre certaines fonctions biologiques.

Ainsi, j'ai déjà parlé des ARNm *zur*, *znu\_C* et *znu\_B* qui sont impliqués dans le transport et l'utilisation du zinc, et qui constituent des cibles potentielles intéressantes de RsaC puisque les trois sont retrouvés en opéron et au même locus que le gène *sodA*. On constate également plusieurs ARNm qui sont impliqués dans l'équilibre oxydo-réduction (redox) de la bactérie comme *feoD*, *rex*, *scdA* ou encore *cymR*. Ce dernier est un facteur de transcription qui a été démontré comme capable de sentir l'équilibre redox au niveau des cystéines (le groupement thiol est oxydé ou réduit en fonction de cet équilibre), ce qui lui permet en cas de stress oxydatif d'activer la transcription de gènes impliqués dans la réponse à ce stress ainsi que d'autres impliqués dans la synthèse et l'import de cystéines (Ji *et al.*, 2012). Ainsi ce facteur exerce une fonction clé dans le maintien de l'homéostasie redox. La délétion de *cymR* augmente la survie de la bactérie dans les macrophages mais atténue la virulence dans un modèle murin (Soutourina *et al.*, 2010). Ceci est intéressant puisque je retrouve dans les données de protéomique (**tableau 2**) la protéine YraA qui est une cystéine protéase putative, alors que je retrouve aussi en **tableau 3** les ARNm *rex* et *ywgN* qui codent tous les deux des protéines impliquées dans l'équilibre redox. A nouveau, ces résultats sont encore trop préliminaires pour pouvoir avancer des conclusions mais ils laissent entrevoir des tendances dans la fonction de

RsaC, qui se recoupe toutes dans l'adaptation au stress oxydatif et le métabolisme des ions métalliques (voir **figure 48** ci-dessous).



**Figure 48 : Modèle schématique du rôle possible de RsaC chez *S. aureus*** – Ce modèle est purement théorique et basé sur les résultats préliminaires que j'ai obtenus. Une flèche indique une activation alors que les barres représentent une inhibition. Les pointillés indiquent un résultat non validé entièrement. Adapté de Soutourina *et al.*, 2010.

#### 4. RsaC et persistance

Ce recoupement de fonctions cellulaires pourrait par ailleurs avoir un effet physiologique sur la résistance à l'internalisation de la bactérie par un type cellulaire particulier : les ostéoblastes. J'avais mentionné en introduction que *S. aureus* peut être internalisée par différents types cellulaires ; les ostéoblastes sont les cellules chargées de la construction du tissu osseux et se trouvent à sa surface. Leur capacité d'internalisation des bactéries permet de lutter contre les infections bactériennes au niveau de ce tissu, appelées ostéomyélites. Une étude récente de transcriptomique a montré que RsaC est l'ARNnc de *S. aureus* le plus fortement induit lors d'une infection des ostéoblastes (Szafranska *et al.*, 2014). Pour vérifier si l'expression de cet ARN peut jouer un rôle dans la survie de *S. aureus*, plusieurs expériences dont les résultats sont présentés en **figure 37** et **figure 38** ont été réalisées. La première a permis simplement de confirmer ce qui avait été observé par Szafranska *et al.*, puisque nous avons retrouvé RsaC comme étant un des ARNnc le plus exprimé dans les ostéoblastes – du moins, parmi tous les autres ARNnc testés dans cette expérience (certains ne sont pas montrés sur la figure). La **figure 38** apporte des données plus troublantes. Dans un premier



temps, on peut constater que la cytotoxicité des cellules est augmentée en absence de RsaC, ce qui irait vers un rôle atténuateur de virulence comme peut le faire l'ARNnc RsaA (Romilly *et al.*, 2014). En revanche, il faut bien prendre en compte sur la **figure 38A** l'écart-type important des données dans chaque condition, et l'absence de données statistiques sur ces résultats puisqu'une seule expérience a été réalisée. Il faut donc être prudent sur ces résultats préliminaires. Pour la seconde partie des expériences, en revanche, si les valeurs présentent aussi un écart-type assez important, les barres des médianes sont ici bien plus éloignées, ce qui sous-entend une différence significative (l'expérience sera bien sûr renouvelée). Dans ce cas, l'absence de *rsaC* entraîne une meilleure persistance de *S. aureus* dans les ostéoblastes. La présence de RsaC diminuerait donc la survie des bactéries dans le milieu intracellulaire, ce qui pourrait être relié à une possible diminution de l'expression de SodA, entraînant alors une chute de la résistance au stress oxydatif causé par les ROS produits par les ostéoblastes. SodA est en effet tout aussi impliquée dans la résistance au stress oxydatif d'origine intracellulaire qu'extra-cellulaire (Karavolos *et al.*, 2003). On pourrait se poser la question de l'utilité pour *S. aureus* d'exprimer un ARN qui ne lui est pas favorable en condition de stress oxydatif (RsaC est l'ARNnc le plus fortement exprimé lors de l'internalisation **figure 37**), ce qui semble aller dans une contre logique. On peut également supposer que RsaC est un ARNnc dont l'action est nécessaire après l'adaptation au stress, ce qui implique une cinétique à prendre en compte dans son expression, et que je n'ai pas encore pu analyser par ces expériences.

En résumé, même si j'observe des tendances sur le rôle de RsaC, nos résultats ne sont pas aussi clairs que ceux observés pour RsaA. Je n'ai effectivement jamais pu voir de phénotype très marqué avec la souche  $\Delta$ *rsaC* – alors que beaucoup de tests physiologiques ont été réalisés par nos collaborateurs de Lyon, tels que des tests d'hémolyse, de formation de biofilms ou encore d'internalisation dans les monocytes (résultats non montrés). Il est fort probable que le véritable rôle de RsaC ne puisse être démontré dans les conditions de laboratoire que nous utilisons. Nous envisageons cependant de tester des modèles d'infections murins (infections aiguës ou chroniques) pour analyser un phénotype possible dans la virulence. Je n'exclue pas que l'action exercée par RsaC est de participer à une régulation fine en réponse au stress oxydatif ou que celui-ci agirait à une étape ultérieure du stress après détoxification des ROS. En

d'autre terme, il pourrait être un signal pour changer le comportement des cellules après adaptation au stress. Ainsi il sera important de déterminer les événements cinétique de la régulation et à quel moment la bactérie a besoin de réprimer l'expression de la protéine SodA *in vivo*. Comme je l'ai soulevé auparavant, RsaC pourrait également avoir un rôle sur l'expression de PsaA, une protéine impliquée dans le transport des ions manganèse. Il a été montré que la présence d'ions  $Mn^{2+}$  est suffisante en soi pour se protéger d'un stress oxydatif. RsaC pourrait donc d'un côté inhiber l'expression de SodA et de l'autre activer celle de PsaA pour protéger la bactérie contre un stress oxydatif qui n'implique pas la présence d'ion superoxydes, et contre lequel la production de la superoxyde dismutase n'est pas nécessaire. Cela pourrait aussi être une façon de limiter le nombre de molécules de SodA produites tout en assurant un influx d'ions  $Mn^{2+}$  pour assurer au mieux la processivité de l'enzyme. Il a en effet également été démontré que la présence des ions  $Mn^{2+}$ , en plus d'être nécessaire à l'activité de SodA, l'amplifient également (Karavolos *et al.*, 2003). Il sera nécessaire de réaliser des expériences supplémentaires, en prenant notamment en compte la région 5' de l'ARN et ses régions répétées pour évaluer leur rôle physiologique, mais nous ne sommes plus si loin de découvrir quels rôles exercent RsaC chez *S. aureus*.

## 5. RsaC, RNase III et protéines auxiliaires : quelles interactions ?

En cherchant à regarder la stabilité de RsaC *in vivo* chez *S. aureus*, j'ai pu démontrer un rôle important de la RNase III, codée par le gène *rnc*, pour cet ARN. Nous avons montré que la stabilité de RsaC est très fortement diminuée lorsque l'enzyme n'est pas exprimée (**figure 25**) mais également lorsque l'enzyme est présente mais n'est pas capable de couper les ARN tout en pouvant s'y lier (**figure 26**). Chez *S. aureus* d'autres ARNm comme *cspA* ont besoin de la RNase III pour leur maturation et leur stabilité (revue en annexe de Durand *et al.*, 2014), à l'inverse l'enzyme peut également dégrader les duplexes d'ARN qui impliquent un ARNnc et son ARNm cible, comme c'est le cas avec RsaA et MgrA (Romilly *et al.*, 2014). Nous avons démontré l'expression d'un ARN anti-sens à RsaC (**figure 27**) qui pourrait s'apparier dans la région 5' de l'ARN, possiblement en amont des premières régions répétées. Cependant, le niveau de cet ARN anti-sens par

rapport à RsaC est extrêmement faible et il n'est pas impossible que le transcrit que nous avons observé provienne d'un phénomène de transcription anti-sens de type pervasive comme décrite par Lasa *et al.*, 2011. Dans cette étude, les chercheurs ont montré qu'il existe chez *S. aureus* un phénomène de transcription généralisée du génome sur les deux brins, ce qui génère un très grand nombre de transcrits anti-sens qui sont rapidement dégradés, en particulier par la RNase III. Cette régulation dite chevauchante pourrait exercer un rôle dans la régulation post-transcriptionnelle, cette régulation étant dépendante des taux respectifs des ARNm et ARN anti-sens. RsaC n'est pas un ARNm en tant que tel (bien que deux ORF putatives aient été prédites dans sa séquence, voir **figure 17**) mais la transcription d'un ARN anti-sens, qui pourrait d'ailleurs être élevée dans des conditions que nous n'avons pas encore définies, pourrait contribuer à ajuster le niveau de RsaC.

Nos résultats ont montré que RsaC a besoin de l'action de coupure de la RNase III pour être stabilisé. Une telle interaction avec l'ARN anti-sens devrait dans ce cas être efficace puisque que RsaC s'accumule au cours de la croissance. L'expression de l'ARN anti-sens reste très faible quelque soit la phase de croissance. Comment la RNase III intervient-elle pour stabiliser RsaC ? Est-ce une action directe ou indirecte de la RNase ? Existe-t-il un processus de maturation de RsaC ? Ceci sous-entendrait qu'en absence de l'enzyme, nous observions un produit de transcription de taille plus importante, ce qui n'a pas été le cas dans la souche sauvage (**figure 25B**). En revanche j'ai observé dans la souche mutante  $\Delta rnc$ , la présence d'un transcrit faible mais reproductible qui est plus long par northern blot. Ce transcrit pourrait correspondre à un pré-ARN devant être mûré pour donner RsaC. La présence de cette bande sous-entend qu'il existe peut-être un « pré-RsaC » qui doit être mûré pour être l'ARNnc stable. Nos expériences de *primer extension* ont montré un signal fort qui pourrait correspondre au +1 de transcription de RsaC, mais il s'agit peut-être du premier nucléotide du transcrit mûré. Dans la région en amont de cette pause de la RT, je n'ai pas retrouvé de séquence consensus en -10 et -35. La recherche du promoteur endogène de RsaC sera donc essentielle pour vérifier si notre TSS est celui de l'ARN RsaC véritable, ou si celui-ci est transcrit en opéron avec *psaA* par exemple.

Enfin une dernière question se pose : existe-t-il des protéines (autre que la RNase III) qui sont capables d'interagir avec RsaC ? J'expliquais auparavant que RsaC, par la présence de ses motifs répétés (dans la partie 5' mais aussi les motifs riches en C dans la partie 3') pourrait séquestrer des protéines qui iraient reconnaître ces motifs spécifiquement. L'approche MAPS peut également être utilisée pour répondre à cette question, en allant séquencer les protéines présentes dans les éluats par spectrométrie de masse. Le seul problème pour l'instant est que je n'ai pas réussi à conduire cette approche avec la version longue de RsaC. J'avais conduit une expérience préliminaire avec la version courte qui m'avaient donné quelques protéines candidates dont la protéine ribosomale S1 – qui était par ailleurs aussi retrouvée enrichi avec MS2-RsaA, et MS2-RsaE (résultats non montrés). Mais je n'ai jamais pu reproduire ces résultats. La question de ces protéines interactantes reste donc encore entière pour RsaC, mais comme l'ARN est retenu sur la colonne de chromatographie d'affinité même en l'absence d'étiquette MS2 (**figure 30B**) il serait préférable d'envisager une autre approche pour cet ARN. Une alternative serait l'utilisation soit d'un ARN portant une biotine à l'une de ses extrémité ou d'un oligonucléotide biotinylé complémentaire à RsaC pour effectuer une chromatographie d'affinité basée sur l'interaction streptavidine-biotine.

Des expériences préliminaires ont été effectuées par Zongfu Wu qui a modifié l'extrémité 3' de RsaC<sub>L</sub> avec une biotine ainsi que sa version courte pour purifier *in vitro* les protéines associées à ces deux ARN en partant d'un extrait brut de *S. aureus* HG001. Cette stratégie semble être prometteuse puisque plusieurs protéines communes aux deux ARN RsaC ont été retenues. Ces protéines incluent une ARN hélicase de type DEAD box CshA, quelques RNases (RNases J1/J2, PNPase, RNase III), mais aussi des enzymes de modifications de l'ARN (pseudouridine synthase). Ces expériences encourageantes devront être renouvelées. La recherche des protéines interagissant à l'ARN par le biais de l'ARN est difficile sans doute à cause du fait que ces interactions sont transitoires. Ainsi il serait sans doute préférable d'utiliser des approches de pontage (UV ou par formaldéhyde) pour stabiliser ces interactions au cours du traitement. Il est cependant intéressant d'avoir retrouvé certaines des enzymes de modifications des ARN. La pseudouridine est une des bases modifiées les plus répandues et nous ne pouvons exclure que ces enzymes pourraient moduler l'activité, la stabilité ou bien encore la

structure des ARN (Marbaniang and Vogel, 2016). Une récente étude réalisée sur de nombreuses bactéries dont *S. aureus*, a montré que de nombreux ARN présentent des méthylations en position N6 de l'adénine. Un enrichissement a été observé pour les ARNm associés à la respiration, au métabolisme des acides aminés, à la réponse au stress et aussi dans divers sARN (Deng et al., 2015). L'analyse des bases modifiées et leurs fonctions est un nouvel aspect de la régulation de l'expression des gènes que nous ne devons plus négliger.

**RsaC est un ARN de *S. aureus* sur lequel mon travail aura permis de mettre en avant des spécificités qui nous avaient auparavant échappé, et notamment une partie 5' très singulière et souche-dépendante. J'ai également pu mettre en évidence l'importance de la RNase III dans sa stabilité, qui laisse supposer l'existence d'un produit de transcription plus long, que cette dernière pourrait cliver pour générer un transcrit stable. Si la fonction exacte de RsaC n'est toujours pas élucidée, les expériences que j'ai menées pointent indéniablement vers un rôle dans la régulation de l'équilibre redox interne et dans la protection contre le stress oxydatif, qui pourrait notamment se faire par la régulation de l'expression de la protéine SodA. D'autres expériences seront nécessaires pour confirmer les travaux que j'ai entrepris au cours de ma thèse, afin de pouvoir replacer RsaC dans les réseaux de régulation de *S. aureus*.**

## Partie 2 : d'un point de vue technique

Comme je l'ai expliqué en début de manuscrit dans mes Objectifs, j'ai voulu développer au laboratoire deux approches pour rechercher les cibles directes des ARNnc de *S. aureus* de manière plus systématique, et mettre en place les outils pour purifier les partenaires protéiques. Les deux approches que j'ai développées au laboratoire, MAPS et Grad-seq, sont complémentaires l'une de l'autre : la première se concentre sur un ARNnc donné alors que la seconde permet de s'intéresser à l'ensemble des ARNnc exprimés *in vivo* dans la cellule.

### 1. L'approche MAPS : un outil puissant, mais perfectionnable

La technique MAPS a été développée par Lalaouna *et al.* (2015) chez *E. coli*. Cette méthode est dérivée de celles de Said *et al.* (2009) et de Desnoyers and Massé (2012) qui repose sur l'expression *in vivo* d'ARN étiqueté en 5' avec deux tiges-boucles reconnues par la protéine du phage MS2. Je l'ai appliquée pour la première fois chez *S. aureus* durant ma thèse en utilisant RsaA comme premier ARN modèle. Sa première cible étant connue (MgrA), le fait de retrouver cet ARNm enrichi en *first hit* dans mes données d'ARN-seq a été une preuve de concept. Après analyse bio-informatique (réalisée par Stefano Marzi) nous avons dégagé une liste de 10 ARNm comme cibles potentielles, pour lesquels j'ai prouvé l'interaction directe et une régulation traductionnelle pour quatre d'entre eux, allant jusqu'à regarder les sites d'interaction pour l'ARNm *flr* (voir Résultats, partie 2).

Ces résultats prouvent que l'approche MS2 est efficace pour aller rechercher des ARNm cibles et constitue un outil puissant pour permettre d'envisager une étude systématique des réseaux de régulation impliquant les ARNnc. Elle pourrait également être utilisée en prenant à la place un ARNm étiqueté pour aller regarder l'ensemble des ARNnc qui sont capables d'interagir avec lui et, possiblement, de réguler son expression et/ou sa stabilité. Il faut noter que cette technique a été réalisée sur une souche sauvage où les RNases sont toujours actives. Il est envisageable de pouvoir réaliser de telles expériences sur des souches mutées au niveau des RNases pour stabiliser les duplexes

formés. Cette approche peut être aussi adaptée à tout type de conditions de culture. La méthode est utilisée à présent par les autres doctorants et chercheurs de l'équipe et des résultats surprenants ont été récemment obtenus sur deux des ARNnc, RsaG et RsaI (résultats non montrés) qui tous deux sont impliqués dans la réponse au métabolisme des sucres.

J'ai aussi montré qu'il est possible d'utiliser des ARN plus longs tels que RsaC. Cependant il semblerait que la taille contribue à retenir l'ARN sur les billes de manière aspécifique. L'utilisation de billes magnétiques pourrait aider pour réduire ce bruit de fond. Néanmoins, là aussi les données de RNA-seq nous ont donné plusieurs cibles potentielles qui interviennent dans des fonctions proches. Nous avons observé très peu d'ARNm communs à RsaA et RsaC apportant une démonstration supplémentaire de la spécificité des interactions visualisées. Cette approche permet aussi d'apporter des informations sur une hiérarchie possible des ARNm cibles. L'approche MAPS se montre complémentaire à la protéomique mais elle nécessite une validation expérimentale pour s'assurer que l'interaction est directe et est fonctionnelle, mais aussi pour déterminer si celle-ci conduit à une répression, activation ou séquestration de la cible.

C'est pour la recherche de protéines interactantes que l'approche s'est révélée moins concluante. Une expérience préliminaire m'avait montré un enrichissement de la protéine ribosomique S1 avec les ARN RsaA, RsaC et RsaE étiquetés, mais je n'ai jamais pu reproduire ces résultats (non montrés). Par ailleurs, ces résultats ont été invalidés par des expériences de retard sur gel qui n'ont montré aucune interaction directe entre RsaA/C et S1. En revanche, cette protéine a été trouvée fortement associée à l'ARNnc RsaI (Marena A., résultats non montrés). Par la suite dans trois expériences distinctes, je n'ai jamais obtenu de protéines qui étaient fortement enrichies par la présence de MS2-RsaA. J'ai utilisé cette approche pour m'assurer s'il existait un équivalent chez *S. aureus* à la protéine Hfq d'*E. coli* qui facilite les appariements entre les ARNnc et leurs ARNm cibles, et/ou stabilise les duplexes entre les deux ARN (Vogel and Luisi, 2011). Le rôle d'Hfq chez *S. aureus* n'étant pas clair – et la protéine n'étant pas impliquée dans la prise en charge des ARNnc (Bonh *et al.*, 2007 ; Huntzinger *et al.*, 2005 ; Boisset *et al.*, 2007) – nous nous sommes demandés si une autre protéine auxiliaire pourrait jouer ce rôle. Au vu des expériences menées, il serait tentant de répondre qu'une telle protéine

n'est peut être pas nécessaire car les duplexes d'ARN chez *S. aureus* seraient suffisamment stables par eux-mêmes (Jousselin et al., 2009). Il faut néanmoins apporter quelques nuances à cette conclusion. Si les duplexes d'ARN semblent être stables chez *S. aureus*, rien ne nous dit qu'il en va de même pour les complexes ARN/protéines et l'approche MS2 n'est pas idéal pour les complexes transitoires et dynamique. Comme décrit plus haut, une étape de pontage au sein des cellules vivantes devrait stabiliser les complexes ribonucléoprotéiques. L'aléa de cette étape c'est que nous risquons de former des complexes ARN/protéines qui ne sont pas biologiquement relevant.

Le MAPS est une approche puissante et relativement fiable chez *S. aureus* avec le principal avantage de purifier les complexes ARN/ARN et ARN/protéines qui se forment *in vivo* dans la cellule. Il ne faut pas oublier que l'étiquette MS2 peut perturber la structure de l'ARN utilisé, entraînant une perte de fonctionnalité (ce qui nécessite donc de réaliser des contrôles en amont). L'un des inconvénients est de ne pouvoir s'atteler qu'à un seul ARN à la fois. C'est dans cette optique que j'ai développé l'approche Grad-seq chez cette bactérie avec l'espoir d'accéder aux interactions ARN-ARN à l'échelle du génome.

## 2. L'approche Grad-seq est-elle réellement utile chez *S. aureus* ?

La méthode de séparation des complexes ribonucléoprotéiques formés *in vivo* chez *S. aureus* par centrifugation sur gradient de glycérol est adaptée des travaux de Wassarman & Storz (2000) qui ont permis de mettre en évidence l'interaction de l'ARN 6S avec l'ARN polymérase. C'est cet ARN qui nous a servi de preuve de concept pour vérifier que cette méthode est bien applicable chez *S. aureus* (**figures 39, 44, 45**). J'ai également démontré que le gradient de glycérol a permis de visualiser une co-migration spécifique entre l'ARNm *mgrA* et RsaA (**figure 40B**). J'ai aussi montré que chaque ARNnc étudié présente un profil de migration qui lui est spécifique. Il est par exemple intéressant de noter que RsaA est présent dans les fractions relativement lourdes du gradient suggérant que celui-ci fait partie intégrante de complexes de taille importante alors que l'ARN transcrit *in vitro* déposé seul sur le gradient n'est présent que dans les



fractions légères (**figure 40A**). Ainsi cette approche devrait aider à classer les ARNnc selon leur profil de migration.

Au départ de mon projet de thèse j'ai voulu utiliser cette approche de gradient de glycérol principalement pour rechercher les protéines interagissant avec les petits ARN de *S. aureus*; en prenant en compte les résultats de Wassarman & Storz (2000) nous espérons visualiser une particule ribonucléoprotéique impliquant les ARNnc (RsaA ou bien encore RsaC). Les gels de protéines que j'ai réalisés, et les données de protéomiques obtenues n'ont pas conduit à un tel enrichissement. Comme pour l'approche MS2, il est possible que les complexes ARN/protéiques se dissocient au cours des différentes étapes de l'expérience. Là encore, une étape de pontage pourrait améliorer. Cependant, cette méthode suggère que nous ne formons pas de RNP stable comme celle observée avec l'ARN 6S. Les données de protéomique réalisées sur deux fractions ont montré un très grand nombre de protéines qui rend l'interprétation difficile, d'autant plus qu'une co-localisation d'une protéine et d'un ARN dans une même fraction du gradient ne signifie pas forcément qu'il y a une interaction. En fait, nous pourrions combiner les approches Grad-Seq et MAPS pour effectuer une deuxième séparation sur chromatographie d'affinité. Une autre critique que je voudrais émettre, est que le protocole que j'ai appliqué, repris de Wassarman & Storz (2000), adopte des paramètres qui facilitent la purification de l'ARN 6S. Ces conditions ne permettent pas par exemple de vérifier si les ARN se retrouvent associés aux ribosomes en cours de traduction ou pas. Par exemple, RsaC et l'ARNIII se retrouvent dans les fractions lourdes où nous avons détecté les sous-unités 30S. Même si nous n'avons pas été en mesure de montrer que RsaC forme des complexes d'initiation par toeprint *in vitro*, il sera important de visualiser les ribosomes en entier et les polysomes afin de s'assurer que RsaC ne code pas pour des petites protéines.

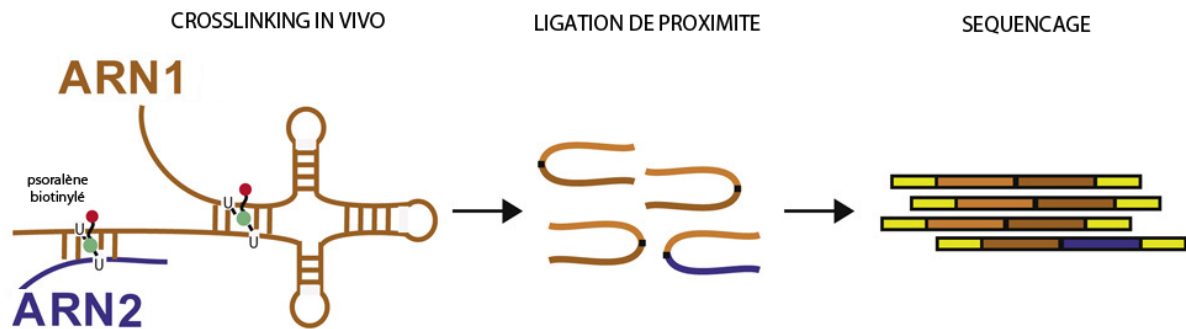
Du côté des interactions ARN-ARN, le bilan est un peu plus positif, car les duplexes d'ARN chez *S. aureus* sont stables et ne sont pas rapidement dégradés. Nous avons obtenu récemment un grand nombre de données provenant de l'ARN-seq qui nécessitent une analyse plus en profondeur, mais avec lesquelles j'ai déjà pu vérifier la séparation de certains ARNnc dans le gradient. J'ai également mis évidence un petit ARN

très fortement exprimé, Sau5971, par rapport aux autres, qui adopte un profil de localisation très spécifique (**figure 45**). Cet ARN pourrait être issu de la région 3'UTR d'un ARNm codant pour une métalloprotéase membranaire de type Api. Cet ARN a été trouvé stabilisé dans une souche qui est caractérisée par un phénotype stable de petite colonie (Abu-Qatouseh et al., 2010). La fonction de cet ARN demeure cependant inconnue. Par rapport à RsaA il sera intéressant de regarder si je retrouve les ARNm que j'ai obtenus par MAPS dans les fractions dans lesquelles RsaA est enrichi. Le principal problème que nous avons pour le moment concerne l'annotation du génome qui est très incomplète. Nous avons de notre côté séquencé la souche HG001 que nous avons récemment annoté avec l'aide de Christine Gaspin (INRA, Toulouse). Cette souche a un génome quasi-identique à celui de HG003, la souche sur laquelle le Grad-seq a été réalisé. L'analyse des données issues de l'ARN-seq sera réalisée par S. Marzi en collaboration avec J. Vogel (Université de Wurzburg), en espérant pouvoir retrouver RsaC avec l'ARNm *sodA*.

Le but du gradient de glycérol était avant tout de pouvoir visualiser l'intégralité des ARNnc et des ARNm qui co-localisent dans les différentes fractions du gradient, avec dans l'idée que la co-localisation et l'enrichissement des deux ARN dans une même fraction seraient les signes d'une interaction. Cependant cette approche est techniquement lourde, très sensible aux variations expérimentales et nécessite un grand nombre de contrôles en amont avant de pouvoir procéder au séquençage. Une nouvelle approche s'affranchit de la mise en place des gradients et semble donner une meilleure couverture des interactions ARN-ARN *in vivo* : il s'agit de la méthode *sequencing of psoralen crosslinked, ligated and selected hybrids* pour SPLASH (Aw et al., 2016) ou bien encore "LIGATION of interacting RNA followed by high-throughput sequencing" pour LIGR-seq (Sharma et al., 2016).

Ces approches ont été développées sur cellules eucaryotes et consiste à ponter les ARN qui interagissent dans la cellule *in vivo* par l'utilisation d'un composé appelé psoralène, qui est de plus biotinylé. Les ARN totaux sont ensuite extraits et fragmentés, et les fragments d'ARN qui sont pontés par le psoralène sont retenus sur des billes de streptavidine, puis sont reliés ensemble. Le psoralène est ensuite éliminé et les

fragments d'ARN chimères sont rétro-transcrits, amplifiés par PCR et séquencés (**figure 49**).



**Figure 49 : Représentation schématique du SPLASH** - Le psoralène biotinylé peut se fixer entre deux molécules d'ARN différentes ou dans les régions d'appariement d'un même ARN. Après fragmentation, les ARN crosslinkés sont ligés entre eux et les chimères sont rétro-transcrits et amplifiées par PCR avant d'être séquencées.

Les données de séquençage ont ainsi permis de révéler à la fois les interactions intramoléculaires (régions en double brin d'un ARN donné), et les interactions intermoléculaires. Associée à des annotations correctes de génome, cette approche pourrait permettre plus rapidement d'obtenir une vision globale des ARN qui interagissent ensemble *in vivo* chez *S. aureus*. Nous pourrions également combiner cette approche au Grad-seq en réalisant cette étape de pontage en amont de la préparation des extraits bruts. Bien que rajouter cette étape de pontage rallonge sensiblement l'expérience, elle nous permettrait de discriminer plus rapidement les ARN qui nous intéressent, et en fonction de leur localisation dans le gradient nous pourrions simplement cibler directement les ARN qui se trouvent soit dans les fractions légères, soit dans les fractions lourdes du gradient.

Quoi qu'il en soit, ces nouvelles approches basées sur le même principe SPLASH/LIGR-seq me semblent plus prometteuses que celle du gradient de glycérol. Elles offrent la possibilité à la fois d'obtenir des informations sur l'existence d'interactions à longue distance des ARN et donc sur leur repliement, et sur les interactions entre ARN (i.e., opérons divergents et convergents, entre ARNm, entre ARN-ARNnc, ARNnc-ARNnc...). Couplée à une étape ultérieure de co-immunoprécipitation utilisant des anticorps contre une protéine donnée (telle que la RNase III) portant une étiquette FLAG, il serait alors possible de classifier les hybrides observés. Une telle approche pourrait être utilisée dans diverses conditions de stress (stress oxydatif, NO) ou dans diverses souches mutées pour analyser la dynamique de ces interactions. Il serait même envisageable

d'effectuer une telle étude sur les bactéries internalisées dans les macrophages pour mieux aborder le rôle des ARN dans les interactions hôte-pathogène. De même l'utilisation de l'approche MAPS dans ces conditions diverses permettra d'aller regarder les variations des cibles d'un ARNnc et nous permettra d'aborder de façon précise leur place au sein des réseaux de régulation de *S. aureus*. Le développement de ces techniques au sein de mon équipe constitue un tournant important pour caractériser ces réseaux, leurs composantes et leurs dynamiques.

# ANNEXES

# Matériel et méthodes

**Note :** la majeure partie du détail des expériences conduites est déjà explicitée dans les « Matériel et méthodes » des deux articles présentés en Résultats.

## 1. Liste des souches et plasmides utilisés

	Caractéristiques	Référence
<b>Souches <i>E. coli</i></b>		
DC10B	Mutant dmc; utilisée pour le clonage de plasmides qui peuvent être directement transformés dans <i>S. aureus</i>	Monk <i>et al.</i> (2012)
<b>Souches <i>S. aureus</i></b>		
RN HG001	Dérivée de 8325-4 restaurée pour le gène <i>rsbU</i>	Herbert <i>et al.</i> (2010)
RN HG003	Dérivée de 8325-4 restaurée pour les gènes <i>rsbU</i> et <i>tcaR</i>	Herbert <i>et al.</i> (2010)
RN4220	Mutant de restriction de 8225-4, utilisée pour clonage	Kreiswirth <i>et al.</i> (1983)
RN6390	Dérivée de 8325-4, déficiente pour le gène <i>rsbU</i>	Peng <i>et al.</i> (1988)
LUG779	RN6390- $\Delta$ rnc, $\Delta$ rnc region :: <i>cat86</i>	Huntzinger <i>et al.</i> (2005)
Newman	Isolat clinique, souche CP5, production importante de <i>clumping factor</i>	Duthie <i>et al.</i> (1952)
COL	Souche de laboratoire résistante à la méticilline	Dyke <i>et al.</i> (1966)
LUG 2498	RN HG001 $\Delta$ rsaC	
LUG 2482	RN HG001 transformée avec pCN51/P3	
LUG 2612	LUG 2498 transformée avec pCN51/P3	
LUG 2625	LUG 2498 transformée avec pLUG 1268	
RN1 pCN51/MS2	RN HG001 transformée avec pCN51/MS2	
RN1 $\Delta$ rsaC		
pCN51/MS2-RsaC	LUG 2498 transformée avec pCN51/MS2-RsaC	
RN1 $\Delta$ rsaC		
pCN51/MS2-RsaC <sub>L</sub>	LUG 2498 transformée avec pCN51/MS2-RsaC <sub>L</sub>	
<b>Plasmides</b>		
pCN51	Vecteur navette entre <i>E. coli</i> et <i>S. aureus</i>	Charpentier <i>et al.</i> (2004)
pCN51/P3	pCN51 avec promoteur P3	
PLUG 1268	pCN51 exprimant RsaC <sub>L</sub> (1117 nts) sous contrôle du promoteur P3	
pCN51/MS2	pCN51 exprimant étiquette MS2 sous contrôle du promoteur P3	
pCN51/MS2-RsaC	pCN51 exprimant MS2-RsaC (544 nts) sous contrôle du promoteur P3	
pCN51/MS2-RsaC <sub>L</sub>	pCN51 exprimant MS2-RsaC <sub>L</sub> (1117 nts) sous contrôle du promoteur P3	
pUC18-rsaA	<i>rsaA</i> (forme courte) sous contrôle du promoteur T7	Romilly <i>et al.</i> (2014)

## 2. Liste des oligonucléotides utilisés

Nom	Séquence	Utilisation
T7-RsaC Fwd	TAATACGACTCACTATAGGGAAAAGCTTTATGTGG	Transcription in vitro RsaC
RsaC Rev	AAAATAGCCACACTCATATG	Transcription in vitro RsaC
T7 RsaC DIG For	TAATACGACTCACTATAGGGACGCCATTCCTACACACTC	Sonde DIG RsaC
RsaC DIG Rev	CAAAGGAAAGGGGCATACAA	Sonde DIG RsaC
T7 RsaCas DIG For	TAATACGACTCACTATAGGGCAAAGGAAAGGGGCATACAA	Sonde DIG RsaC antisens
RsaCas DIG Rev	ACGCCATTCCTACACACTC	Sonde DIG RsaC antisens
PstI-MS2 For	CGCCTG CAGCGTACACCATCAGGGTACGTTTTTCGTACACCATCAGGGTACGTTT	clonage MS2
BamHI-MS2 Rev	CGCGGATCCAAACGTACCCTGATGGTGTACGAAAAACGTACCCTGATGGTGTACG	clonage MS2
RsaC PrExt	ATTGTATTAGAAGAACTTCT	Primer Extension RsaC/SHAPE
RsaC PrExt 2	GTGGTAAAAGAGATGAAGG	Primer Extension RsaC/SHAPE
RsaC PrExt 3	GCGATCGGGAACGGTATA	Primer Extension RsaC/SHAPE
RsaC PrExt 4	CATTA AACACAGGGTGTTA	Primer Extension RsaC/SHAPE
RsaC-514	GTGAGTATACTGTAGCCTAC	SHAPE
RsaC-416	GCTAATAAATATCCAATAA	SHAPE
RsaC-288	ATTTTCAGATCATCCCTCTT	SHAPE
RsaC-PCR1for	AAAGGATCCTGGTGTGCCTATCATT	Construction plasmide d'inactivation
RsaC-PCR1Rev	AAACTCGAGTGGGACTATTAGCCTTCGACTT	Construction plasmide d'inactivation
RsaC-PCR2rev	ACGCGTCGACTTATTATTTATTGCATCTG	Construction plasmide d'inactivation
RsaC-PCR2for	AATCTCGAGTGAATACACTTAACTTCACC	Construction plasmide d'inactivation
RsaC 679291	TGATGCGTCGAGGATATGAA	Contrôle inactivation
RsaC 683506	TGGGAGATGCAATGAGTCAC	Contrôle inactivation
BamHI rsaC RN1 Reverse	GGGGATCCggtgaagtaagtgtattcac	Construction plasmide complémentation
PstI rsaC RN1 Forward	CCCTGCAGaagtgaaggtaaatagttccc	Construction plasmide complémentation
gyrB F	GGTGGCGACTTTGATCTAGC	qPCR
gyrB R	TTATACAACGGTGGCTGTGC	qPCR
RsaC Forward	CGTCCCTTCATCTCTTTTAACCA	qPCR
RsaC Reverse	CGTCCCTTCATCTCTTTTAACCA	qPCR
sodMrev	ATC TAA AGT GCC CCA CTG CG	qPCR
sodMfw	CGG AAG CGA TGA GGA TGT CA	qPCR
sodA rev2	CCA ACC TGA ACC AAA GCG TG	qPCR
sodA fow2	ACG TAA TAA TGG CGG TGG ACA	qPCR
PSA2 fow	TTG GTG GAG ACA ACG TCG AT	qPCR
PSA2 rev	ACC AAC CGT TAC CAG TCT CT	qPCR
nhk1 fow	AAACGAGAGTGAACCTCCGCA	qPCR
nhk1 rev	CCGTCACAAACCATTCCAAC	qPCR

### 3. Méthodes expérimentales

#### *Test Rifampicine*

Les souches d'intérêt (HG001-Wt, RN6390-Wt, RN6390- $\Delta rnc$ ) sont mises en culture dans 50 ml de milieu BHI à 37°C, en partant de pré-cultures des même souches réalisées la veille (une colonie dans 2 ml BHI). Après 5h de culture et vérifications de la mesure de la  $DO_{600nm}$ , la rifampicine est ajoutée dans chaque culture à une concentration finale de 300  $\mu\text{g/ml}$  au  $T_0$ . Aux temps  $t = 2, 4, 8, 15, 30, 45$  et 60 minutes, 5 ml de culture sont prélevées, les cellules sont centrifugées, reprises dans 1 ml de solution RNA Pro (MP Biomedicals) et les ARN totaux sont extraits en suivant le protocole du kit FastRNA Pro Blue Kit du fabricant. Pour chaque échantillon, 10  $\mu\text{g}$  d'ARN totaux sont déposés sur gel d'agarose dénaturant et le Northern blot est réalisé comme indiqué dans Romilly *et al.*, 2014.

#### *Primer Extension*

Les réactions de *primer extension* ont été réalisées en suivant scrupuleusement le protocole indiqué dans Helfer *et al.*, 2013. Pour les réactions d'extension comme celles de séquençage, 20  $\mu\text{g}$  d'ARN totaux de HG001 ont été utilisés comme matrice de départ. La migration des produits de réaction s'est faite sur un gel de polyacrylamide 6%-urée 8 M.

#### *SHAPE*

Les réactions de SHAPE conduites par Dr. Zong-Fu Wu ont été réalisées selon le protocole indiqué par Helfer *et al.*, 2013. Les données acquises ont ensuite été analysées à l'aide du programme QuShape (Karabiber *et al.*, 2015).

#### *Quantification relative des ARNnc par RT-qPCR*

Les ARN sont quantifiés par le spectrophotomètre NanoDrop et 1  $\mu\text{g}$  d'ARN total est rétro-transcrit en ADNc en utilisant le kit Reverse Transcriptase System (Promega). L'ADNc est utilisé comme matrice pour amplification par RT-PCR en utilisant l'appareil



LightCycler (Roche) et les oligonucléotides spécifiques sont indiqués dans la liste précédente. Les quantités relatives d'amplicons pour chaque gène sont déterminées en utilisant le même procédé de RT-PCR au contrôle interne (gène *gyrB* qui code la sous-unité B de l'ADN gyrase). Les expériences de RT-qPCR ont été réalisées en triplicats.

#### **Test de stress oxydatif au MV**

La sensibilité au MV a été mesurée par le test de viabilité tel que décrit par Karavolos *et al.*, 2003. Des milieux BHI frais ont étéensemencés avec une culture de *S. aureus* incubée sur la nuit pour une  $DO_{600nm} \sim 0,05$ . Les cellules sont soumises au MV (10 nM) en début de phase exponentielle de croissance ( $t = 2h$ ). Des aliquotes de cultures sont prélevées aux temps choisis et le nombre de bactéries viables est déterminé par ensemencement sur gélose solide GP EryT. Les expériences ont été réalisées en triplicats.

#### **Mesure d'activité de SodA (activité SOD)**

Pour le calcul de l'activité SOD, les cellules ont été collectées en fin de phase stationnaire de croissance ( $DO_{600nm} \sim 9$ , ou  $\sim 4$  en présence de MV), lavées avec du PBS et lysées par le stress osmotique et utilisation de la lysostaphine tel que décrit dans Clements *et al.*, 1999. La concentration en protéines totales a été déterminée par le test de Bradford (Bio-Rad) et les échantillons sont stockés à  $-80^{\circ}C$  jusqu'à utilisation. Les activités de SodA et SodM ont été déterminées par coloration négative sur gel de polyacrylamide en conditions natives (Valderas *et al.*, 2012).

Des quantités égales de protéines (20  $\mu g$ ) sont déposées sur un gel de polyacrylamide 12% dans un tampon glycine sans SDS. Après la migration, le gel est imbibé dans 1,225 nM de NBT pendant 45 min et ensuite lavé avec de l'eau distillée. Le gel est plongé dans une solution contenant 0,028 nM de riboflavine et 28 mM de TEMED (tetraméthylène diamine) pendant 45 min. Le gel est ensuite exposé à la lumière pour initier la réaction photo-chimique. L'activité SOD est visualisée par une zone claire entourée d'un fond bleu foncé.

### Test de production endogène de ROS

Pour la quantification des ROS, les cellules sont prélevées après 6h de croissance (1 ml à  $DO_{600nm} = 1$ ), lavées deux fois avec le tampon PBS et reprises dans ce même tampon. Une sonde 2'7'-dichlorofluorescéine (Sigma) est ajoutée aux cellules en suspension, pour une concentration finale de 10  $\mu$ M. L'intensité de fluorescence (FI) est mesurée pendant 7h avec un spectrofluorimètre dans des plaques à 96 puits (Ex/Em = 485/538 nm). La valeur de FI est normalisée par l'absorbance à 600 nm de chaque culture pour la quantification. Les expériences ont été conduites en triplicats.

### Gradient de glycérol

#### 1. Préparation des extraits bruts

Les cultures de *S. aureus* d'intérêt (HG001Wt, HG003-Wt) sont cultivées en milieu BHI à 37°C. Le temps de culture dépend de la phase de croissance dans laquelle l'ARN est exprimé (pour RsaC, 6h, la  $DO_{600nm} \sim 5/6$ ). Les cellules sont ensuite récoltées, centrifugées et resuspendues dans 1 ml de Tampon R (20 mM Tris-HCl pH 7,5 ; 10 mM Acétate de  $Mg^{2+}$  ; 150 mM NaCl ; 2 mM DTT ; 1 mM EDTA) à 4°C. Les cellules sont placées dans un Tube du kit FastRNA Pro Blue (MP Biomedicals) (un tube contient 0,7g de billes de verre) et sont lysées par l'appareil FastPrep avec les réglages force 6.0 pendant 40 s. Les tubes sont ensuite centrifugés 15 min, 13000 rpm, 4°C et les surnageants, constituant les extraits bruts, sont doucement récupérés et placés à 4°C avant d'être déposés sur gradient.

#### 2. Préparation des gradients

Deux solutions de tampon R avec 10% et 40% de glycérol sont réalisées. Le gradient est préparé dans un tube à centrifugeuse Beckman-Coulter Polyallomer (14\*89 mm). Dans le tube, 5 ml de la solution à 10% est déposé, puis 5 ml de la solution à 40% (en la déposant directement dans le fond du tube, elle ira soulever la solution légère), en veillant à ne pas mélanger les deux solutions. Après fermeture avec le bouchon adéquat

(25 mm hauteur, BioComp), les tubes sont placés symétriquement sur le porteur MagnaBase lui-même déposé sur la plaque magnétique de l'appareil Gradient Master (BioComp). Le gradient est réalisé par l'appareil en sélectionnant le programme adéquat. Une fois le gradient réalisé, les tubes sont placés dans les godets Beckman SW41Ti refroidis et laissés à 4°C jusqu'à utilisation.

### 3. Centrifugation et collection des gradients

Pour chaque gradient, un extrait brut est déposé doucement (par capillarité sur la surface interne du tube). Les godets sont refermés et placés dans l'ultracentrifugeuse avec le rotor SW41Ti. La centrifugation est réalisée pendant 16h, 36000 rpm, 4°C. La collecte du gradient est réalisée en combinant un appareil AKTA avec un collecteur de fractions de Gradient de Densité Auto-Densi Flow (Labconco). Le collecteur de fractions est préalablement lavé avec une solution d'éthanol 20% puis équilibré avec une solution de tampon R à 10% de glycérol. La DO<sub>280nm</sub> est mesurée tout au long de la collecte. Les réglages de l'AKTA utilisés sont les suivants : fraction 0,6 ml ; flow : 1 ml/min ; position : 8 (qui permet de récolter le gradient dans des tubes eppendorfs 1,5 ml).

Sur chaque fraction, les ARN sont extraits par traitement consécutif au phénol-chloroforme puis chloroforme/alcool isoamylique puis précipités à l'éthanol froid. Les tubes sont centrifugés, les culots d'ARN lavés à l'éthanol 80% et repris dans 20 µl d'H<sub>2</sub>O traitée au DEPC. Les concentrations sont mesurées par Nanodrop et les ARN sont traités 15 min à la DNase I (dans le tampon adéquat, ThermoScientific) à 37°C. Le volume réactionnel est ajusté à 100 µl avec de l'eau et une nouvelle extraction puis précipitation des ARN est réalisée. Après lavage à l'éthanol 80%, les culots sont séchés et repris dans 15 µl d'eau traitée au DEPC. Les ARN peuvent être ensuite envoyés à séquencer.

# Bibliographie

Aw, J.G.A., Shen, Y., Wilm, A., Sun, M., Lim, X.N., Boon, K.-L., Tapsin, S., Chan, Y.-S., Tan, C.-P., Sim, A.Y.L., et al. (2016). In Vivo Mapping of Eukaryotic RNA Interactomes Reveals Principles of Higher-Order Organization and Regulation. *Mol. Cell* 62, 603–617.

Bohn, C., Rigoulay, C., and Bouloc, P. (2007). No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*. *BMC Microbiol.* 7, 10.

Boisset, S., Geissmann, T., Huntzinger, E., Fechter, P., Bendridi, N., Possedko, M., Chevalier, C., Helfer, A.C., Benito, Y., Jacquier, A., et al. (2007). *Staphylococcus aureus* RNAlII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev* 21, 1353–1366.

Busch, A., Richter, A.S., and Backofen, R. (2008). IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* 24, 2849–2856.

Chambers, H.F., and DeLeo, F.R. (2009). Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nature Reviews Microbiology* 7, 629–641.

Chauhan, A., Bernardin, A., Mussard, W., Kriegel, I., Esteve, M., Ghigo, J.-M., Beloin, C., and Semetey, V. (2014). Preventing Biofilm Formation and Associated Occlusion by Biomimetic Glycocalyxlike Polymer in Central Venous Catheters. *Journal of Infectious Diseases* 210, 1347–1356.

Cheung, A.L., Nishina, K.A., Trotonda, M.P., and Tamber, S. (2008). The SarA protein family of *Staphylococcus aureus*. *The International Journal of Biochemistry & Cell Biology* 40, 355–361.

Clements, M.O., and Foster, S.J. (1999). Stress resistance in *Staphylococcus aureus*. *Trends Microbiol.* 7, 458–462.

Clements, M.O., Watson, S.P., and Foster, S.J. (1999). Characterization of the major superoxide dismutase of *Staphylococcus aureus* and its role in starvation survival, stress resistance, and pathogenicity. *J. Bacteriol.* *181*, 3898–3903.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szczesniak, M.W., Gaffney, D.J., Elo, L.L., Zhang, X., et al. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology* *17*.

Cowan, S.T., Shaw, C., and Williams, R.E. (1954). Type strain for *Staphylococcus aureus* Rosenbach. *J. Gen. Microbiol.* *10*, 174–176.

Cramton, S.E., Schnell, N.F., Götz, F., and Brückner, R. (2000). Identification of a new repetitive element in *Staphylococcus aureus*. *Infect. Immun.* *68*, 2344–2348.

Crosby, H.A., Schlievert, P.M., Merriman, J.A., King, J.M., Salgado-Pabón, W., and Horswill, A.R. (2016). The *Staphylococcus aureus* Global Regulator MgrA Modulates Clumping and Virulence by Controlling Surface Protein Expression. *PLOS Pathogens* *12*, e1005604.

Delihias, N. (2011). Impact of small repeat sequences on bacterial genome evolution. *Genome Biol Evol* *3*, 959–973.

Deng, X., Chen, K., Luo, G.-Z., Weng, X., Ji, Q., Zhou, T., and He, C. (2015). Widespread occurrence of N6-methyladenosine in bacterial mRNA. *Nucleic Acids Res.* *43*, 6557–6567.

Desnoyers, G., and Massé, E. (2012). Noncanonical repression of translation initiation through small RNA recruitment of the RNA chaperone Hfq. *Genes Dev.* *26*, 726–739.

Deurenberg, R.H., and Stobberingh, E.E. (2008). The evolution of *Staphylococcus aureus*. *Infection, Genetics and Evolution* *8*, 747–763.

Dowding, J.E. (1977). Mechanisms of gentamicin resistance in *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy* *11*, 47–50.

Durand, S., Tomasini, A., Braun, F., Condon, C., and Romby, P. (2015). sRNA and mRNA turnover in Gram-positive bacteria. *FEMS Microbiology Reviews* 39, 316–330.

Fagerlund, A., Granum, P.E., and Håvarstein, L.S. (2014). *Staphylococcus aureus* competence genes: mapping of the SigH, ComK1 and ComK2 regulons by transcriptome sequencing. *Mol. Microbiol.* 94, 557–579.

Fechter, P., Caldelari, I., Lioliou, E., and Romby, P. (2014). Novel aspects of RNA regulation in *Staphylococcus aureus*. *FEBS Letters* 588, 2523–2529.

Han, J., He, L., Shi, W., Xu, X., Wang, S., Zhang, S., and Zhang, Y. (2014). Glycerol Uptake Is Important for L-Form Formation and Persistence in *Staphylococcus aureus*. *PLoS ONE* 9, e108325.

Huntzinger, E., Boisset, S., Saveanu, C., Benito, Y., Geissmann, T., Namane, A., Lina, G., Etienne, J., Ehresmann, B., Ehresmann, C., et al. (2005). *Staphylococcus aureus* RNAlII and the endoribonuclease III coordinately regulate *spa* gene expression. *EMBO J* 24, 824–835.

Ibarra, J.A., Pérez-Rueda, E., Carroll, R.K., and Shaw, L.N. (2013). Global analysis of transcriptional regulators in *Staphylococcus aureus*. *BMC Genomics* 14, 1.

Ji, Q., Zhang, L., Sun, F., Deng, X., Liang, H., Bae, T., and He, C. (2012). *Staphylococcus aureus* CymR Is a New Thiol-based Oxidation-sensing Regulator of Stress Resistance and Oxidative Response. *Journal of Biological Chemistry* 287, 21102–21109.

Jousselin, A., Metzinger, L., and Felden, B. (2009). On the facultative requirement of the bacterial RNA chaperone, Hfq. *Trends Microbiol.* 17, 399–405.

Karabiber, F., McGinnis, J.L., Favorov, O.V., and Weeks, K.M. (2013). QuShape: Rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. *RNA* 19, 63–73.

Karavolos, M.H. (2003). Role and regulation of the superoxide dismutases of *Staphylococcus aureus*. *Microbiology* 149, 2749–2758.

Khoon, L.Y., and Neela, V. (2010). Secretome of *Staphylococcus aureus*. *African Journal of Microbiology Research* 4, 500–508.

KOCH, H.U., HAAS, R., and FISCHER, W. (1984). The role of lipoteichoic acid biosynthesis in membrane lipid metabolism of growing *Staphylococcus aureus*. *European Journal of Biochemistry* 138, 357–363.

Lasa, I., Toledo-Arana, A., and Gingeras, T.R. (2012). An effort to make sense of antisense transcription in bacteria. *RNA Biol* 9, 1039–1044.

Lindsay, J.A., and Holden, M.T.G. (2004). *Staphylococcus aureus*: superbug, super genome? *Trends in Microbiology* 12, 378–385.

Lioliou, E., Sharma, C.M., Caldelari, I., Helfer, A.-C., Fechter, P., Vandenesch, F., Vogel, J., and Romby, P. (2012). Global Regulatory Functions of the *Staphylococcus aureus* Endoribonuclease III in Gene Expression. *PLoS Genetics* 8, e1002782.

Lowy, F.D. (1998). *Staphylococcus aureus* infections. *N. Engl. J. Med.* 339, 520–532.

Malachowa, N., and DeLeo, F.R. (2010). Mobile genetic elements of *Staphylococcus aureus*. *Cellular and Molecular Life Sciences* 67, 3057–3071.

Marbaniang, C.N., and Vogel, J. (2016a). Emerging roles of RNA modifications in bacteria. *Curr. Opin. Microbiol.* 30, 50–57.

Marbaniang, C.N., and Vogel, J. (2016b). Emerging roles of RNA modifications in bacteria. *Current Opinion in Microbiology* 30, 50–57.

Marshall, J.H., and Wilmoth, G.J. (1981). Pigments of *Staphylococcus aureus*, a series of triterpenoid carotenoids. *Journal of Bacteriology* 147, 900–913.

Novick, R.P., Christie, G.E., and Penadés, J.R. (2010). The phage-related chromosomal islands of Gram-positive bacteria. *Nature Reviews Microbiology* 8, 541–551.

Orenstein, A. (2011). The Discovery and Naming of *Staphylococcus aureus*. Periodical [serial Online]. Date. Available From.

Otto, M. (2014). *Staphylococcus aureus* toxins. *Current Opinion in Microbiology* 17, 32–37.

Perichon, B., and Courvalin, P. (2009). VanA-Type Vancomycin-Resistant *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy* 53, 4580–4587.

Peschel, A., and Otto, M. (2013). Phenol-soluble modulins and staphylococcal infection. *Nat. Rev. Microbiol.* 11, 667–673.

Poupel, O., Moyat, M., Groizeleau, J., Antunes, L.C.S., Gribaldo, S., Msadek, T., and Dubrac, S. (2016). Transcriptional Analysis and Subcellular Protein Localization Reveal Specific Features of the Essential WalkR System in *Staphylococcus aureus*. *PLoS ONE* 11, e0151449.

Prat, C., Bestebroer, J., de Haas, C.J.C., van Strijp, J.A.G., and van Kessel, K.P.M. (2006). A new staphylococcal anti-inflammatory protein that antagonizes the formyl peptide receptor-like 1. *J. Immunol.* 177, 8017–8026.

Purves, J., Blades, M., Arafat, Y., Malik, S.A., Bayliss, C.D., and Morrissey, J.A. (2012a). Variation in the genomic locations and sequence conservation of STAR elements among staphylococcal species provides insight into DNA repeat evolution. *BMC Genomics* 13, 515.

Purves, J., Blades, M., Arafat, Y., Malik, S.A., Bayliss, C.D., and Morrissey, J.A. (2012b). Variation in the genomic locations and sequence conservation of STAR elements among staphylococcal species provides insight into DNA repeat evolution. *BMC Genomics* 13, 515.



Abu-Qatouseh, L.F., Chinni, S.V., Seggewiss, J., Proctor, R.A., Brosius, J., Rozhdestvensky, T.S., Peters, G., von Eiff, C., and Becker, K. (2010). Identification of differentially expressed small non-protein-coding RNAs in *Staphylococcus aureus* displaying both the normal and the small-colony variant phenotype. *J. Mol. Med.* *88*, 565–575.

Queck, S.Y., Jameson-Lee, M., Villaruz, A.E., Bach, T.-H.L., Khan, B.A., Sturdevant, D.E., Ricklefs, S.M., Li, M., and Otto, M. (2008). RNAIII-Independent Target Gene Control by the *agr* Quorum-Sensing System: Insight into the Evolution of Virulence Regulation in *Staphylococcus aureus*. *Molecular Cell* *32*, 150–158.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. *Nat. Biotechnol.* *29*, 24–26.

Rochat, T., Bouloc, P., Yang, Q., Bossi, L., and Figueroa-Bossi, N. (2012a). Lack of interchangeability of Hfq-like proteins. *Biochimie* *94*, 1554–1559.

Rochat, T., Bouloc, P., Yang, Q., Bossi, L., and Figueroa-Bossi, N. (2012b). Lack of interchangeability of Hfq-like proteins. *Biochimie* *94*, 1554–1559.

Romeo, T., Vakulskas, C.A., and Babitzke, P. (2013). Post-transcriptional regulation on a global scale: form and function of Csr/Rsm systems. *Environ. Microbiol.* *15*, 313–324.

Ruiz de los Mozos, I., Vergara-Irigaray, M., Segura, V., Villanueva, M., Bitarte, N., Saramago, M., Domingues, S., Arraiano, C.M., Fechter, P., Romby, P., et al. (2013). Base pairing interaction between 5'- and 3'-UTRs controls *icaR* mRNA translation in *Staphylococcus aureus*. *PLoS Genet.* *9*, e1004001.

Sesto, N., Wurtzel, O., Archambaud, C., Sorek, R., and Cossart, P. (2013). The excludon: a new concept in bacterial antisense RNA-mediated gene regulation. *Nat. Rev. Microbiol.* *11*, 75–82.

Sharma, E., Sterne-Weiler, T., O'Hanlon, D., and Blencowe, B.J. (2016). Global Mapping of Human RNA-RNA Interactions. *Mol. Cell* *62*, 618–626.

Showsh, S.A., De Boever, E.H., and Clewell, D.B. (2001). Vancomycin Resistance Plasmid in *Enterococcus faecalis* That Encodes Sensitivity to a Sex Pheromone Also Produced by *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy* 45, 2177–2178.

Soutourina, O., Dubrac, S., Poupel, O., Msadek, T., and Martin-Verstraete, I. (2010). The pleiotropic CymR regulator of *Staphylococcus aureus* plays an important role in virulence and stress response. *PLoS Pathog.* 6, e1000894.

Strasters, K.C., and Winkler, K.C. (1963). Carbohydrate metabolism of *Staphylococcus aureus*. *Microbiology* 33, 213–229.

Szafranska, A.K., Oxley, A.P.A., Chaves-Moreno, D., Horst, S.A., Roßlenbroich, S., Peters, G., Goldmann, O., Rohde, M., Sinha, B., Pieper, D.H., et al. (2014). High-Resolution Transcriptomic Analysis of the Adaptive Response of *Staphylococcus aureus* during Acute and Chronic Phases of Osteomyelitis. *mBio* 5, e01775–14.

Thammavongsa, V., Kim, H.K., Missiakas, D., and Schneewind, O. (2015). Staphylococcal manipulation of host immune responses. *Nature Reviews Microbiology* 13, 529–543.

Theodore, T.S., and Schade, A.L. (1965). Carbohydrate metabolism of iron-rich and iron-poor *Staphylococcus aureus*. *Microbiology* 40, 385–395.

Tomasini, A., François, P., Howden, B.P., Fechter, P., Romby, P., and Caldelari, I. (2014). The importance of regulatory RNAs in *Staphylococcus aureus*. *Infection, Genetics and Evolution* 21, 616–626.

Tormo-Más, M.A., Mir, I., Shrestha, A., Tallent, S.M., Campoy, S., Lasa, I., Barbé, J., Novick, R.P., Christie, G.E., and Penadés, J.R. (2010). Moonlighting bacteriophage proteins derepress staphylococcal pathogenicity islands. *Nature* 465, 779–782.

Tsubakishita, S., Kuwahara-Arai, K., Sasaki, T., and Hiramatsu, K. (2010). Origin and Molecular Evolution of the Determinant of Methicillin Resistance in Staphylococci. *Antimicrobial Agents and Chemotherapy* 54, 4352–4359.

Vogel, J., and Luisi, B.F. (2011). Hfq and its constellation of RNA. *Nat. Rev. Microbiol.* 9, 578–589.

Wang, B., and Muir, T.W. (2016). Regulation of Virulence in *Staphylococcus aureus*: Molecular Mechanisms and Remaining Puzzles. *Cell Chemical Biology* 23, 214–224.

# Publications

## **sRNA and mRNA turnover in Gram-positive bacteria**

**Sylvain Durand<sup>1</sup>, Arnaud Tomasini<sup>2</sup>, Frédérique Braun<sup>1</sup>, Ciarán Condon<sup>1\*</sup>, and Pascale Romby<sup>2\*</sup>**

\*Co-corresponding authors

<sup>1</sup>CNRS FRE 3630 (affiliated with Univ. Paris Diderot, Sorbonne Paris Cité),  
Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France

<sup>2</sup>Architecture et Réactivité de l'ARN, Université de Strasbourg, CNRS, IBMC, 15 rue René  
Descartes, F-67084 Strasbourg, France

### **Abstract**

It is widely recognized that RNA degradation plays a critical role in gene regulation when fast adaptation of cell growth is required to respond to stress and changing environmental conditions. Bacterial ribonucleases acting alone or in concert with various *trans*-acting regulatory factors are important mediators of RNA degradation. Here, we will give an overview of what is known about ribonucleases in several Gram-positive bacteria, their specificities and mechanisms of action. In addition, we will illustrate how sRNAs act in a coordinated manner with ribonucleases to regulate the turnover of particular mRNA targets, and the complex interplay existing between the ribosome, the ribonucleases and RNAs.

Keywords: Gram-positive bacteria, ribonucleases, sRNA, mRNA stability



## REVIEW ARTICLE

**sRNA and mRNA turnover in Gram-positive bacteria**Sylvain Durand<sup>1</sup>, Arnaud Tomasini<sup>2</sup>, Frédérique Braun<sup>1</sup>, Ciarán Condon<sup>1,\*</sup>  
and Pascale Romby<sup>2</sup><sup>1</sup>CNRS FRE 3630 (affiliated with Univ. Paris Diderot, Sorbonne Paris Cité), Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France and <sup>2</sup>Architecture et Réactivité de l'ARN, Université de Strasbourg, CNRS, IBMC, 15 rue René Descartes, F-67084 Strasbourg, France\*Corresponding author: Institut de Biologie Physico-chimique, 13 rue Pierre et Marie Curie F-75005 Paris, France. Tél: 158415123; E-mail: [condon@ibpc.fr](mailto:condon@ibpc.fr)  
One sentence summary: This review provides an update of the ribonucleases, their mechanisms of action and their roles in regulation in Gram-positive bacteria.

Editor: Wolfgang Hess

**ABSTRACT**

It is widely recognized that RNA degradation plays a critical role in gene regulation when fast adaptation of cell growth is required to respond to stress and changing environmental conditions. Bacterial ribonucleases acting alone or in concert with various *trans*-acting regulatory factors are important mediators of RNA degradation. Here, we will give an overview of what is known about ribonucleases in several Gram-positive bacteria, their specificities and mechanisms of action. In addition, we will illustrate how sRNAs act in a coordinated manner with ribonucleases to regulate the turnover of particular mRNA targets, and the complex interplay existing between the ribosome, the ribonucleases and RNAs.

**Keywords:** Gram-positive bacteria; ribonucleases; sRNA; mRNA stability

**INTRODUCTION**

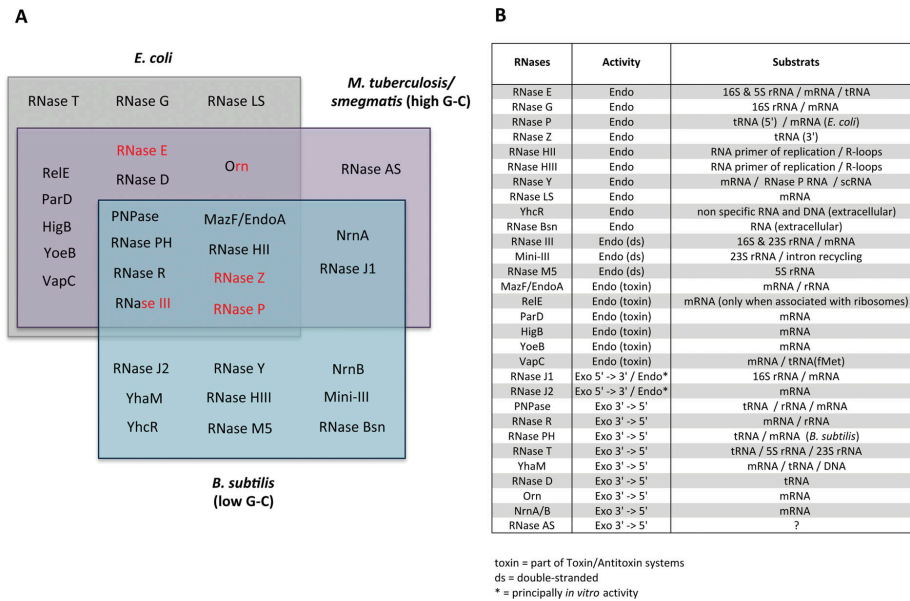
Regulation of mRNA degradation is a key method of controlling gene expression to allow bacteria to adapt to their environmental conditions. The genome sequencing projects of the late 90s revealed significant differences in the degradation machineries of Gram-positive and Gram-negative bacteria (Condon and Putzer 2002). For example, RNase E which is the major ribonuclease involved in mRNA degradation in *Escherichia coli* is absent in *Bacillus subtilis*, the paradigm of Gram-positive organisms (Fig. 1). Moreover, an RNase can be essential in one organism and not in another. It is therefore clear that many of the blueprints that have been established for both regulated and constitutive RNA turnover in the model enterobacterium *E. coli* have to be re-established in the Firmicutes. Recent studies have permitted a detailed characterization of the RNases present in *B. subtilis* that have been extended to other Gram-positive bacteria (reviewed in Condon and Bech-

hofer 2011; Morrison and Dunman 2011; Jester, Romby and Lioliou 2012).

The degradation of any specific mRNA is influenced by a number of factors including its secondary structure and its rate of translation. These parameters can be affected by RNA-binding proteins and, as has been shown more recently, by regulatory RNAs. Regulatory RNAs encoded in *cis* (antisense or asRNA) or in *trans* (small or sRNA) with respect to their targets are now recognized as important actors in the modulation of gene expression. Although these regulatory RNAs act through a wide variety of mechanisms, most of them described in the literature block translation and/or affect the stability of their mRNA targets (Lalaouna *et al.* 2013). The majority of these pioneering studies have been performed in Gram-negative bacteria such as *E. coli* or *Salmonella typhimurium*. In these bacteria, the Sm-like protein Hfq is required to protect the sRNA from degradation and to stimulate basepairing interactions between the sRNA and mRNA (e.g. De Lay and Gottesman 2011; Vogel and Luisi 2011; Régnier

Received: 23 December 2014; Accepted: 1 March 2015

© FEMS 2015. All rights reserved. For permissions, please e-mail: [journals.permissions@oup.com](mailto:journals.permissions@oup.com)



**Figure 1.** Ribonucleases and their functions. (A) Ribonucleases found in *E. coli*, *B. subtilis* and *M. tuberculosis*. Essential RNases are indicated in red. RNase III and Orn were only found essential for cell growth in some bacteria (see the text). In this review, only RNases involved in mRNA degradation have been discussed in detail, such as RNase J1/J2, RNase Y, PNPase and RNase III (see the text). (B) Activity (endo- vs exoribonucleolytic) and substrates of RNases in panel A. The toxins listed are all type II.

and Hajnsdorf 2013; Sauer 2013; Wagner 2013). Hfq has also been proposed to recruit RNase E to the binding site of the sRNA and thus facilitate RNA degradation (Morita and Aiba 2011; Prevost et al. 2011). In Gram-positive bacteria, recent transcriptome data have shown that regulatory RNAs are also widely present, but Hfq does not seem to play a major role (Bohn, Rigoulay and Bouloc 2007; Dambach, Imov and Winkler 2013; Hammerle et al. 2014), again forcing a re-thinking of an established paradigm.

The differences in the mRNA degradation machinery (Fig. 1) and in the mode of action of regulatory RNAs have incentivized their characterization in Gram-positive bacteria to the level currently enjoyed in Gram-negative organisms. In this review, we will describe progress made in the determination of the mRNA degradation pathways in Gram-positive bacteria and how the RNases involved are influenced by the action of regulatory sRNAs.

## mRNA DEGRADATION IN GRAM-POSITIVE BACTERIA

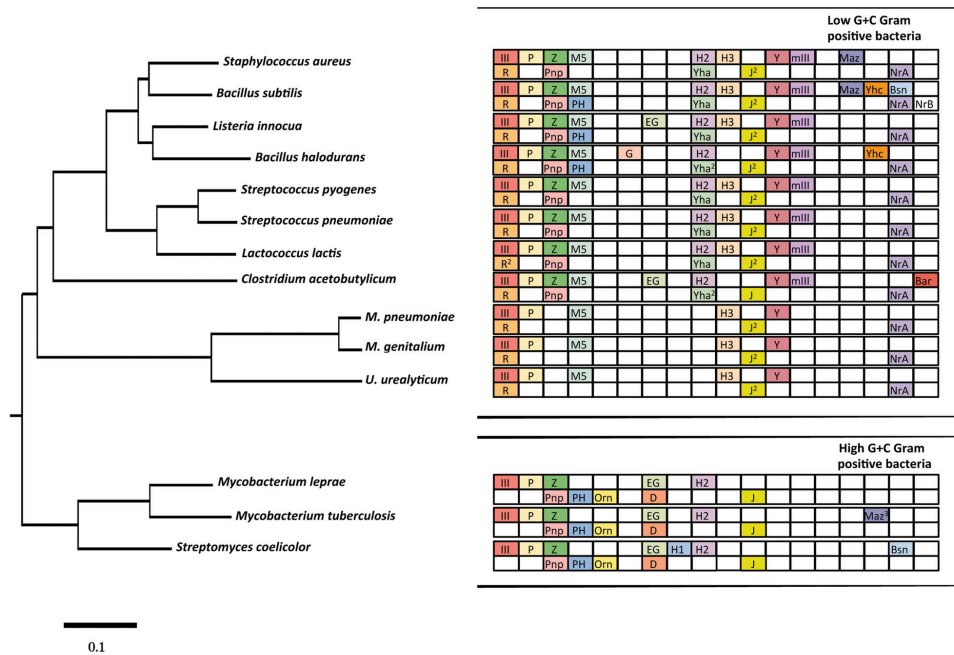
*Escherichia coli* and *B. subtilis* have different sets of endo- and exoribonucleases to degrade mRNAs (Fig. 1). A few of them including 3'-5' exoribonucleases (PNPase, RNase R, RNase PH) and several endoribonucleases (RNase III, RNase P, RNase Z) are conserved in Gram-negative as well as in low GC and high GC Gram-positive bacteria (Fig. 2). In contrast, one of the main RNases (the endoribonuclease Y) involved in mRNA degradation in the low GC Gram-positive bacteria is absent from high GC Gram-positive bacteria such as *Mycobacterium tuberculosis*, which instead encode the endoribonuclease E/G (Figs 1 and 2). Below we will focus

on the major enzymes required for the degradation of mRNAs in Gram-positive bacteria.

## Key enzymes involved in mRNA degradation in Firmicutes

### Ribonuclease Y

The key endoribonuclease of *E. coli* mRNA turnover, RNase E, has been replaced by RNase Y in *B. subtilis* (Fig. 1). RNase Y was first shown to be responsible for the endonucleolytic cleavage between the *cggR* and *gapA* open reading frames of the *gapA* operon (Commichau et al. 2009) and in S-adenosylmethionine riboswitch turnover (Shahbadian et al. 2009) in *B. subtilis*. *CggR* encodes a transcriptional regulator of the *gapA* gene encoding glyceraldehyde 3-phosphate dehydrogenase (GAPDH). RNase Y cleavage of this mRNA allows differential expression of these two genes (100-fold more GAPDH than CggR). Depletion of RNase Y also increased the half-life of bulk RNA more than 2-fold in *B. subtilis* (Shahbadian et al. 2009), the first suggestion that RNase Y could have an important role in global mRNA degradation in this organism, equivalent to RNase E in *E. coli*. Interestingly, both enzymes have similar cleavage specificities (AU-rich single-stranded regions) and are localized to the cytoplasmic membrane. In the case of RNase E, membrane targeting is through an amphipathic helix and the enzyme rapidly diffuses around the inner membrane to form short-lived foci, which have been attributed to transient RNA degradation bodies (Strahl et al. 2015). Such a dynamic clustering of RNase Y has not yet been observed in Gram-positive bacteria, which is targeted to membranes through an N-terminal transmembrane domain.



**Figure 2.** The phylogenetic distribution of Gram-positive ribonucleases. The endonucleases are shown in the top row and the exonucleases in the bottom row of blocks for each species. The phylogenetic relationship between the different organisms was calculated by comparing 16S rRNA sequences using Clustal X. The phylogenetic tree was drawn using Phylodendron (<http://iubio.bio.indiana.edu/soft/molbio/java/apps/trees/>). The abbreviations for the different RNases are as follows: III = RNase III; P = RNase P; Z = RNase Z; M5 = RNase M5; G = RNase G; EG = RNase E/G; H1 = RNase H1; H2 = RNase H2; H3 = RNase H3; Y = RNase Y; mIII = MiniRNase III; Maz = MazF/EndoA; Yhc = YhcR; Bsn = RNase Bsn; Bar = Barnase; R = RNase R; II = RNase II; Pnp = polynucleotide phosphorylase; PH = RNase PH; Om = oligoribonuclease; T = RNase T; D = RNase D; Yha = RNase Yha; J = RNase J; NrA = NanoRNase A; NrB = NanoRNase B. \*Number potential orthologues present. This figure is an updated version of a figure shown in Condon and Putzer (2002).

Like RNase E, RNase Y endoribonuclease activity is thought to be stimulated by a 5' monophosphate, although this has only been documented so far for one substrate, the *yitJ* riboswitch (Shahbadian *et al.* 2009). While RNase Y was originally thought to be essential in *B. subtilis* (as RNase E is in *E. coli*), this was recently shown not to be the case (Figaro *et al.* 2013). *Bacillus subtilis* strains completely lacking RNase Y are viable, but they grow slowly and have pleiotropic phenotypes (failure to sporulate or become competent for DNA uptake, for example). The *my* deletion mutant can be combined with *mc* and *pnp*, but not *rnjA* (Figaro *et al.* 2013), suggesting that it is not possible to inactivate both major endo and 5'-exonucleolytic pathways (see below). Three different global transcriptome analyses of RNase Y depleted *B. subtilis* strains showed that between 13 and 23% of individual protein-encoding genes and many non-coding RNA genes were upregulated, confirming a global role in mRNA turnover for this enzyme.

In *Staphylococcus aureus* and *Streptococcus pyogenes*, the gene-encoding RNase Y was first identified in a screen for mutants with attenuated virulence and called *cvfA* (Kaito *et al.* 2005). Although deletion mutants of *cvfA* have only slightly reduced growth rates in both cases compared to wild-type, transcriptome analyses have led to divergent conclusions as to the global importance of RNase Y in mRNA turnover in these two organisms. In *S. pyogenes*, deletion of the RNase Y/*CvfA* gene led to an up-

regulation of a significant number (14%) of genes in stationary phase (Kang, Caparon and Cho 2010) and a 2-fold increase global mRNA half-life in late-log phase (Chen *et al.* 2013), suggesting a global role for RNase Y similar to that observed in *B. subtilis*. In *S. aureus*, however, only a small subset (4%) of mRNAs and sRNAs were upregulated in the absence of RNase Y (Marincola *et al.* 2012), suggesting that there may be some functional redundancy with another endoribonuclease in this organism. Although some Firmicutes, such as *Listeria* and *Clostridia*, do have an ortholog of the RNase E catalytic domain (RNase E/G) to potentially provide such functional redundancy, *S. aureus* is not one of these organisms (Fig. 2). In *S. aureus*, RNase Y is required for the processing and stabilization of the transcript encoding the global virulence regulatory system SaeRS, and was shown to activate the synthesis of exotoxins independently of the *agr* and *sae* pathways by an as yet undefined mechanism (Marincola *et al.* 2012).

#### Ribonucleases J1 and J2

RNases J1 and J2 have been shown to form a 5'-3' exoribonuclease complex in *B. subtilis* involved in the 5'-3' degradation of mRNAs and in the maturation of the 5'-end of 16S ribosomal RNA (Mathy *et al.* 2007). Indeed, RNase J1 was the first 5'-3' exoribonuclease identified in prokaryotes and its discovery explained how mRNA can be greatly stabilized by a stalled ribosome or a

secondary structure near the 5' end, even if the rest of the mRNA is completely free of ribosomes (Bechhofer and Zen 1989). Although both RNase J1 and J2 have been shown to additionally have endoribonuclease activity *in vitro*, they are thought to act primarily as a 5'-3' exoribonuclease complex *in vivo*, with RNase J1 providing most of the activity and RNase J2 helping to stabilize (or regulate) RNase J1 in both *S. aureus* (Linder, Lemeille and Redder 2014) and *B. subtilis* (Mathy et al. 2010). RNase J2 appears to play a more important role in *S. mutans*, where it has been proposed to act as an endonuclease independently of RNase J1 (Liu et al. 2015). Most Firmicutes have at least two RNase J paralogs, while outside of this phylum typically only one ortholog is observed per genome. In *B. subtilis*, RNase J1 was originally thought to be essential until it was shown that knockout of the gene was possible (Figaro et al. 2013). The lack of RNase J1 also causes pleiotropic effects in cell growth, cell morphology and in the development of competence and sporulation, while *B. subtilis* RNase J2 mutants grow normally. In *S. pyogenes*, both RNase J1 and J2 mutants are thought to be non-viable (Bugrysheva and Scott 2010), while in *S. aureus* and *S. mutans*, neither enzyme is essential, although in the former growth is restricted to a very narrow window around 37°C (Linder, Lemeille and Redder 2014; Chen et al. 2015; Liu et al. 2015). Synthetic lethality of the *B. subtilis* *mjA* deletion has been tested with a number of the other key RNase mutants: *mjA mc* mutants are viable (unpublished results); while we were unable to make *mjA pnp* or *mjA my* double mutants (Figaro et al. 2013). The inability to delete both the *pnp* and *mjA* genes in the same strain suggests that cells cannot function in the absence of one or other of the major exonucleolytic (5' or 3') pathways.

*Bacillus subtilis* cells severely depleted for RNase J1 show increased abundance of about 21% of its transcripts (Durand et al. 2012a), considerably more than an earlier study performed under milder depletion conditions (Mader et al. 2008). A global role for the exoribonuclease activity of RNase J1 in mRNA turnover was also seen in *S. aureus*, as well as for the correct maturation of the 5' end of both the 16S rRNA and the RNA subunit of RNase P (Linder, Lemeille and Redder 2014). Degradation of a number of specific mRNAs was shown to be dependent on both RNase J1 and J2 in *S. pyogenes* (Bugrysheva and Scott 2010), but a global analysis of the role these enzymes in mRNA decay has not yet been performed.

### Ribonuclease III

Almost all bacteria have the double-strand-specific endoribonuclease RNase III, with *Deinococcus radiodurans* being a notable exception (Drider and Condon 2004). In the archaea, RNase III activity has been replaced by bulge-helix-bulge nuclease, which has similar specificity (Heinemann, Soll and Randau 2010). The catalytic domain of RNase III is also found conserved in the two enzymes Dicer and Drosha, which are involved in siRNA and miRNA processing in eukaryotes (Jaskiewicz and Filipowicz 2008). Bacterial RNase III is primarily known for its role in the processing of tRNAs. However, the enzyme is non-essential in most organisms studied so far with the notable exception of *B. subtilis*, where RNase III is required to prevent expression of two prophage-encoded type I toxins via asRNA regulation. Deletion of RNase III has limited impact on mRNA expression levels in both *B. subtilis* and *E. coli*, with only 11–12% of mRNAs affected in either organism (Stead et al. 2011; Durand et al. 2012a). In *B. subtilis*, many of these effects were shown to be indirect and occur at the transcriptional level, so the actual number of mRNAs directly cleaved by RNase III is considerably smaller (Durand et al. 2012a). In *S. pyogenes*, RNase III is a key partner required for

the RNA-mediated immunity CRISPR/Cas system against phages and plasmids (Deltcheva et al. 2011). A small sRNA called tracrRNA directs RNase III-mediated maturation of the short repeat-spacer-derived crRNA to silence the foreign DNA in a sequence-specific manner.

Co-immunoprecipitation of RNAs bound to RNase III identified a significant number of specific sRNA and mRNA substrates of this enzyme in *S. aureus* (Lioliou et al. 2012). In addition, *S. aureus* RNase III was shown to contribute to the correct maturation of rRNAs and tRNAs, to autoregulate its synthesis by cleaving the coding region of its own mRNA, to enhance the stability of *cspA* mRNA by cleaving in the 5' UTR and to cut overlapping 5' UTRs of divergently transcribed genes. In this organism, RNase III has also been proposed to play a role in the removal of background transcriptional noise from both strands of the chromosome (Lasa et al. 2011), a phenomenon not clearly seen in *B. subtilis* (Durand et al. 2012a) but recently observed in *E. coli* (Lybecker et al. 2014b). Whether these cryptic transcripts possess regulatory functions or are by-products of transcription events awaits further experimental work (Lasa, Toledo-Arana and Gíngeras 2012; Lybecker, Bilusic and Raghavan 2014a). Interestingly, it was recently suggested that pervasive transcription might be considered as a genome surveillance mechanism for DNA damage, enabling a robust action of the nucleotide excision repair machinery (Kamathapu and Nudler 2015). Studies performed in *S. aureus* and *Listeria monocytogenes* also revealed the unexpected but common presence of mRNAs with overlaps of either their entire length or over their 5' or 3'-untranslated regions, which were degraded in a RNase III-dependent manner (Lasa et al. 2011; Lioliou et al. 2012; Ruiz de los Mozos et al. 2013). In line with these observations, a new gene organization has been discovered in *L. monocytogenes*, the so-called excludon, where overlapping UTRs regulate the expression of divergent genes encoding proteins with opposing functions (Wurtzel et al. 2012). Besides antisense regulation, RNase III plays also an important role in the regulation of some specific mRNAs by sRNAs in the Firmicutes, similar to observations in the Enterobacteria (see examples below).

### PNPase

In 1996, Luttinger, Hahn and Dubnau (1996) performed a mini Tn-10 insertion screen to identify mutants of *B. subtilis* impaired in competence development. One of the genes identified was *comR*, which was renamed *pnpA* due to its strong sequence similarity to the PNPase gene of *E. coli*. Biochemical studies suggested that PNPase plays an important role in mRNA degradation in *B. subtilis* (Deutscher and Reuven 1991; Wang and Bechhofer 1996) and the degradation products of several mRNAs, such as *ermC*, *rpsO* or the *trp* RNA leader (Bechhofer and Wang 1998; Oussenko et al. 2005; Deikus and Bechhofer 2009) were shown to be dependent on PNPase. A recent global transcriptome analysis (Liu et al. 2014) confirmed that PNPase is an important player in global mRNA degradation in *B. subtilis*. This work showed that the steady-state levels of 10% of *B. subtilis* mRNAs (412 genes) were increased >1.5-fold in the *pnpA* mutant strain. Among these mRNAs, 178 were highly dependent on PNPase activity and none of the other 3' exoribonucleases could replace PNPase in the degradation process. The number of direct mRNA targets of PNPase was probably underestimated for different reasons. First, only half of the *B. subtilis* genes were detected in the condition of the study. Second, PNPase is likely to have some functional redundancy with the three other 3'-5' exoribonucleases in the cell. Last, the effect of PNPase on some mRNAs might be masked if the first endoribonucleolytic cleavage, which



allows access to PNPase, occurs early in the mRNA. In such cases, the average signal over the whole transcript might not be significantly affected.

A role of PNPase in regulating virulence gene expression in *S. aureus* has been observed recently (Numata *et al.* 2014). Surprisingly, the phenotype of the mutant strain deficient in RNase Y, i.e. decreased hemolysin production, was suppressed by disrupting the *pppA* gene. It has been suggested that RNase Y, in addition to being an endoribonuclease, can convert RNAs bearing 2'-3' cyclic phosphate groups (resulting from cleavage by toxin-type RNases) into 3'-phosphorylated RNAs, which are far more resistant to PNPase activity than RNAs with 2'-3' cyclic phosphates. The authors thus propose a model whereby RNase Y and PNPase competitively interfere with the degradation of some mRNAs involved in virulence (Numata *et al.* 2014). An alternative explanation can be imagined, however. If the initial destabilizing cleavage by RNase Y occurred in the 3'-UTR, the lack of degradation by PNPase might stabilize the cleaved RNA sufficiently to yield a protein product.

As in *E. coli*, PNPase from *B. subtilis* is sensitive to RNA secondary structure (Deikus and Bechhofer 2007). *Bacillus subtilis* has at least three more 3'-5' exoribonucleases that help to degrade mRNAs: RNase PH acts by phosphorolysis like PNPase, while RNase R and YhaM are hydrolytic enzymes. RNase R has been shown to help PNPase to degrade structured RNAs (Oussenko *et al.* 2005). Moreover, *E. coli* PNPase activity can be helped both by action of the helicase RhlB in the context of the degradosome and by the polyadenylation of the 3' end of mRNA (Khemici *et al.* 2004) synthesized either by the poly(A) polymerase or by PNPase. Polyadenylation of mRNA has also been observed in *B. subtilis*, but the enzyme(s) responsible for this phenomenon and its significance are unknown (Cao and Sarkar 1993; Campos-Guillen *et al.* 2005).

Interestingly, a recent genetic screen in *E. coli* has shown that PNPase plays an important role in sRNA regulation by protecting these molecules from degradation by RNase E or other ribonucleases, by binding without degradation (De Lay, Schu and Gottesman 2013). PNPase was also shown to be responsible for sRNA degradation when they are not associated with Hfq (Andrade *et al.* 2012). Nothing is known in Gram-positive bacteria about the role of PNPase in sRNA regulation. However, recent work in *L. monocytogenes* revealed a unique CRISPR-like sRNA (clustered regularly interspaced short palindromic repeats) whose DNA interference activity and RNA turnover depend on PNPase, although the mechanism is not yet fully understood (Sesto *et al.* 2014).

#### Enzymes involved in mRNA degradation in Actinobacteria

Although the degradation machinery is remarkably different between low GC Gram-positive and Gram-negative bacteria, Gram-positive bacteria with a high GC content such as the Actinobacteria *M. tuberculosis* and *M. smegmatis* have a mixture of RNases found in both types of bacteria (Figs 1 and 2). Indeed, the Actinobacteria encode like *E. coli*, RNase E/G, RNase D, oligo-RNases (Orn) and a plethora of type II toxins with RNase activity. They also have the 5'-3' exoribonuclease J and the nano-RNase (NmA) found in *B. subtilis* (Taverniti *et al.* 2011; Grosjean *et al.* 2014). RNase E/G is essential in *M. tuberculosis* while RNase J mutants are viable, suggesting a broader role of RNase E in mRNA degradation than RNase J, even if this hypothesis needs to be confirmed (Taverniti *et al.* 2011). RNase III and RNase J have been given particular attention in *Streptomyces* species because of

their impact on the expression of genes involved in secondary metabolism and antibiotic production. The mechanisms have not yet been clearly established and tend to suggest indirect effects (Gatewood *et al.* 2012; Lee, Gatewood and Jones 2013; Jones *et al.* 2014). RNase III is, however, responsible for rRNA maturation in *Streptomyces* and this was suggested to be critical for the translation of large multicistronic mRNA transcripts (for a review, see Liu *et al.* 2013).

#### Pathways of mRNA degradation in Firmicutes

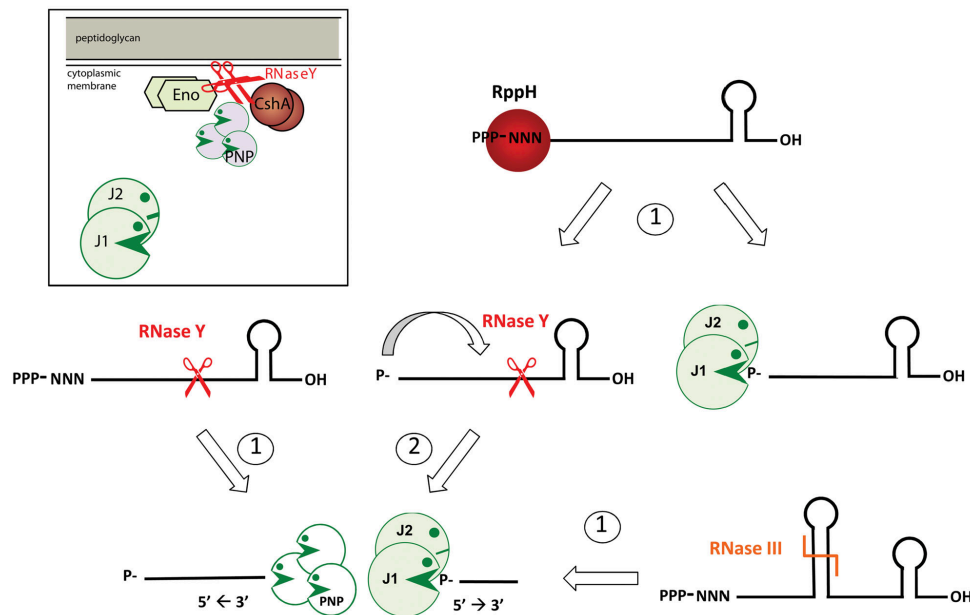
Primary mRNA transcripts in bacteria are protected at their 5'-end by a triphosphate group, known to inhibit the activity of many RNases (both endo- and exoribonucleases) involved in mRNA degradation (Fig. 3). They are also typically protected at their 3'-end by a Rho-independent terminator that blocks 3'-5' exoribonuclease attack. Initiation of mRNA degradation must therefore override one of these two protective elements. Based on global transcriptome data and detailed studies on individual mRNAs, several different pathways of mRNA degradation have been proposed in Gram-positive bacteria. One degradation initiation pathway is similar to that found in Gram-negative bacteria but with different enzymes. In this case, the limiting step is an endoribonuclease cut to render the mRNA accessible to exoribonucleases or yet other endoribonuclease cleavages. In *B. subtilis* and other Firmicutes, this endoribonuclease can be RNase Y, specific for single-stranded RNA, or to a lesser extent RNase III, which cleaves double-stranded RNA (Fig. 3). These RNases leave a 5' monophosphate extremity which can then be attacked by the 5'-3' exoribonuclease RNase J1/J2 complex or cleaved further by RNase Y. The 3'-end of the upstream cleavage product is degraded by 3'-5' exoribonucleases, principally PNPase in *B. subtilis* (Fig. 3).

In the alternative degradation pathway, the 5' triphosphate of the mRNA can be converted to a 5'-monophosphate by an RNA pyrophosphohydrolase (RppH) (Hsieh *et al.* 2013), first discovered in *E. coli* (Deana, Gelesnik and Belasco 2008). After removal of the triphosphate, the mRNA can be degraded either directly by the 5'-3' exoribonuclease RNase J1/J2 complex or potentially be subjected to stimulated cleavage by RNase Y (Fig. 3). Analysis of the growth phenotypes of *rppH rny* and *rppH mjA* double mutant strains compared to the respective single RNase mutants suggested that RppH may preferentially act in the same degradation pathway as RNase J1. Indeed, the *rppH mjA* double has the same doubling time as a single *mjA* mutant, whereas the doubling time of the *rppH rny* double mutant is significantly greater than that of the *rny* strain alone (Figaro *et al.* 2013).

Importantly, not all mRNAs are substrates for RppH. In *B. subtilis*, this enzyme has a strong preference for a G-residue in the second position of its mRNA substrate and requires at least three unpaired nucleotides (nts) to act efficiently (Hsieh *et al.* 2013; Piton *et al.* 2013). Preliminary data also suggest that other RppH exist in *B. subtilis* that could affect a larger number of mRNAs (Hsieh *et al.* 2013; Condon *et al.* unpublished results).

#### Protein complexes involved in RNA degradation

Messenger RNA degradation can be performed by RNases in complex with other proteins, as has been demonstrated with the discovery of the RNase E-based degradosome in *E. coli*. These interactions are known to optimize mRNA degradation through allosteric activation in some cases and, in others, concentrate different activities in the same subcellular location. Several RNase



**Figure 3.** A schematic view of the pathways involved in RNA degradation. Insert: a degradosome-like complex has been proposed in *B. subtilis* structured primarily around RNase Y (Lehnik-Habrink et al. 2011). J1 and J2 are dual endo- and 5'-3' exo-ribonucleases, PNP is the 3' exoribonuclease, Eno is for enolase and CshA is the major RNA helicase associated with the degradosome. The organization of this degradosome-like complex seems to be also conserved in *S. aureus* (Roux, DeMuth and Dunman 2011). Primary mRNA transcripts in bacteria are protected at their 5'-end by a tri-phosphate group. Initiation of mRNA degradation can involve an endoribonuclease (RNase Y or RNase III) cut, which is the limiting step. This step generates a 5' monophosphate extremity, which can be attacked by the 5'-3' exoribonuclease, RNase J1/J2 complex or cleaved further by RNase Y. The 3'-end of the upstream cleavage product is degraded by 3'-5' exoribonucleases, principally PNPase in *B. subtilis*. In the alternative degradation pathway, the 5' tri-phosphate of the mRNA can be converted to a 5'-monophosphate by an RppH. After removal of the tri-phosphate, the mRNA can be degraded by the 5'-3' exoribonuclease RNase J1/J2 complex or by RNase Y.

complexes have been proposed in the Firmicutes, structured around enzymes other than RNase E (Fig. 3, insert).

#### The RNA degradosome

In *E. coli*, and in other Enterobacteria, the RNA degradosome is organized around RNase E, which has an N-terminal catalytic domain (~500 amino acids) and an equally large intrinsically unfolded C-terminal domain, peppered with structured microdomains that can interact with accessory proteins (Ait-Bara, Carpousis and Quentin 2014). *Escherichia coli* RNase E principally binds PNPase, the DEAD-box RhlB helicase and the glycolytic enzyme enolase. However, the composition of the RNase E-based degradosome shows some degree of plasticity from organism to organism and from one growth condition to another. The *Caulobacter crescentus* degradosome, for example, contains aconitase instead of enolase (Hardwick et al. 2011). In *E. coli*, RhlB, CsdA and RhlE helicases are interchangeable in vitro (Khemici et al. 2004), and RhlB can be replaced by CsdA in the cold (Prud'homme-Genereux et al. 2004). Furthermore, a number of proteins have been shown to bind non-stoichiometrically to the degradosome in *E. coli*, to modulate or regulate its activity. In the high GC Gram-positive Actinobacterium, *Streptomyces coelicolor*, the PNPase interacting domain is found at the N-terminus of the RNase E (Lee and Cohen 2003), while the catalytic domain occupies a central location in the protein. In *M. tuberculosis*, RNase E was found to interact with an inor-

ganic polyphosphate/ATP-NAD kinase (Ppnk), an acetyltransferase and GroEL but the meaning of these interactions is unknown (Kovacs et al. 2005).

In *B. subtilis*, a degradosome-like complex has been proposed, structured primarily around RNase Y (Fig. 3, insert). A number of pairwise interactions were identified by bacterial two-hybrid (B2H) assay, cross-linkage followed by strep-tagged pull-down assay and surface plasmon resonance (SPR) analysis. Like RNase E, RNase Y has an intrinsically unstructured domain, in this case from amino acids 30 to 200 that also have a predicted propensity to form a coiled-coil. RNase Y principally forms dimers in vitro. Although the coiled-coil domain and the transmembrane domains form the strongest self-interaction in B2H assays, it is clear that all domains, the KH domain, the HD domain and the C-terminal domain of RNase Y contribute to its dimerization (Lehnik-Habrink et al. 2011). Only the full-length RNase Y protein gives a positive B2H reaction with its proposed partners below (Lehnik-Habrink et al. 2011). In this review, we will only deal with the more convincing of these interactions.

A reciprocal interaction was first seen between *B. subtilis* RNase Y and enolase by B2H assay. Although a positive interaction between enolase and the negative leucine-zipper control was seen in the same test and many laboratories systematically eliminate enolase as a false positive in two-hybrid assays because of its high abundance in the cell, the interaction between RNase Y and enolase was further supported by

cross-linked pull-down assay (Lehnik-Habrink *et al.* 2011) and by SPR analysis (Newman *et al.* 2012). The RNase Y-enolase interaction seems to be conserved in *S. pyogenes* as demonstrated by cross-linked pull-down assay (Kang, Caparon and Cho 2010) and in *S. aureus* by B2H (Roux, DeMuth and Dunman 2011).

A reciprocal B2H interaction was also observed between *B. subtilis* RNase Y and both PNPase (Commichau *et al.* 2009) and the CshA RNA helicase (Lehnik-Habrink *et al.* 2010). The RNase Y-PNPase interaction was confirmed by SPR, and the equilibrium dissociation constant is in the nanomolar range (Newman *et al.* 2012). The specificity of the RNase Y-CshA interaction was also confirmed in a domain swapping experiment with CshB helicase, that showed that the C-terminal domain of CshA is responsible for the protein-protein interaction (Lehnik-Habrink *et al.* 2010). Although it has not been shown that RNase Y can interact with more than one of these partners at a time, it is nonetheless an intriguing case of convergent evolution that *B. subtilis* and *S. aureus* can potentially form an RNase Y-based degradosome containing enolase, PNPase and the CshA RNA helicase. The stoichiometry of such a complex (unlike the *E. coli* degradosome, the partners of the putative *B. subtilis* complex have only been seen in Western blots), and whether any of these interactions are functionally relevant, remained to be experimentally determined. Although weak interaction between CshA and RNase Y was detected by B2H assay in *S. aureus*, no evidence for a direct interaction between PNPase and RNase Y was observed, suggesting that complex formation varies within the Firmicutes (Roux, DeMuth and Dunman 2011).

CshA has additionally been proposed to interact with a number of other partners in B2H assays and/or cross-link pull-down assays. These include phosphofructokinase, enolase, PNPase, RNase J1, and the DEAD box helicases CshB and DeadYxiN in *B. subtilis* (Lehnik-Habrink *et al.* 2010) and the protein subunit of RNase P (RnpA) in *S. aureus* (Olson *et al.* 2011; Roux, DeMuth and Dunman 2011). This is quite a large number of potential partners for the 55 kDa CshA protein and one wonders whether some of the weaker positive interactions, seen in the B2H assays or in cross-linked pull-down assays, are not simply tethered through RNA. A number of additional minor pairwise interactions identified by these techniques have led authors to present models of large degradosome complexes in *B. subtilis* with up to eight interacting protein partners that seem very premature (Lehnik-Habrink *et al.* 2010; Lehnik-Habrink *et al.* 2011; Roux, DeMuth and Dunman 2011). Studies performed on *S. aureus* CshA have suggested that the RNA helicase contributes to mRNA degradation (Oun *et al.* 2013). Indeed, deletion of the *cshA* gene resulted in dysregulation of biofilm formation and hemolysis due to the dysregulation of *agr* mRNA stability, encoding the quorum-sensing system (Oun *et al.* 2013).

#### The RNase J1/J2 complex

RNase J1 and RNase J2 are present in similar numbers (2500–3000 molecules cell<sup>-1</sup>) in *B. subtilis* and copurify in stoichiometric amounts from *B. subtilis* or when co-expressed in *E. coli*, suggesting that the RNase J1/J2 complex is the primary form of these enzymes *in vivo* (Mathy *et al.* 2010). The complex has been confirmed by strong positive interactions in numerous B2H and yeast two-hybrid (Y2H) assays in both *B. subtilis* and *S. aureus* (Commichau *et al.* 2009; Mathy *et al.* 2010; Roux, DeMuth and Dunman 2011). The complex consists primarily of heterotetramers at high concentrations *in vitro*, but easily dissociate to heterodimers, which maybe the more relevant form at physiological concentrations. As mentioned above, the primary role of *B. subtilis* and *S. aureus* RNase J2 appears to stabilize or modu-

late the activity of RNase J1 (Mathy *et al.* 2010; Linder, Lemeille and Redder 2014; Gilet *et al.* 2015). There has been also some controversy about a direct interaction of RNase J1 with RNase Y in the proposed degradosome complex. An interaction has been proposed based on B2H and cross-linked pull-down assay (Commichau *et al.* 2009; Lehnik-Habrink *et al.* 2011). However, extensive Y2H (RNase J1 as bait) or Y3H screens (RNase J1/J2 complex as bait) failed to identify any interacting partners other than RNase J1 and J2 (Mathy *et al.* 2010). Furthermore, no evidence for additional interacting partners was seen in either a FLAG-tagged pull-down assay of RNase J1 (Mathy *et al.* 2010) or in direct measurements of a potential interaction between RNase Y and RNase J1 by SPR (Newman *et al.* 2012). Finally, no interaction was seen in B2H assays of *S. aureus* RNase Y and either RNase J1 or J2. All these negative results lead us to believe that RNase J1/J2 and the putative RNase Y-based degradosome act as independent complexes in the Firmicutes (Fig. 3, insert).

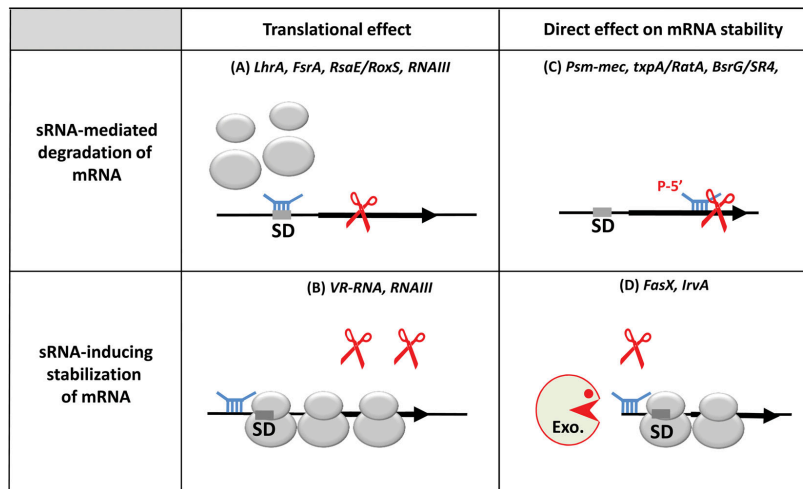
### MODULATION OF mRNA DEGRADATION BY NON-CODING RNAS (sRNA AND asRNA)

To initiate degradation, RNases require access to the mRNA and have to deal with secondary structure, RNA-binding proteins and, particularly, translating ribosomes that potentially obscure RNase cleavage sites. An obvious strategy to avoid competition with ribosomes is to block translation. Regulatory RNA seems to be the perfect player for this role (Figs 4 and 5). Indeed, small regulatory RNAs, and in particular those acting *in trans* (sRNAs) often block translation by basepairing interactions with the Shine-Dalgarno (SD) sequence of mRNA targets. The mRNA free of ribosomes can then be degraded by ribonucleases. Numerous sRNAs seem to act in this way in Gram-positive bacteria and modulate mRNA degradation indirectly (Brantl and Bruckner 2014). Here, we present specific cases where links between the degradation machinery and sRNAs have been revealed (Figs 4 and 5). For a complete list of sRNAs studied in the Firmicutes and their modes of action, see the recent review (Brantl and Bruckner 2014). We will refer to small non-coding RNAs that form imperfect duplexes with their mRNA targets, as sRNAs, and to anti-sense RNAs, which form fully complementary interactions with the mRNAs since they are encoded on the same gene locus, as asRNAs.

#### Modulation of mRNA degradation via modification of mRNA translation: destabilization

##### *Listeria monocytogenes* LhrA sRNA

LhrA was identified in *L. monocytogenes* due to its binding to Hfq. The half-life of LhrA is >30 min in wild-type strains and is decreased to <3 min in a  $\Delta$ hfq mutant strain (Christiansen *et al.* 2004). This sRNA was shown to post-transcriptionally regulate the *lmo0850* mRNA, encoding a small peptide of unknown function, by basepairing interactions close to the ribosome binding sites (RBS). Toeprint and  $\beta$ -galactosidase assays showed that LhrA blocks translation initiation and also reduces the level of *lmo0850* mRNA (Nielsen *et al.* 2010) (Fig. 4A). The RNases involved in either the turnover of the LhrA sRNA or the target *lmo0850* mRNA have not yet been characterized. Interestingly, *L. monocytogenes* has both a short form RNase E homolog and RNase Y, as well as the 5'-3' exoribonucleases J1/J2. The RNase E/G protein of *Listeria* lacks the C-terminal domain where Hfq is thought to bind in *E. coli*. This observation raises the question of whether



**Figure 4.** Various ways that small RNAs (sRNAs) regulate mRNA stability. (A and B) Degradation can be a consequence of the effect of sRNA on translation. (A) The repression of translation is often subsequently followed by rapid mRNA degradation while (B) the activation of mRNA translation protects the mRNA from RNases. (C) Binding of sRNA to mRNA recruits a specific RNase to destabilize the mRNA. (D) Binding of sRNA can induce a specific mRNA processing site in the 5' UTR that leads to stabilizing effect. Alternatively, binding of the sRNA (or regulatory mRNA) can prevent the access of an RNase to prevent the degradation of target mRNA.

and how Hfq interacts with RNase E/G or other RNases in this organism.

#### *Bacillus subtilis* FsrA sRNA

The FsrA sRNA was identified in *B. subtilis* and is involved in the iron-sparing response (Gaballa et al. 2008). The level of several mRNAs encoding genes linked to central metabolism, such as *citB* and *sdhCAB*, are upregulated in an FsrA mutant strain (Smaldone et al. 2012). The ribonucleases involved in these regulatory phenomena are still unknown. Interestingly, FsrA requires a number of small basic proteins (FbpA, B and C) to act efficiently on its targets. For example, the regulation of the *lutABC* operon by FsrA requires the small protein FbpB for full repression. The authors proposed that FbpA, B and C could fulfill the role of Hfq to stimulate basepairing between FsrA and *lutABC* and possibly even attract the degradation machinery. In particular, the FbpB protein plays a greater role than FsrA in regulating the levels of the *lutABC* mRNA (Smaldone et al. 2012). The mode of action of FsrA is still speculative but its potential interaction with the RBS of the *lutA*, *citB* and *sdhCAB* mRNAs suggested that FsrA would block their translation to provoke mRNA degradation (Fig. 4A).

#### RsaE/RoxS sRNA

RsaE was first identified in *S. aureus* and this is the sole trans-acting sRNA to be conserved in *B. subtilis*, apart from the ubiquitous 6S RNA (Geissmann, Marzi and Romby 2009; Bohn et al. 2010). This RNA regulates several targets linked to folate and central metabolism in *S. aureus* (Geissmann, Marzi and Romby 2009; Bohn et al. 2010). In *B. subtilis*, the RsaE homolog (which is called RoxS) also regulates some targets linked to central metabolism, but plays an even greater role in controlling genes involved in oxidative stress or redox homeostasis in response to nitric oxide (NO) (Durand et al. 2015). In most cases studied, RsaE/RoxS is predicted to bind the SD sequences of target mRNAs in *S. aureus* and *B. subtilis* leading to translational inhibition

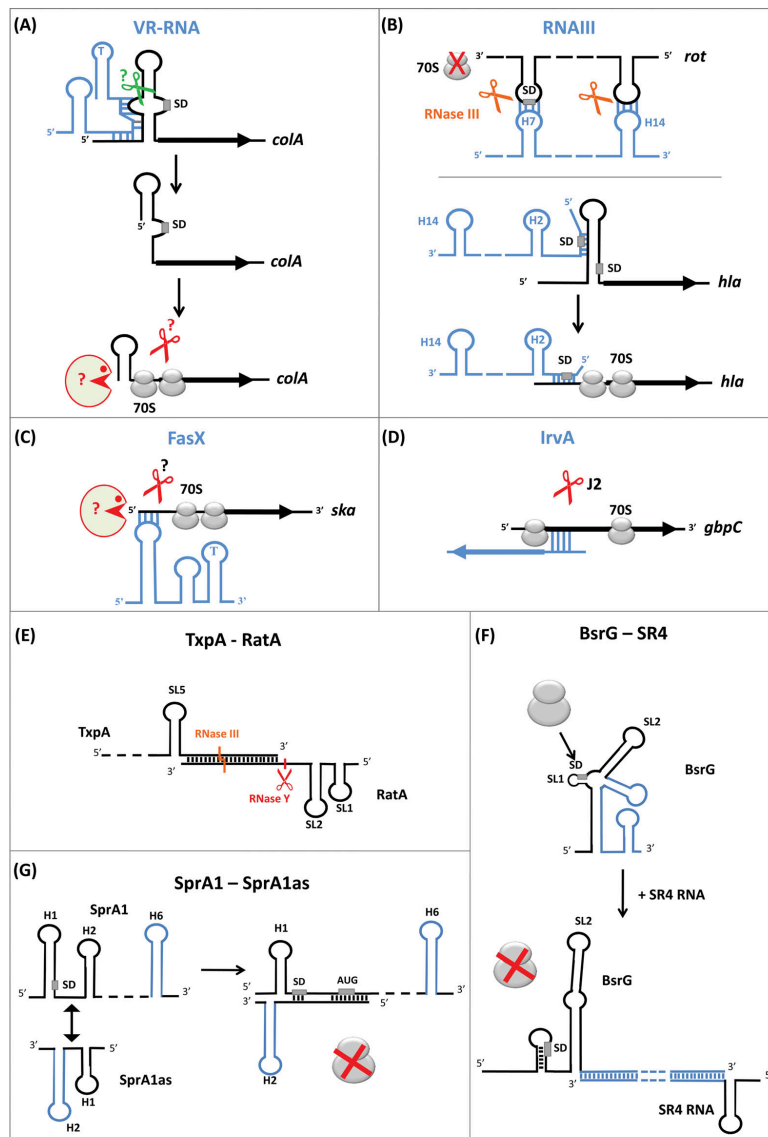
(Fig. 4A). However, this may reflect a bias in target prediction programs that often search for basepairing interactions around translation initiation sites. In *B. subtilis*, RNase Y and RNase III are important enzymes involved in the regulation of target mRNAs stability in response to RoxS binding (Durand et al. 2015). More unexpectedly, RNase Y intervenes at an additional level by processing the 5' end of RoxS removing about 20 nts. Processing of RoxS was shown to expand the repertoire of targets recognized by this sRNA in *B. subtilis* (Durand et al. 2015). This study reveals a complex interplay between RNases and RoxS to regulate its mRNA targets.

#### Modulation of mRNA degradation via modification of mRNA translation: stabilization

Under specific conditions, sRNAs can also stabilize mRNAs using a variety of mechanisms. The most common mechanism involves a conformational change of the mRNA structure upon binding of the sRNA that liberates the RBS to recruit the ribosome and to enhance translation (Fig. 4B). The activation of translation then protects the mRNA from degradation. In a second mechanism, binding of the sRNA creates or reveals a specific processing site of the mRNA, which renders the RBS accessible for ribosome binding and stabilizes the mRNA (Fig. 4B). The VR-RNA in *Clostridium perfringens* and RNAIII in *S. aureus* use these strategies and will be described below.

#### *Clostridium perfringens* VR-RNA

The VR-RNA is a 386-nucleotide sRNA encoded in the genome of *C. perfringens* (Obana et al. 2010). This sRNA binds the 5' UTR of the *colA* mRNA, encoding a collagenase. Binding of the sRNA triggers cleavage of the *colA* mRNA at the 3' edge of the VR-RNA binding site, which has two consequences: (1) it renders the SD sequence accessible to the ribosome and (2) it creates a stable stem-loop structure at the 5' end of the mRNA with no 5'



**Figure 5.** Mechanisms of sRNA-mediated regulation in various Gram-positive bacteria. (A) *Clostridium perfringens* VR-RNA binds the 5' UTR of the *colA* mRNA, encoding a collagenase and triggers cleavage of the *colA* mRNA, which in turn stabilizes the mRNA (Obana et al. 2010). The sRNA is in blue, the mRNA target is colored in black, SD is for Shine-Dalgarno sequence, 70S is the ribosome, RNases are represented by scissors. (B) *Staphylococcus aureus* RNase III represses the translation of *rot* mRNA by sequestering the SD sequence, and activates the translation of *hla* mRNA through mRNA conformational changes (Novick 2003). (C) *Streptococcus pyogenes* FasX sRNA forms basepairing interactions with the first 9 nts of the *ska* mRNA to create a stable RNA helix at the 5' end of the mRNA (Ramirez-Pena et al. 2010). (D) The 5' UTR of the *S. mutans* *irvA* mRNA basepairs with the coding sequence of *gbpC* mRNA to block cleavage by RNase J2 (Liu et al. 2015). (E) *Bacillus subtilis* type I toxin/antitoxin system *TxpA/RatA* (Silvaggi, Perkins and Losick 2005). The 3' ends of *RatA* forms a large duplex with the *txpA* mRNA, which is cleaved by RNase III. (F) *Bacillus subtilis* *BsrG/SR4* is a temperature-dependent type I toxin-antitoxin system (Jahn et al. 2012). The basepairing interactions between *SR4* and *bsrG* mRNA (in blue) lead to structural changes of the mRNA around the SD, preventing the ribosome (70S) binding. (G) The short asRNA, *SprA1-AS*, binds to the RBS of *SprA1* through imperfect basepairing interactions to prevent translation of a small toxic peptide (Sayed, Jousselein and Felden 2011). The Rho-independent terminator hairpin of *SprA1-AS* (in blue), which is fully complementary to the 3' end of *SprA1*, does not interact with *SprA1*.

overhang (Figs 4B and 5A). These modifications allow stabilization of the *colA* mRNA. Indeed, the half-life of the mRNA is four minutes in a wild-type strain but decreases to less than two minutes in a  $\Delta$ VR-RNA strain (Obana et al. 2010). This study also shows that ribosome binding is essential to protect the *colA* mRNA from degradation. The identity of the RNase responsible for the cleavage after VR-RNA binding is still unknown but RNase III is not involved in this process (Obana et al. 2010). RNase Y could be a good candidate for this cleavage.

#### **Staphylococcus aureus RNAIII**

RNAIII is longRNA (514 nts) that was identified in *S. aureus* as one of the main intracellular effectors of the *agr* quorum-sensing system (Novick et al. 1993). This RNA encodes  $\delta$ -hemolysin but also acts as a regulatory RNA. RNAIII possesses 14 hairpin motifs and three of them, all containing stretches of consecutive C-residues, are used to fulfill its repressor activity. The expression of several genes is negatively regulated by RNAIII (*spa*, SA1000, *rot*, *lytM* and *coa*) using a shared mechanism, i.e. blocking translation upon basepairing interactions with the SD sequence (Boisset et al. 2007) (Figs 4A, B and 5B). Depending on the mRNA signals, the repressor RNAIII-mRNA complexes adopt different topologies. For instance, loop 13 of RNAIII binds to the SD sequence of the *spa* mRNA to form a long imperfect duplex while the two apical loops H7 and H14 of RNAIII form loop-loop interactions with two G-rich loops including the SD sequence of *rot* mRNA (Fig. 5B). In both cases, the complexes formed are accompanied by RNase III cleavages, leading to a functional inactivation of the mRNAs.

In contrast, RNAIII also increases the level of the *hla* mRNA encoding  $\alpha$ -hemolysin (Morfeldt et al. 1995; Novick 2003). Although RNAIII is most probably consumed together with its repressed mRNA targets (Boisset et al. 2007), the yield of RNAIII is sufficiently abundant at the late-exponential growth phase to activate the translation of *hla* mRNA. In this case, the 5' part of RNAIII binds to the 5' UTR of *hla* to impair the formation of a secondary structure which normally traps the SD sequence of the *hla* gene (Fig. 5B). The basepairing interactions with RNAIII stimulate the translation of *hla* mRNA, which is then presumably responsible for the increase in mRNA levels (Morfeldt et al. 1995). Whether RNase III cleaved *hla* mRNA bound to RNAIII to generate a shorter 5' UTR of *hla* mRNA has not been analyzed. The enhanced *hla* mRNA levels are also partially explained by RNAIII-dependent repression of the *rot* mRNA, which encodes a transcriptional repressor of toxin genes, such as *hla* (Geisinger et al. 2006; Boisset et al. 2007).

In this way, the quorum-sensing-dependent RNAIII is involved in a particular regulatory network motif, a double feed-forward loop that behaves as a double selector switch to ensure fine-tuned coordination of the inverse expression of two sets of genes (adhesins and exotoxins), tight regulation and filtering of noisy signals (Nitzan et al. 2015).

#### **Modulation of mRNA degradation without affecting translation**

In the Enterobacteria, the MicC sRNA binds the coding sequence of the *ompD* mRNA and directly promotes its degradation. It has been proposed that Hfq not only promotes the interaction between MicC and *ompD*, but also recruits RNase E to this site to activate mRNA degradation, without inhibiting translation (Pfeiffer et al. 2009). It has been further proposed that a 5' monophosphate group on the MicC sRNA can stimulate cleavage of *ompD* mRNA by RNase E (Bandyra et al. 2012). Such a di-

rect role of sRNA on mRNA stability, without interfering with translation cannot be excluded in Gram-positive bacteria (Fig. 4C and D). Indeed, five regulatory RNAs (three sRNAs, FasX, *IrvA*, Psm-mec and two antisense RNAs, *RatA* and *SR4*) are able to influence mRNA degradation without affecting translation (see below). Different players are presumably involved in these regulatory events, since RNase E is absent from most Firmicutes and Hfq seems to play a less important role in this type of regulation (Geisinger et al. 2006; Heidrich et al. 2006; Bohn, Rigoulay and Bouloc 2007; Boisset et al. 2007; Gaballa et al. 2008; Hammerle et al. 2014).

#### **Streptococcus pyogenes FasX sRNA**

The FasX sRNA was identified in *S. pyogenes* (Ramirez-Pena et al. 2010). One of the identified targets of FasX is the *ska* mRNA, encoding the virulence factor streptokinase. The *ska* mRNA is more stable in a wild-type strain than in a  $\Delta$ *fasX* strain. FasX forms basepairing interactions with the first 9 nts of the *ska* mRNA to create a stable RNA helix directly at the 5' end of the mRNA (Figs 4D and 5C). One prediction is that this conformation blocks access to both the 5'-3' exoribonuclease activity of RNase J1, which requires a single-stranded 5'-extension of at least 5 nts to gain access to the mRNA. RppH would also be predicted to be inhibited by this double-stranded conformation of the 5' end, assuming that it has a similar specificity (at least three single-stranded residues with G at position 2) to *B. subtilis* RppH (Piton et al. 2013). When FasX is absent, the model would predict that the 5' triphosphate of the *ska* mRNA, which starts with four consecutive G-residues, would be removed by *S. pyogenes* RppH and the mRNA would then be attacked by RNase J1. Although this model has not yet been tested, it has been shown that neither RNase Y nor PNPase is involved in this process (Ramirez-Pena et al. 2010) (Fig. 5C).

#### **Streptococcus mutans *irvA* regulatory RNA**

The *irvA* gene encodes a putative transcriptional regulator, originally thought to repress the so-called dextran-dependent aggregation (DDAG) stress response in *S. mutans*. However, a recent study has shown that, in fact, it is the 5' UTR of the *irvA* mRNA (and not the repressor encoded by the ORF) that is responsible for the DDAG minus phenotype observed in *irvA* deletion strains. The 5' UTR of the fully intact *irvA* mRNA behaves as a *trans*-acting regulatory RNA that stabilizes the *gbpC* mRNA about 10-fold (Fig. 5D). GbpC is the key surface-exposed lectin responsible for the DDAG+ phenotype under stress conditions. The 5' UTR of *irvA* interacts with the coding sequence of *gbpC* (about 110 nt downstream of the GbpC initiation codon) and blocks cleavage by RNase J2 at this site and apparently another about 100 nts further downstream (Liu et al. 2015). This system is a very nice example of a dual function messenger and regulatory RNA (reminiscent of *S. aureus* RNAIII, where the 3' UTR provides the regulatory function).

#### **Staphylococcus aureus Psm-mec sRNA**

The Psm-mec sRNA, like *irvA* and RNAIII, is a dual-function RNA in *S. aureus*. It encodes a cytolysin of the phenol-soluble modulins (PSM) and also acts as a regulatory RNA. Psm-mec mRNA binding causes a 2-fold decrease in the half-life of the *agrA* mRNA via RNase III cleavage (Fig. 4C; Kaito et al. 2013). Psm-mec mRNA also inhibits the synthesis of *AgrA* by binding to its coding sequence (more than 200 nts downstream of the start codon) but it is not clear whether this interaction is required to promote *agrA* degradation (Kaito et al. 2013). Depending on the strain of

*Staphylococcus*, the targets of the Psm-mec RNA appear to be different (Cheung et al. 2014).

#### *Bacillus subtilis* txpA/RatA

The type I toxin/antitoxin system txpA/RatA was identified in *B. subtilis*. The antitoxin is an asRNA, which inhibits the expression of the small toxic peptide TxpA expressed from the opposite strand (Silvaggi, Perkins and Losick 2005). The 3' ends of RatA and the txpA mRNA overlap by about 120 nts. Repression of TxpA toxin expression occurs via RNase III, which cleaves in the complementary region between txpA and RatA (Figs 4C and 5E). This cleavage is absolutely essential to silence TxpA expression. Indeed, an RNase III mutant is lethal due to the expression of this toxin and another called YonT, silenced using a similar mechanism (Durand, Gilet and Condon 2012b). In contrast to txpA, the half-life of RatA is not significantly affected by the RNase III deletion, which was surprising since this RNase cleaves generally both strands of the RNA duplex. This can be explained by the fact that RatA is synthesized in 15-fold excess over the txpA mRNA. In this context, most of RatA in the cell is not paired to txpA and follows a classical mRNA degradation pathway involving RNase Y cleavage around position +90, upstream of the base-pairing region with txpA. After this cleavage, the downstream product is degraded by RNase J1 and the upstream product of RatA cleavage is attacked by PNPase (Durand, Gilet and Condon 2012b). This example is one of the best characterized system in terms of the RNases involved in mRNA degradation mediated by small asRNA.

#### *Bacillus subtilis* bsrG/SR4

BsrG/SR4 is a temperature-dependent type I toxin/antitoxin system, similar to txpA/RatA, also identified in *B. subtilis* (Jahn et al. 2012). The bsrG mRNA is 294 nts in length and encodes a toxic hydrophobic peptide of 38 amino acids. Its expression is regulated by an asRNA (SR4) complementary to its 3'-end. Like txpA/RatA, the bsrG/SR4 duplex is cleaved by RNase III (Figs 4C and 5F). The half-life of both RNAs also depends on RNase R and RNase Y, presumably when they are not basepaired with each other.

The basepairing interactions between SR4 and bsrG lead to a secondary structure change of the bsrG mRNA around the RBS (Fig. 5F). This rearrangement extends the stem-loop structure encompassing the SD sequence from 4 to 8 nts, inhibiting bsrG translation (Jahn and Brantl 2013). It is not known whether this translation inhibition is necessary to allow RNase III cleavage further downstream.

#### Modulation of translation without affecting mRNA degradation

It should be noted that it is also possible for sRNAs to modulate translation without affecting mRNA degradation, as exemplified by the *S. aureus* SprA1/SprA1as type I toxin/antitoxin system. To avoid the toxicity of the peptide expressed from SprA1 during *S. aureus* growth, the stable sprA1 mRNA is repressed by high amounts of the unstable antitoxin SprA1as (Sayed et al. 2012). The ribonucleases involved in the degradation pathway of SprA1as are not known. Contrary to most type I antitoxin/toxin systems, SprA1as prevents translation of sprA1 through imperfect basepairing interactions explaining why RNase III does not cleave the SprA1-SprA1as duplex (Fig. 5G). Surprisingly, the functional domain of the asRNA does not involve the Rho-independent terminator hairpin of SprA1as, which is fully complementary to the 3' end of sprA1 (Sayed, Jousselin and Felden 2011). Instead, the binding region is located in its 5' part that is

partially complementary to the RBS of sprA1 (Fig. 5C). Overproduction of SprA1as has no effect on sprA1 mRNA levels, suggesting that this system functions solely at the translational level.

## CONCLUDING REMARKS

The different studies presented here show that regulatory RNAs in several Gram-positive bacteria have an important role in modulating RNase access to specific mRNAs to target them for degradation as a function of growth conditions, as shown in *Enterobacteriaceae*. However, modulation of mRNA degradation by sRNA and the RNases involved in these processes are still largely unknown in Gram-positive bacteria and need to be further explored. RNase III seems to play an important role in regulating degradation via antisense RNA in several type I toxin/antitoxin systems and in various sRNA-mediated regulatory events in both *S. aureus* (RNAIII, RsaE, psm-mec) and *B. subtilis* (RoxS). Nevertheless, studies of other sRNAs (FasX, VR-RNA, IrvA) suggest that RNases Y, J1 and even J2 may also be important players in mRNA degradation promoted by sRNA. Because these RNases are thought to be able to form diverse complexes, this leads to the question of how these complexes might be involved in sRNA-mediated mRNA degradation *in vivo*. Furthermore, the composition of these RNA degrading complexes might be different depending on growth conditions, as has been observed in *E. coli*.

Hfq plays a central role in RNA-mediated regulation in *E. coli* by stabilizing sRNAs, by facilitating basepairing interactions with their targets and by stimulating mRNA degradation by RNase E (Bandyra et al. 2012). A similar mechanism could exist in Gram-positive bacteria but most of them have RNase Y in place of RNase E (Fig. 1). LhrA in *L. monocytogenes* is currently the sole sRNA from a Gram-positive organism requiring Hfq to stimulate basepairing interactions with its target *imo0850* (Christiansen et al. 2004). While other examples will undoubtedly be uncovered, especially in *Listeria* and *Clostridium* where Hfq appears to play a more important role than in *S. aureus* or *B. subtilis*, these observations make it unlikely that Hfq is the major factor mediating mRNA degradation by sRNA through RNase Y throughout the Firmicutes. It is interesting in this regard that Hfq from *E. coli* can be replaced by Hfq from either *L. monocytogenes* or *C. difficile* (Caillet et al. 2014). Both of these organisms have an RNase E/G homolog. In contrast, *S. aureus* and *Borrelia burgdorferi* do not have RNase E and their Hfq proteins are unable to fully complement a  $\Delta hfq$  strain of *E. coli* (Vecerek et al. 2008; Rochat et al. 2012). Thus, it could be interesting to determine whether there is any correlation between the interchangeability of Hfq and the degradation machinery present in a particular organism.

*Bacillus subtilis* has numerous annotated RNA-binding proteins of which at least 14 have unknown functions. Moreover, the group of J. Helmann has shown that the small basic proteins FbpA, B and C play a role in regulation by the sRNA FsrA, suggesting that other RNA chaperones can substitute for Hfq in some Firmicutes. These proteins could also help to modulate RNase activity on target mRNAs as has been suggested for FbpB in *B. subtilis* (Smaldone et al. 2012). In *S. aureus*, several RNA-binding proteins such as the RNA helicase CshA (Oun et al. 2013) and the transcriptional regulatory factor SarA (Morrison et al. 2012) have been shown to play major regulatory roles in RNA metabolism, although their mechanisms of action are not yet fully understood. A conserved RNA-binding protein YbeY, which contains structural domain similar to the Agonate protein, has been shown to regulate the accumulation of Hfq-dependent and -independent sRNAs and the target mRNAs in *Sinorhizobium*

meliloti (Pandey et al. 2014). Although this protein is conserved in Firmicutes, its proposed function as an RNase (or RNA chaperone) has not been studied in these bacteria. Approaches to fractionate and identify the various classes of sRNP might also provide some clues on the implication of the RNases and RNA chaperones in sRNA-dependent pathways.

Many sRNAs, among them FsrA, RsaE and FasX, have C-rich regions (CRRs) predicted to interact with G-rich sequences such as SD elements in their mRNA targets in bacteria. At least 3 sRNAs in *B. subtilis* and 11 in *S. aureus* have these CRRs suggesting that these sRNAs belong to a family of regulatory RNAs. The CRR, in addition to representing a potential seed sequence for the interaction with mRNA targets, could represent a binding site for specific proteins and could play a role in sRNA stability, as it was seen for CRR found in the 3' UTR of some mRNAs in eukaryotes (Makeyev and Liebhaber 2002). Indeed, unpublished results show that mutations in the some of the C-rich domains of *B. subtilis* RsaE/RoxS render it more unstable than the wild-type sequence, even though the secondary structure is predicted to be unaffected by these mutations (Durand et al. 2015). More recently, a C-rich sequence motif in the 3' UTR of *icaR* mRNA was shown to affect the translation of its own mRNA most probably by favoring a 5'-3' UTR interaction and by recruiting RNase III to induce rapid degradation of the mRNA (Ruiz de los Mozos et al. 2013).

Finally, in addition to non-coding RNA, the modulation of RNase activity or stability can also be achieved by a direct binding of particular proteins to RNases, e.g. RraA to RNase E (Lee et al. 2003; Zhao et al. 2006), by RNase post-translational modification, e.g. acetylation of RNase R (Liang, Malhotra and Deutscher 2011; Liang and Deutscher 2012) or by modification of the 5' end of *E. coli* and *Streptomyces* RNAs, e.g. modification by NAD (Chen et al. 2009; Cahova et al. 2014). The potential for these types of regulation must also be considered in Gram-positive bacteria even if no specific examples have been characterized to date.

## FUNDING

This work was supported by funds from the CNRS (UPR 9073, UPR 9002), Université Paris VII-Denis Diderot (CC), Université de Strasbourg (PR), the Agence Nationale de la Recherche. This work has been published under the framework of two LABEX programs: ANR-Dynamo (CC) and ANR-10-LABX-0036.NETRNA (PR) that benefit from a funding from the state managed by the French National Research Agency as part of the Investments for the future program.

**Conflict of interest.** None declared.

## REFERENCES

- Ait-Bara S, Carpousis AJ, Quentin Y. RNase E in the gamma-Proteobacteria: conservation of intrinsically disordered non-catalytic region and molecular evolution of microdomains. *Mol Genet Genomics* 2014, DOI: 10.1007/s00438-014-0959-5.
- Andrade JM, Pobre V, Matos AM, et al. The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. *RNA* 2012;18:844-55.
- Bandyra KJ, Said N, Pfeiffer V, et al. The seed region of a small RNA drives the controlled destruction of the target mRNA by the endoribonuclease RNase E. *Mol Cell* 2012;47:943-53.
- Bechhofer DH, Wang W. Decay of ermC mRNA in a polynucleotide phosphorylase mutant of *Bacillus subtilis*. *J Bacteriol* 1998;180:5968-77.
- Bechhofer DH, Zen KH. Mechanism of erythromycin-induced ermC mRNA stability in *Bacillus subtilis*. *J Bacteriol* 1989;171:5803-11.
- Bohn C, Rigoulay C, Bouloc P. No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*. *BMC Microbiol* 2007;7:1-10.
- Bohn C, Rigoulay C, Chabelskaya S, et al. Experimental discovery of small RNAs in *Staphylococcus aureus* reveals a riboregulator of central metabolism. *Nucleic Acids Res* 2010;38:6620-36.
- Boisset S, Geissmann T, Huntzinger E, et al. *Staphylococcus aureus* RNAPIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Gene Dev* 2007;21:1353-66.
- Brantl S, Bruckner R. Small regulatory RNAs from low-GC Gram-positive bacteria. *RNA Biol* 2014;11:443-56.
- Bugrysheva JV, Scott JR. The ribonucleases J1 and J2 are essential for growth and have independent roles in mRNA decay in *Streptococcus pyogenes*. *Mol. Microbiol* 2010;75:731-43.
- Cahova H, Winz ML, Hofer K, et al. NAD capture Seq indicates NAD as a bacterial cap for a subset of regulatory RNAs. *Nature* 2014, DOI: 10.1038/nature14020.
- Caillet J, Gracia C, Fontaine F, et al. Clostridium difficile Hfq can replace *Escherichia coli* Hfq for most of its function. *RNA* 2014;20:1567-78.
- Campos-Guillen J, Bralley P, Jones GH, et al. Addition of poly(A) and heteropolymeric 3' ends in *Bacillus subtilis* wild-type and polynucleotide phosphorylase-deficient strains. *J Bacteriol* 2005;187:4698-706.
- Cao GJ, Sarkar N. Poly(A) RNA in *Bacillus subtilis*: identification of the polyadenylation site of flagellin mRNA. *FEMS Microbiol Lett* 1993;108:281-5.
- Chen X, Liu N, Khajotia S, et al. RNases J1 and J2 are critical pleiotropic regulators in *Streptococcus mutans*. *Microbiology* 2015, in press. DOI: 10.1099/mic.0.000039.
- Chen YG, Kowtoniuk WE, Agarwal I, et al. LC/MS analysis of cellular RNA reveals NAD-linked RNA. *Nat Chem Biol* 2009;5:879-81.
- Chen Z, Itzek A, Malke H, et al. Multiple roles of RNase Y in *Streptococcus pyogenes* mRNA processing and degradation. *J Bacteriol* 2013;195:2585-94.
- Cheung GY, Villaruz AE, Joo HS, et al. Genome-wide analysis of the regulatory function mediated by the small regulatory psm-mec RNA of methicillin-resistant *Staphylococcus aureus*. *Int J Med Microbiol* 2014;304:637-44.
- Christiansen JK, Larsen MH, Ingmer H, et al. The RNA-binding protein Hfq of *Listeria monocytogenes*: role in stress tolerance and virulence. *J Bacteriol* 2004;186:3355-62.
- Commichau FM, Rothe FM, Herzberg C, et al. Novel activities of glycolytic enzymes in *Bacillus subtilis*: interactions with essential proteins involved in mRNA processing. *Mol Cell Proteomics* 2009;8:1350-60.
- Condon C, Bechhofer DH. Regulated RNA stability in the Gram positives. *Curr Opin Microbiol* 2011;14:148-54.
- Condon C, Putzer H. The phylogenetic distribution of bacterial ribonucleases. *Nucleic Acids Res* 2002;30:5339-46.
- Dambach M, Irnov I, Winkler WC. Association of RNAs with *Bacillus subtilis* Hfq. *PLoS One* 2013;8:e55156.
- De Lay N, Gottesman S. Role of polynucleotide phosphorylase in sRNA function in *Escherichia coli*. *RNA* 2011;17:1172-89.
- De Lay N, Schu DJ, Gottesman S. Bacterial small RNA-based negative regulation: Hfq and its accomplices. *J Biol Chem* 2013;288:7996-8003.



- Deana A, Celesnik H, Belasco J. The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature* 2008;**451**:355–8.
- Deikus G, Bechhofer DH. Initiation of decay of *Bacillus subtilis* trp leader RNA. *J Biol Chem* 2007;**282**:20238–44.
- Deikus G, Bechhofer DH. *Bacillus subtilis* trp Leader RNA: RNase J1 endonuclease cleavage specificity and PNPase processing. *J Biol Chem* 2009;**284**:26394–401.
- Deltcheva E, Chylinski K, Sharma CM, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 2011;**471**:602–7.
- Deutscher MP, Reuven NB. Enzymatic basis for hydrolytic versus phosphorolytic mRNA degradation in *Escherichia coli* and *Bacillus subtilis*. *P Natl Acad Sci USA* 1991;**88**:3277–80.
- Drider D, Condon C. The continuing story of endoribonuclease III. *J Mol Microb Biotech* 2004;**8**:195–200.
- Durand S, Braun F, Lioliou E, et al. A nitric oxide regulated small RNA controls expression of genes involved in redox homeostasis in *Bacillus subtilis*. *PLoS Genet* 2015;**11**:e1004957.
- Durand S, Gilet L, Bessieres P, et al. Three essential ribonucleases-RNase Y, J1, and III-control the abundance of a majority of *Bacillus subtilis* mRNAs. *PLoS Genet* 2012a;**8**:e1002520.
- Durand S, Gilet L, Condon C. The essential function of *B. subtilis* RNase III is to silence foreign toxin genes. *PLoS Genet* 2012b;**8**:e1003181.
- Figaro S, Durand S, Gilet L, et al. *Bacillus subtilis* mutants with knockouts of the genes encoding ribonucleases RNase Y and RNase J1 are viable, with major defects in cell morphology, sporulation, and competence. *J Bacteriol* 2013;**195**:2340–8.
- Gaballa A, Antelmann H, Aguilar C, et al. The *Bacillus subtilis* iron-sparing response is mediated by a Fur-regulated small RNA and three small, basic proteins. *P Natl Acad Sci USA* 2008;**105**:11927–32.
- Gatewood ML, Bralley P, Weil MR, et al. RNA-Seq and RNA immunoprecipitation analyses of the transcriptome of *Streptomyces coelicolor* identify substrates for RNase III. *J Bacteriol* 2012;**194**:2228–37.
- Geisinger E, Adhikari RP, Jin R, et al. Inhibition of rot translation by RNase III, a key feature of agr function. *Mol Microbiol* 2006;**61**:1038–48.
- Geissmann T, Marzi S, Romby P. A search for small noncoding RNAs in *Staphylococcus aureus* reveals a conserved sequence motif for regulation. *Nucleic Acids Res* 2009;**37**:7239–59.
- Gilet L, DiChiara JM, Figaro S, et al. Small stable RNA maturation and turnover in *Bacillus subtilis*. *Mol Microbiol* 2015;**95**:270–82.
- Grosjean H, Breton M, Sirand-Pugnet P, et al. Predicting the minimal translation apparatus: lessons from the reductive evolution of mollicutes. *PLoS Genet* 2014;**10**:e1004363.
- Hammerle H, Amman F, Vecerek B, et al. Impact of Hfq on the *Bacillus subtilis* transcriptome. *PLoS One* 2014;**9**:e98661.
- Hardwick SW, Chan VS, Broadhurst RW, et al. An RNA degradationosome assembly in *Caulobacter crescentus*. *Nucleic Acids Res* 2011;**39**:1449–59.
- Heidrich N, Chinali A, Gerth U, et al. The small untranslated RNA SR1 from the *Bacillus subtilis* genome is involved in the regulation of arginine catabolism. *Mol Microbiol* 2006;**62**:520–36.
- Heinemann IU, Soll D, Randau L. Transfer RNA processing in archaea: unusual pathways and enzymes. *FEBS Lett* 2010;**584**:303–9.
- Hsieh PK, Richards J, Liu Q, et al. Specificity of RppH-dependent RNA degradation in *Bacillus subtilis*. *P Natl Acad Sci USA* 2013;**110**:8864–9.
- Jahn N, Brantl S. One antitoxin—two functions: SR4 controls toxin mRNA decay and translation. *Nucleic Acids Res* 2013;**41**:9870–80.
- Jahn N, Preis H, Wiedemann C, et al. BsrG/SR4 from *Bacillus subtilis*—the first temperature-dependent type I toxin-antitoxin system. *Mol Microbiol* 2012;**83**:579–98.
- Jaskiewicz L, Filipowicz W. Role of Dicer in posttranscriptional RNA silencing. *Curr Top Microbiol Immunol* 2008;**320**:77–97.
- Jester BC, Romby P, Lioliou E. When ribonucleases come into play in pathogens: a survey of gram-positive bacteria. *Int J Microbiol* 2012;**2012**:592196.
- Jones SE, Leong V, Ortega J, et al. Development, antibiotic production, and ribosome assembly in *Streptomyces venezuelae* are impacted by RNase J and RNase III deletion. *J Bacteriol* 2014;**196**:4253–67.
- Kaito C, Kurokawa K, Matsumoto Y, et al. Silkworm pathogenic bacteria infection model for identification of novel virulence genes. *Mol Microbiol* 2005;**56**:934–44.
- Kaito C, Saito Y, Ikuo M, et al. Mobile genetic element SCCmec-encoded psm-mec RNA suppresses translation of agrA and attenuates MRSA virulence. *PLoS Pathog* 2013;**9**:e1003269.
- Kamathapu V, Nudler E. Rethinking transcription coupled DNA repair. *Curr Opin Microbiol* 2015;**24**:15–20.
- Kang SO, Caparon MG, Cho KH. Virulence gene regulation by CvfA, a putative RNase: the CvfA-enolase complex in *Streptococcus pyogenes* links nutritional stress, growth-phase control, and virulence gene expression. *Infect Immun* 2010;**78**:2754–67.
- Khemic V, Toesca I, Poljak L, et al. The RNase E of *Escherichia coli* has at least two binding sites for DEAD-box RNA helicases: functional replacement of RhlB by RhlE. *Mol Microbiol* 2004;**54**:1422–30.
- Kovacs L, Csanadi A, Megyeri K, et al. Mycobacterial RNase E-associated proteins. *Microbiol Immunol* 2005;**49**:1003–7.
- Lalaouna D, Simoneau-Roy M, Lafontaine D, et al. Regulatory RNAs and target mRNA decay in prokaryotes. *Biochim Biophys Acta* 2013;**1829**:742–7.
- Lasa I, Toledo-Arana A, Dobin A, et al. Genome-wide antisense transcription drives mRNA processing in bacteria. *P Natl Acad Sci USA* 2011;**108**:20172–7.
- Lasa I, Toledo-Arana A, Gingeras TR. An effort to make sense of antisense transcription in bacteria. *RNA Biol* 2012;**9**:1039–44.
- Lee JH, Gatewood ML, Jones GH. RNase III is required for actinomycin production in *Streptomyces antibioticus*. *Appl Environ Microb* 2013;**79**:6447–51.
- Lee K, Cohen SN. A *Streptomyces coelicolor* functional orthologue of *Escherichia coli* RNase E shows shuffling of catalytic and PNPase-binding domains. *Mol Microbiol* 2003;**48**:349–60.
- Lee K, Zhan X, Gao J, et al. RraA, a protein inhibitor of RNase E activity that globally modulates RNA abundance in *E. coli*. *Cell* 2003;**114**:623–34.
- Lehnik-Habrink M, Newman J, Rothe FM, et al. RNase Y in *Bacillus subtilis*: a natively disordered protein that is the functional equivalent of RNase E from *Escherichia coli*. *J Bacteriol* 2011;**193**:5431–41.
- Lehnik-Habrink M, Pfortner H, Rempeters L, et al. The RNA degradationosome in *Bacillus subtilis*: identification of CshA as the major RNA helicase in the multiprotein complex. *Mol Microbiol* 2010;**77**:958–71.
- Liang W, Deutscher MP. Post-translational modification of RNase R is regulated by stress-dependent reduction in the acetylating enzyme Pka (YfiQ). *RNA* 2012;**18**:37–41.

- Liang W, Malhotra A, Deutscher MP. Acetylation regulates the stability of a bacterial protein: growth stage-dependent modification of RNase R. *Mol Cell* 2011;44:160–6.
- Linder P, Lemelle S, Redder P. Transcriptome-wide analyses of 5'-ends in RNase J mutants of a gram-positive pathogen reveal a role in RNA maturation, regulation and degradation. *PLoS Genet* 2014;10:e1004207.
- Liolou E, Sharma CM, Caldeleri I, et al. Global regulatory functions of the *Staphylococcus aureus* endoribonuclease III in gene expression. *PLoS Genet* 2012;8:e1002782.
- Liu B, Deikus G, Bree A, et al. Global analysis of mRNA decay intermediates in *Bacillus subtilis* wild-type and polynucleotide phosphorylase-deletion strains. *Mol Microbiol* 2014;94:41–55.
- Liu G, Chatr KF, Chandra G, et al. Molecular regulation of antibiotic biosynthesis in streptomycetes. *Microbiol Mol Biol R* 2013;77:112–43.
- Liu N, Niu G, Xie Z, et al. The *Streptococcus mutans* *irvA* gene encodes a trans-acting riboregulatory mRNA. *Mol Cell* 2015;57:179–90.
- Luttinger A, Hahn J, Dubnau D. Polynucleotide phosphorylase is necessary for competence development in *Bacillus subtilis*. *Mol Microbiol* 1996;19:343–56.
- Lybecker M, Bilusic I, Raghavan R. Pervasive transcription: detecting functional RNAs in bacteria. *Transcription* 2014a;5:e944039.
- Lybecker M, Zimmermann B, Bilusic I, et al. The double-stranded transcriptome of *Escherichia coli*. *P Natl Acad Sci USA* 2014b;111:3134–9.
- Mader U, Zigl I, Kretschmer J, et al. mRNA processing by RNases J1 and J2 affects *Bacillus subtilis* gene expression on a global scale. *Mol Microbiol* 2008;70:183–96.
- Makeyev AV, Liebhaber SA. The poly(C)-binding proteins: a multiplicity of functions and a search for mechanisms. *RNA* 2002;8:265–78.
- Marincola G, Schäfer T, Behler J, et al. RNase Y of *Staphylococcus aureus* and its role in the activation of virulence gene. *Mol Microbiol* 2012;85:817–32.
- Mathy N, Benard L, Pellegrini O, et al. 5'-to-3' exonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA. *Cell* 2007;129:681–92.
- Mathy N, Hebert A, Mervelet P, et al. *Bacillus subtilis* ribonucleases J1 and J2 form a complex with altered enzyme behaviour. *Mol Microbiol* 2010;75:489–98.
- Morfeldt E, Taylor D, von Gabain A, et al. Activation of alpha-toxin translation in *Staphylococcus aureus* by the trans-encoded antisense RNA, RNAIII. *EMBO J* 1995;14:4569–77.
- Morita T, Aiba H. RNase E action at a distance: degradation of target mRNAs mediated by an Hfq-binding small RNA in bacteria. *Gene Dev* 2011;25:294–8.
- Morrison JM, Anderson KL, Beenken KE, et al. The staphylococcal accessory regulator, *SarA*, is an RNA-binding protein that modulates the mRNA turnover properties of late-exponential and stationary phase *Staphylococcus aureus* cells. *Front Cell Infect Microbiol* 2012;2:26.
- Morrison JM, Dunman PM. The modulation of *Staphylococcus aureus* mRNA turnover. *Future Microbiol* 2011;6:1141–50.
- Newman JA, Hewitt L, Rodrigues C, et al. Dissection of the network of interactions that links RNA processing with glycolysis in the *Bacillus subtilis* degradosome. *J Mol Biol* 2012;416:121–36.
- Nielsen JS, Lei LK, Ebersbach T, et al. Defining a role for Hfq in Gram-positive bacteria: evidence for Hfq-dependent antisense regulation in *Listeria monocytogenes*. *Nucleic Acids Res* 2010;38:907–19.
- Nitzan M, Fechter P, Peer A, et al. A defense-offense multi-layered regulatory switch in a pathogenic bacterium. *Nucleic Acids Res* 2015;43:1357–69.
- Novick RP, Ross HF, Projan SJ, et al. Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J* 1993;12:3967–75.
- Novick RP. Autoinduction and signal transduction in the regulation of staphylococcal virulence. *Mol Microbiol* 2003;48:1429–49.
- Numata S, Nagata M, Mao H, et al. CvfA protein and polynucleotide phosphorylase act in an opposing manner to regulate *Staphylococcus aureus* virulence. *J Biol Chem* 2014;289:8420–31.
- Obana N, Shirahama Y, Abe K, et al. Stabilization of *Clostridium perfringens* collagenase mRNA by VR-RNA-dependent cleavage in 5' leader sequence. *Mol Microbiol* 2010;77:1416–28.
- Olson PD, Kuechenmeister LJ, Anderson KL, et al. Small molecule inhibitors of *Staphylococcus aureus* RnpA alter cellular mRNA turnover, exhibit antimicrobial activity, and attenuate pathogenesis. *PLoS Pathog* 2011;7:e1001287.
- Oun S, Redder P, Didier JP, et al. The CshA DEAD-box RNA helicase is important for quorum sensing control in *Staphylococcus aureus*. *RNA Biol* 2013;10:157–65.
- Oussenko IA, Abe T, Ujiiie H, et al. Participation of 3'-to-5' exonucleases in the turnover of *Bacillus subtilis* mRNA. *J Bacteriol* 2005;187:2758–67.
- Pandey SP, Winkler JA, Li H, et al. Central role for RNase YbeY in Hfq-dependent and Hfq-independent small RNA regulation in bacteria. *BMC Genomics* 2014;15:121–37.
- Pfeiffer V, Papenfort K, Lucchini S, et al. Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol* 2009;16:840–6.
- Piton J, Larue V, Thillier Y, et al. *Bacillus subtilis* RNA deprotection enzyme RppH recognizes guanosine in the second position of its substrates. *P Natl Acad Sci USA* 2013;110:8858–63.
- Prevost K, Desnoyers G, Jacques JF, et al. Small RNA-induced mRNA degradation achieved through both translation block and activated cleavage. *Gene Dev* 2011;25:385–96.
- Prud'homme-Genereux A, Beran RK, Iost I, et al. Physical and functional interactions among RNase E, polynucleotide phosphorylase and the cold-shock protein, CsdA: evidence for a 'cold shock degradosome'. *Mol Microbiol* 2004;54:1409–21.
- Ramirez-Pena E, Trevino J, Liu Z, et al. The group A *Streptococcus* small regulatory RNA FasX enhances streptokinase activity by increasing the stability of the *ska* mRNA transcript. *Mol Microbiol* 2010;78:1332–47.
- Régnier P, Hajnsdorf E. The interplay of Hfq, poly(A) polymerase I and exonucleases at the 3' ends of RNAs resulting from Rho-independent termination: a tentative model. *RNA Biol* 2013;10:602–9.
- Rochat T, Bouloc P, Yang Q, et al. Lack of interchangeability of Hfq-like proteins. *Biochimie* 2012;94:1554–9.
- Roux CM, DeMuth JP, Dunman PM. Characterization of components of the *Staphylococcus aureus* mRNA degradosome holoenzyme-like complex. *J Bacteriol* 2011;193:5520–6.
- Ruiz de los Mozos I, Vergara-Irigaray M, Segura V, et al. Base pairing interaction between 5'- and 3'-UTRs controls *icaR* mRNA translation in *Staphylococcus aureus*. *PLoS Genet* 2013;9:e1004001.
- Sauer E. Structure and RNA-binding properties of the bacterial LSm protein Hfq. *RNA Biol* 2013;10:610–8.

- Sayed N, Jousselin A, Felden B. A cis-antisense RNA acts in trans in *Staphylococcus aureus* to control translation of a human cytolytic peptide. *Nat Struct Mol Biol* 2011;**19**:105–12.
- Sayed N, Nonin-Lecomte S, Réty S, et al. Functional and structural insights of a *Staphylococcus aureus* apoptotic-like membrane peptide from a toxin-antitoxin module. *J Biol Chem* 2012;**287**:43454–63.
- Sesto N, Touchon M, Andrade JM, et al. A PNPase dependent CRISPR System in *Listeria*. *PLoS Genet* 2014;**10**:e1004065.
- Shahbadian K, Jamali A, Zig L, et al. RNase Y, a novel endoribonuclease, initiates riboswitch turnover in *Bacillus subtilis*. *EMBO J* 2009;**28**:3523–33.
- Silvaggi JM, Perkins JB, Losick R. Small untranslated RNA anti-toxin in *Bacillus subtilis*. *J Bacteriol* 2005;**187**:6641–50.
- Smaldone GT, Antelmann H, Gaballa A, et al. The FsrA sRNA and FbpB protein mediate the iron-dependent induction of the *Bacillus subtilis* lutABC iron-sulfur-containing oxidases. *J Bacteriol* 2012;**194**:2586–93.
- Stead MB, Marshburn S, Mohanty BK, et al. Analysis of *Escherichia coli* RNase E and RNase III activity in vivo using tiling microarrays. *Nucleic Acids Res* 2011;**39**:3188–203.
- Strahl H, Turlan C, Khalid S, et al. Membrane recognition and dynamics of the RNA degradosome. *PLoS Genet* 2015;**11**:e1004961.
- Taverniti V, Forti F, Ghisotti D, et al. *Mycobacterium smegmatis* RNaseJ is a 5'-3' exo-/endoribonuclease and both RNaseJ and RNase E are involved in ribosomal RNA maturation. *Mol Microbiol* 2011;**82**:1260–76.
- Vecerek B, Rajkowitsch L, Sonnleitner E, et al. The C-terminal domain of *Escherichia coli* Hfq is required for regulation. *Nucleic Acids Res* 2008;**36**:133–43.
- Vogel J, Luisi BF. Hfq and its constellation of RNA. *Nat Rev Microbiol* 2011;**9**:578–89.
- Wagner EG. Cycling of RNAs by Hfq. *RNA Biol* 2013;**10**:619–26.
- Wang W, Bechhofer DH. Properties of a *Bacillus subtilis* polynucleotide phosphorylase deletion strain. *J Bacteriol* 1996;**178**:2375–82.
- Wurtzel O, Sesto N, Mellin JR, et al. Comparative transcriptomics of pathogenic and non-pathogenic *Listeria* species. *Mol Syst Biol* 2012;**8**:583.
- Zhao M, Zhou L, Kawarasaki Y, et al. Regulation of RraA, a protein inhibitor of RNase E-mediated RNA decay. *J Bacteriol* 2006;**188**:3257–63.

## NOUVELLE

## L'ARN non codant RsaA favorise la persistance et atténue la virulence de *Staphylococcus aureus*

Claire Lays<sup>1-5</sup>, Cédric Romilly<sup>6</sup>, Arnaud Tomasini<sup>6</sup>,  
Isabelle Caldelari<sup>6</sup>, Yvonne Benito<sup>1-5,7</sup>, Philippe Hammann<sup>8</sup>,  
Thomas Geissmann<sup>1-5</sup>, Sandrine Boisset<sup>1-5,7</sup>, Pascale Romby<sup>6</sup>,  
François Vandenesch<sup>1-5,7</sup>

> Commensale de la peau et des muqueuses, *Staphylococcus aureus* est surtout connue pour être une bactérie pathogène opportuniste de l'homme. Elle est responsable de nombreuses infections, qu'il s'agisse d'atteintes cutanées bénignes ou d'infections plus profondes mettant en jeu le pronostic vital. Le « succès » du développement d'une infection dépend essentiellement de l'expression coordonnée et séquentielle d'une multitude de facteurs de virulence et de gènes accessoires. Cette coordination est assurée par un réseau de régulation extrêmement complexe composé de facteurs protéiques mais également d'ARN régulateurs. Au cours de la dernière décennie et grâce à l'essor de la génomique, les ARN régulateurs ont pris une place prépondérante au sein des réseaux de régulation, aussi bien chez les procaryotes que chez les eucaryotes. Chez les bactéries, ces ARN régulateurs sont divisés en plusieurs classes : les ARNm agissant en *cis* (*riboswitches*), les ARN antisens codés sur le même locus génétique que leur ARN cible (ARNas), et les ARN issus de régions intergéniques (ARNs) [1]. Chez *S. aureus*, plusieurs études ont identifié un grand nombre d'ARN régulateurs codés sur le *core* génome et sur les éléments mobiles [2]. Malgré la progression de nos connaissances concernant ces ARN et la caractérisation du rôle de certains d'entre eux dans la régulation du métabolisme et de la virulence, l'élucidation de leurs

fonctions reste toujours un défi. En utilisant une combinaison d'approches *in vitro* et *in vivo*, nous avons élucidé le mécanisme d'action et la fonction de l'un de ces ARNs, appelé RsaA [3]. Nous avons montré que RsaA réprime la traduction de l'ARNm *mgrA* qui code pour un régulateur transcriptionnel global. Ce dernier induit des phénotypes qui ont un rôle clef lors du processus infectieux.

### RsaA contrôle *mgrA*, un régulateur majeur chez *S. aureus*

Le gène *rsaA* est situé au niveau d'une région intergénique [4]. RsaA est un ARN de 138 nucléotides très stable composé de 3 tiges-boucles et qui contient un motif conservé UCCC en simple brin. Cette séquence a été décrite précédemment dans divers ARN régulateurs comme étant appropriée pour engager l'interaction avec la séquence Shine et Dalgarno des ARNm cibles. RsaA est exprimé lors de la phase post exponentielle de croissance de la bactérie et s'accumule en phase stationnaire, sous le contrôle du facteur alternatif de transcription  $\sigma^B$ . Grâce à une analyse différentielle du protéome couplée à la recherche des interactions ARNm-RsaA, nous avons caractérisé une des principales cibles de RsaA, l'ARNm *mgrA*. La protéine correspondante est un régulateur pléiotrope majeur chez *S. aureus*, responsable de la régulation positive ou négative de plus de 100 gènes [5]. RsaA agit comme un ARN antisens en interagissant avec deux

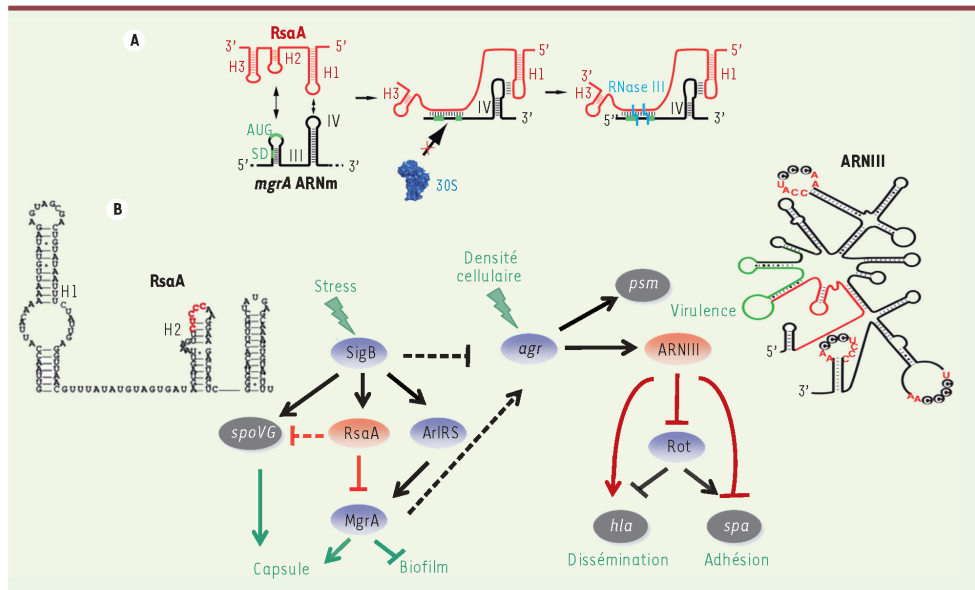
<sup>1</sup> CIRI, centre international de recherche en infectiologie, université de Lyon, Faculté de médecine-site Laennec, 7, rue Guillaume Paradin, 69372 Lyon, France ;  
<sup>2</sup> Inserm U1111, Lyon, France ;  
<sup>3</sup> École normale supérieure de Lyon, France ;  
<sup>4</sup> Université Lyon 1, Lyon, France ;  
<sup>5</sup> CNRS UMR5308, Lyon, France ;  
<sup>6</sup> architecture et réactivité de l'ARN, université de Strasbourg, CNRS, IBMC, Strasbourg, France ;  
<sup>7</sup> Hospices civils de Lyon, Lyon, France ;  
<sup>8</sup> plate-forme de protéomique, IBMC, Strasbourg, France.  
claire.lays@inserm.fr

régions de l'ARNm qui sont essentielles à la régulation : une interaction boucle-boucle implique les boucles apicales des motifs en tige-boucle HI de RsaA et IV de l'ARNm *mgrA* localisé dans la région codante ; la seconde interaction fait intervenir la séquence riche en C de RsaA avec la séquence Shine et Dalgarno de l'ARNm *mgrA* (Figure 1A). Cette seconde interaction empêche la fixation du ribosome pour inhiber la traduction de l'ARNm *mgrA*. Ce duplex est ensuite dégradé par la ribonucléase III (RNase III). Ce mécanisme d'action n'est pas sans rappeler celui de la régulation de l'ARNm *coa* par l'ARNIII, l'effecteur intracellulaire du système de densité cellulaire *agr* [6]. Par ailleurs, il est fort probable que RsaA prenne place dans le réseau de régulation qui répond à l'ARNIII (Figure 1B). En effet, le système *agr* est soumis à de nombreux signaux externes et intracellulaires. Par exemple, le facteur  $\sigma^B$  et MgrA auraient une action différente sur le système *agr*, permettant de relier la réponse au stress et à la virulence.

### RsaA atténue la virulence aiguë, mais facilite la formation d'un biofilm

Régulateur pléiotrope de nombreux gènes, MgrA est connu entre autres comme activateur de la synthèse de la capsule (structure facultative externe permettant à la bactérie de se protéger contre les attaques du système immunitaire) et comme répresseur de





**Figure 1. RsaA et ses réseaux de régulation.** A. Schéma résumant le mécanisme de régulation par RsaA. RsaA se fixe à l'ARNm *mgrA* et inhibe sa traduction en empêchant la fixation de la petite sous-unité 30S du ribosome, et recrute la RNase III pour induire la dégradation du duplex. B. RsaA est activé par  $\sigma^B$  et, en retour, réprime la traduction de MgrA. RsaA est ainsi indirectement relié au réseau de régulation de l'ARNIII, puisque MgrA active l'expression de l'opéron *agrACDB* alors que  $\sigma^B$  le réprime. Les flèches indiquent une activation et les barres une répression. En bleu sont indiqués les régulateurs protéiques, en rouge les ARN régulateurs et en gris les facteurs de virulence. Les lignes rouges correspondent aux régulations post-transcriptionnelles et les lignes noires aux régulations transcriptionnelles. Les événements de régulation potentiellement indirects sont montrés par les lignes en pointillés. Les structures secondaires de RsaA et de l'ARNIII sont présentées avec leur motifs riches en C qui ciblent les ARNm.

la formation de biofilm (structure tridimensionnelle multicellulaire assurant le maintien local de l'infection et participant à la lutte contre le système immunitaire et les conditions environnementales défavorables [7]). Grâce à l'utilisation de mutants pour RsaA et MgrA, nous avons démontré par différentes expériences *in vitro* que l'inactivation de *rsaA* réprime la formation du biofilm tout en activant la production de capsule via la régulation de la traduction de MgrA. Ces deux phénotypes sont étroitement liés au processus infectieux chez l'homme. En effet, le biofilm est fortement impliqué dans le maintien des infections chroniques (infection sur cathéter et sur biomatériaux implantés [8]) alors que la capsule est associée aux infections

aiguës (abcès, bactériémie). Comme l'expression de RsaA influence de manière opposée la synthèse de capsule et la formation de biofilm, nous avons testé l'impact de cet ARN sur les deux types d'infections (Figure 2). Pour cela, nous avons développé deux modèles d'infections chez la souris, le modèle d'infection sur cathéter (révélant le rôle du biofilm) et le modèle de bactériémie par injection intrapéritonéale (mettant en évidence le rôle de la capsule). Dans le premier modèle, un cathéter sous-cutané est infecté chez la souris par une souche exprimant ou non RsaA (Figure 2A). Nous avons suivi le développement de l'infection dans le cathéter (Figure 2A) et au niveau des tissus environnants (Figure 2B). Seules les bactéries exprimant RsaA se

sont développées au site d'infection ainsi que dans les tissus environnants, démontrant que RsaA est impliqué dans le développement et le maintien de l'infection chronique. Lors de la bactériémie expérimentale par injection intrapéritonéale, nous avons mis en évidence que l'expression de RsaA réduisait la persistance de la souche au site infectieux (Figure 2C) ainsi que sa dissémination sanguine (Figure 2D). Ainsi, il apparaît que RsaA influence la virulence de la bactérie en faveur d'une infection chronique tout en limitant la sévérité de l'infection aiguë. La forte conservation de cet ARN chez l'ensemble des espèces du genre *Staphylococcus* [9] nous incite à proposer que le gène codant pour RsaA pourrait avoir été positivement sélectionné au



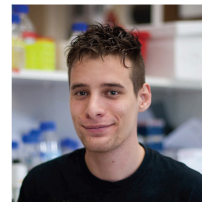
## Curriculum Vitae

### Arnaud Tomasini

26 ans, titulaire Permis B

1, rue des Colverts  
67100 Strasbourg  
06 79 39 69 61

[arnaud.tomasini@gmail.com](mailto:arnaud.tomasini@gmail.com)



## Doctorant en Biologie Moléculaire et Microbiologie

### FORMATION

- Depuis 2012 **Doctorant** en Biologie Moléculaire – Equipe du Dr. Pascale Romby, à l'IBMC, UPR 9002, Strasbourg, encadré par le Dr. Isabelle Caldelari – soutenance prévue au 16 septembre 2016  
Sujet : Recherche des cibles des ARN non codants de *Staphylococcus aureus* et de leurs réseaux de régulation : implication pour la virulence et la réponse au stress
- 2010-2012 **Master** Biologie des Microorganismes – mention Très Bien. Université de Strasbourg
- 2007-2010 **Licence** Biologie Cellulaire et Physiologie – mention Bien. Université de Strasbourg

### RECOMPENSES

- 2015 **Prix du projet innovant** reçu aux Doctoriales ® d'Alsace 2015
- 2012 **Financement de ma thèse** : classé 1<sup>er</sup>/45 au concours de l'École Doctorale pour l'obtention d'un contrat ministériel.

### EXPERIENCE PROFESSIONNELLE

- 2012 **Stage de Master 2** (6 mois) - Laboratoire de P. Romby à l'UPR 9002, IBMC, Strasbourg. Recherche des partenaires protéiques associés à un petit ARN non codant de *Staphylococcus aureus*. Encadrant : Isabelle Caldelari
- 2011 **Stage volontaire** (2 mois) – Laboratoire de P. Romby à l'UPR 9002, IBMC, Strasbourg. Etude de l'expression génétique par les ARN non codants de *Staphylococcus aureus*. Encadrant : Isabelle Caldelari.
- 2010 **Stage volontaire** (1 mois) – Laboratoire de R. Marquet & J.-C. Paillart à l'UPR 9002, IBMC, Strasbourg. Clonage de variants tétracystéine de la protéine APOBEC3G. Encadrant : Jean-Christophe Paillart.
- 2010 **Stage volontaire** (2 mois) – Laboratoire de P. Bertin, UMR 7516, Institut de Botanique, Strasbourg. Encadrant : Florence Ploetze. Métabolisme de l'arsenic chez les *Thiomonas spp.*

#### ➤ Enseignement

- 2012-2014 **Monitorat** en Travaux Pratiques et Dirigés, enseignement de la biologie moléculaire et microbiologie aux étudiants en 1<sup>ère</sup> année de Licence, Université de Strasbourg, responsable : Pascaline Ullmann

#### ➤ Formations Doctorales

- Depuis 2012 **Formations scientifiques et socio-professionnelles** : La publication scientifique : acteurs, circuits et enjeux ; ateliers d'écriture scientifique ; travail vocal et expression corporelle ; Sciences et société ; Connaissance scientifique et réflexion éthique ; participation aux Doctoriales ® d'Alsace 2015

### COMPETENCES TECHNIQUES

- Biologie** microbiologie moléculaire, physiologie bactérienne, ARN, protéines, interactions acide nucléiques-protéines, protéomes bactériens, transcription *in vitro*, traduction *in vitro*, clonages, PCR, manipulation de la radioactivité (toeprint, gelshift, filtration sur nitrocellulose), western blot, northern blot. Connaissances du fonctionnement des laboratoires, de la recherche académique.
- Langues** Anglais (lu écrit parlé). 2015 : TOEIC obtenu (score 975/990)
- Informatique** Microsoft Office, Photoshop, Illustrator, maîtrise des environnements MacOS et Windows, maîtrise outils web, environnement Wordpress.

### PUBLICATIONS ET COMMUNICATIONS

- Article **L'ARN non codant RsaA favorise la persistance et atténue la virulence**

- de *Staphylococcus aureus*.** (2014) *Med Sci* **30** 839-841  
Romilly C., Lays C., Tomasini A., Caldelari I., Benito Y., Hammann P., Geissmann T., Boisset S., Romby P. & Vandenesch F.
- A non-coding RNA promotes bacterial persistence and decreases virulence by regulating a regulator in *Staphylococcus aureus*.** (2014) *PLoS Pathog.* 10(3)  
Romilly C., Lays C., Tomasini A., Caldelari I., Benito Y., Hammann P., Geissmann T., Boisset S., Romby P. & Vandenesch F.
- Reuves **sRNA and mRNA turnover in Gram-positive bacteria.** (2015) *FEMS Microbiol Rev.*, in press.  
Durand S., Tomasini A., Braun F., Condon C., and Romby P.
- The importance of regulatory RNAs in *Staphylococcus aureus*** (2014) *Infection, Genetics and Evolution* **21** 616-626  
Tomasini A., François P., Howden B. P., Fechter P., Romby P. & Caldelari I.
- Présentations orales  
-Séminaire de Microbiologie de Strasbourg Strasbourg (France), mars 2016  
-Ecole Thématique de Microbiologie Moléculaire Carry-Le-Rouet (France), Octobre 2014  
-Séminaire interne de l'IBMC Strasbourg (France), Décembre 2012
- Posters  
-Labex Meeting Bischoffsheim (France), Novembre 2014  
-3rd Molecular Microbiology Meeting Würzburg (France), Mai 2014  
-Congrès Labex NetRNA Strasbourg (France), Avril 2014  
-Séminaire de Microbiologie de Strasbourg Strasbourg (France), Mars 2014  
-Congrès SFBM "SifrARN" Bischoffsheim (France), Novembre 2013

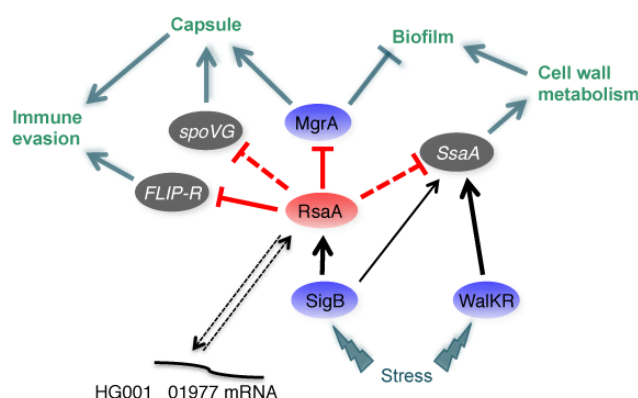
## ACTIVITES EXTRA-PROFESSIONNELLES

- **Activités lors de mon cursus universitaire**
- 2015 **Elu au Conseil Plénier du Collège Doctoral Européen de Strasbourg** – Représentant des doctorants domaine « Santé », Université de Strasbourg
- 2012 **Elu au Conseil Scientifique de l'Ecole Doctorale** – Représentant des doctorants, Université de Strasbourg
- 2010 **Intervenant linguistique au Centre de Ressources des Langues** – Institut de Botanique, Université de Strasbourg
- **Centres d'intérêt**
- Loisirs Cinéma, lecture, musique (saxophone), divertissement numérique.  
Depuis mai 2013 : rédacteur bénévole (actualités, interviews, critiques et dossiers) pour le site DC Planet ([www.dcplanet.fr](http://www.dcplanet.fr)) ; depuis mai 2015 : rédacteur en chef du site (gestion des plannings de publications, organisation d'interviews et conférence, réalisation émissions audios et reportages vidéo, relations associations culturelles et services de presse).
- Sports Course à pied, badminton, tennis



## Identification des cibles primaires des ARN non codant de *Staphylococcus aureus* et de leurs réseaux de régulation : mise au point des approches MAPS et Grad-seq

*S. aureus* est une bactérie pathogène opportuniste de l'homme qui pose un grave problème de santé publique dû à l'émergence et la dissémination de souches multi-résistantes aux antibiotiques. Le pouvoir pathogène de *S. aureus* est conféré par un très grand nombre de facteurs de virulence, dont l'expression est finement régulée à de multiples niveaux. Les effecteurs de cette régulation sont à la fois des protéines et des ARN non codants (ARNnc) aussi appelés ARN régulateurs. Une très grande diversité d'ARN régulateurs existe chez *S. aureus* mais je me suis concentré au cours de ma thèse sur la classe majoritaire, qui sont des ARNnc qui régulent la traduction d'ARNm en s'appariant directement avec leur cible au niveau du site de liaison du ribosome. Ces ARN sont impliqués dans de complexes réseaux de régulation qui permettent de contrôler la physiologie de la cellule ainsi que sa virulence. Pour élargir nos connaissances de ces réseaux, j'ai développé deux approches méthodologiques, appelées MAPS et Grad-seq, que j'ai appliquées *in vivo* chez *S. aureus* en utilisant deux ARNnc comme modèles d'étude : RsaA et RsaC. L'application du MAPS a permis d'identifier de nouvelles cibles directes pour RsaA et des cibles potentielles pour RsaC. L'approche Grad-Seq est un outil puissant pour suivre globalement les ARNnc et leurs ARNm cibles *in vivo* mais demande encore des ajustements. J'ai également pu déterminer un rôle probable pour l'ARNnc RsaC dans la régulation de l'homéostasie oxydo-réductive de *S. aureus*, en lien avec la résistance au stress oxydatif et avec la persistance lors de l'internalisation de la bactérie par des ostéoblastes.



**des ARN non codant de *Staphylococcus aureus* et de leurs réseaux de régulation  
: mise au point des approches MAPS et  
Grad-seq**

## Résumé

*S. aureus* est une bactérie pathogène opportuniste de l'homme qui pose un grave problème de santé publique. Le pouvoir pathogène de *S. aureus* est conféré par un très grand nombre de facteurs de virulence, dont l'expression est finement régulée à de multiples niveaux. Les effecteurs de cette régulation sont à la fois des protéines et des ARN non codants (ARNnc) aussi appelés ARN régulateurs. Je me suis concentré au cours de ma thèse sur la classe majoritaire qui sont les ARNnc qui régulent la traduction d'ARNm. Ils sont impliqués dans de complexes réseaux de régulation qui permettent de contrôler la physiologie de la cellule ainsi que sa virulence. Pour élargir nos connaissances de ces réseaux, j'ai développé deux approches méthodologiques, appelées MAPS et Grad-seq, que j'ai appliquées *in vivo* chez *S. aureus* en utilisant RsaA et RsaC comme modèles. L'application du MAPS a permis d'identifier de nouvelles cibles directes pour RsaA et des cibles potentielles pour RsaC. L'approche Grad-Seq est un outil puissant mais demande encore des ajustements. J'ai également pu déterminer un rôle probable pour l'ARNnc RsaC dans la régulation de l'homéostasie oxydo-réductive de *S. aureus*, en lien avec la résistance au stress oxydatif et avec la persistance lors de l'internalisation par les ostéoblastes.

Mot-clés : staphylococcus aureus, ARN régulateurs, ARN non codants, MAPS

## Résumé en anglais

*S. aureus* is an opportunistic pathogen of the human species which can express a large array of virulence factors whose expression is under tight regulation at multiple levels. The regulation can be done by proteins and by particular molecules of RNA called non-coding RNA (ncRNA). I focused during my thesis on the main category of ncRNA in *S. aureus*, which are regulating the translation of mRNA. These ARNs are involved in complex regulatory networks, impacting the physiology of the bacterial cell and its virulence. To understand further these networks, I developed two methodological approaches *in vivo* in *S. aureus*, called MAPS and Grad-seq, which were applied using RsaA and RsaC as models of studies. MAPS allowed to find new direct targets of RsaA and plausible targets for RsaC. The Grad-Seq method showed to be a powerful tool but still needs refinements. I also could determine a possible role for RsaC in the regulation of oxydo-reductive homeostasis, in direct link with oxydative stress resistance and persistence during internalisation by osteoblasts.

Key-words : staphylococcus aureus, regulatory RNAs, non coding RNAs, MAPS